申报	系列：教学科研并重型
	专业：应用经济学
	职称：副教授

业绩成果材料

（申报人的业绩成果材料包括论文、科研项目、获奖以及其他成果等）

单　位（二级单位） 经济管理学院

姓　名 **赵纯凯**

材料核对人：　　单位盖章：

核对时间：

华南农业大学制

目　录

一、教学研究业绩

教学研究项目：关于校级教改项目的立项通知及有关佐证材料.....4

二、科研项目

1. 主持：关于国家自然科学基金青年项目的立项通知及有关佐证材料.....13

2. 主持：关于广东省哲学社会科学规划项目的立项通知及有关作证材料..... 17

3. 主持：关于农业农村部软科学项目的立项合同及有关作证材料.....20

三、论文、著作等

1. 检索证明.....46

2. 以第一作者发表本专业论文情况

2. 1. Rising housing prices and marriage delays in China: Evidence from the urban land transaction policy..... 53

2. 2. Is digital finance environmentally friendly in China? Evidence from shared-bike trips 70

2. 3. ICT in education can improve students' achievements in rural China: The role of parents, educators and authorities.....85

2. 4. School nutritious feeding and cognitive abilities of students in poverty: Evidence from the nutrition improvement program in China ..110

2. 5. 来自军营的印记———军人职业与老年人健康 123

2. 6. Military Service and Health Outcomes of the Elderly in China....140

2. 7. Understanding government pollution control actions and well-being:

Evidence from micro individuals..... 169
2. 8. The effect of digital finance on residents' happiness: The case of mobile payments in China..... 188
3. 以通讯作者发表本专业论文情况
3. 1. Official turnover and sustainable development in China..... 225
3. 2. Online markets, offline happiness: E-commerce development and subjective well-being in rural China..... 249
3. 3. More is less: Homeroom teachers' administrative duties and students' achievements in China..... 278
3. 4. The long-term effect of childhood left-behind experiences on entrepreneurial activities in adulthood: Empirical evidence from China... 295
3. 5. Can national forest city construction mitigate air pollution in China? Evidence from a quasi- natural experiment..... 313
3. 6. Did innovative city constructions reduce carbon emissions? A quasi- natural experiment in China..... 335
3. 7. Beyond the scenery: Understanding the impact of rural tourism development on household consumption in China..... 361
3. 8. Teachers' administrative positions and students' non-cognitive abilities in China: evidence from a quasi-natural experiment..... 401

【佐证材料切记与目录页所列页码对应，不要用图片格式的材料进行打印。】

教学研究业绩

附件1

华南农业大学2023年校级教改项目立项名单

序号	项目编号	项目类别	项目名称	项目负责人	项目负责人单位
1	JG2023001	重点项目	高阶思维视域下《种子生物学》“四维融合”混合式教学模式改革与实践	周玉亮	农学院
2	JG2023002	重点项目	新林科视域下森林培育学知识图谱课程建设与教学模式创新	邱权	林学与风景园林学院
3	JG2023003	重点项目	一流专业建设背景下风景园林规划设计课程教学改革研究	陈崇贤	林学与风景园林学院
4	JG2023004	重点项目	一流农科高校基础实验课程思政教学一体化改革与实践——以大学物理实验课程为例	劳媚媚	基础实验与实践训练中心
5	JG2023005	重点项目	双一流高校“科-教-思”融合培养学生高阶思维的实验教学改革研究——以遗传学实验为例	李楠	基础实验与实践训练中心
6	JG2023006	重点项目	工程结构设计软件课程的"融合+分层"教学	李文雄	水利与土木工程学院
7	JG2023007	重点项目	基于竞赛和创新方法提升大学生创新能力的实证研究	易欣	材料与能源学院
8	JG2023008	重点项目	基于创新创业能力培养的“赛教融合”《养羊学》课程改革与实践	柳广斌	动物科学学院
9	JG2023009	重点项目	人工智能赋能生态化大学英语混合教学改革研究	苏君	外国语学院
10	JG2023010	重点项目	基于创新能力培养的《电路实验》教学改革研究与实践	王建华	电子工程学院(人工智能学院)
11	JG2023011	重点项目	新时代高校思政课“问题链”教学模式研究 ——聚焦《习近平新时代中国特色社会主义思想概论》课程	何艳玲	马克思主义学院
12	JG2023012	重点项目	三产融合理论在《茶叶生物化学》中的应用	张钰乾	园艺学院
13	JG2023013	重点项目	工程认证背景下基于知识图谱的新工科一流专业建设和提升的研究与实践	王金凤	数学与信息学院、软件学院

14	JG2023014	重点项目	融合数据分析思维和学科交叉的《线性代数》课程教学创新与实践	张伟峰	数学与信息学院、软件学院
15	JG2023015	重点项目	基于创新能力培养的统计学专业数据分析类实验课程改革探索与实践	周燕	数学与信息学院、软件学院
16	JG2023016	招标项目	基于知识图谱的农业院校课程思政建设探索—以公共数学基础课为例	张娜	数学与信息学院、软件学院
17	JG2023017	重点项目	基于新兽医人才培养的《分子生物学》课程改革与实践探索	沈永义	兽医学院
18	JG2023018	重点项目	思政视角下《国际金融》课程教学改革与实践	周超	经济管理学院
19	JG2023019	重点项目	新农科背景下基于“数字标本”的智慧植保实验体系研究与实践	李云锋	植物保护学院
20	JG2023020	重点项目	低碳农业背景下基于一特两驱教学模式的环境土壤学课程体系构建	林庆祺	资源环境学院
21	JG2023021	重点项目	“双一流”建设视野下安全教育模式嵌入环境化学实验室课程体系的实践与优化	高婷	资源环境学院
22	JG2023022	重点项目	“三台协同，以美育人”——以课程教学为基础的高校舞蹈美育建设研究与实践	郑琳喆	艺术学院
23	JG2023023	重点项目	劳动教育融入设计类专业课程路径探索 ——以《纤维艺术造型设计》课程为例	林汉聪	艺术学院
24	JG2023024	招标项目	人工智能背景下的新文科艺术专业人才培养 研究与创新：以动画专业为例	王柯	艺术学院
25	JG2023025	重点项目	“新农科”背景下《植物学》课程教学创新改革研究与实践	白玫	生命科学学院
26	JG2023026	重点项目	基于真人图书馆的“就业创业+思政”融合课程体系改革研究——以法学专业为例	李玮舜	人文与法学学院
27	JG2023027	重点项目	生成式人工智能背景下公共管理类课程教学模式创新研究与实践	张小娟	公共管理学院
28	JG2023028	重点项目	基于课程思政融合及项目驱动教学创新方法的研究与实践	孔莲芳	工程学院
29	JG2023029	招标项目	工程教育认证背景下的教学过程评价体系研究——以《区块链原理与技术》课程为例	肖媚燕	数学与信息学院、软件学院

30	JG2023030	招标项目	面向乡村振兴的农业高校传媒类新文科人才培养研究	胡辉	艺术学院
31	JG2023031	重点项目	以教育数字化转型为契机的《林木遗传育种 学》混合式教学创新研究	李培	林学与风景园林学院
32	JG2023032	重点项目	国际合作办学背景下建成环境历史课的中外融合与创新	李晓雪	林学与风景园林学院
33	JG2023033	重点项目	“双碳”背景下《包装材料学》课程教学改革与实践	刘涛	食品学院
34	JG2023034	重点项目	构建《遗传学》课程思政教学体系的探索与实践	刘自强	农学院
35	JG2023035	重点项目	基于数据挖掘的土木工程专业人才培养方案优化研究	张巍	水利与土木工程学院
36	JG2023036	重点项目	以培养创新人才为导向的《分析化学实验》教学改革与实践	熊亚红	材料与能源学院
37	JG2023037	重点项目	课程思政引领的《药剂学实验》 线上线下混合式教学探究与实践	胡洋	材料与能源学院
38	JG2023038	重点项目	课程链模式下农林院校学生用外语讲中国农耕文化故事能力的培养	秦建华	外国语学院
39	JG2023039	重点项目	基于人工智能技术的大学物理课程思政教学研究与实践	徐初东	电子工程学院(人工智能学院)
40	JG2023040	重点项目	“大思政课”视域下高校思政课实践教学 协同育人创新模式研究与实践	林晓燕	马克思主义学院
41	JG2023041	重点项目	基于文本挖掘的课程思政元素挖掘研究--以应用统计学为例	张建桃	数学与信息学院、软件学院
42	JG2023042	重点项目	新农科背景下《动物解剖与组织胚胎学》课程思政改革创新	马勇江	兽医学院
43	JG2023043	重点项目	学科交叉融合的植物检疫学线上线下一体化教学实践	何晓芳	植物保护学院
44	JG2023044	重点项目	测绘地理信息类专业“学赛研”三位一体实践教学改革研究	刘惠明	资源环境学院
45	JG2023045	重点项目	新农科建设背景下高校图书馆开展中华优秀传统文化教育的探索与实践	黎小妮	图书馆

46	JG2023046	重点项目	大学生创新创业训练计划培育机制改革及育人实效提升研究——以华南农业大学为例	萧润正	党委学生工作部（党委研究生工
47	JG2023047	重点项目	以赋能乡村振兴为导向的一流动画专业优化提升建设的探索和实践	涂先智	艺术学院
48	JG2023048	重点项目	面向工程教育认证的农业院校车辆工程人才培养探索	郭嘉明	工程学院
49	JG2023049	一般项目	面向社会服务的风景园林规划设计类课程教学改革与实践	汤辉	林学与风景园林学院
50	JG2023050	一般项目	基于创新能力培养的《林木细胞生物学》课程教学方法改革与实践	刘林川	林学与风景园林学院
51	JG2023051	一般项目	乡村振兴背景下规划专业人才软技能培养体系研究	王凌	林学与风景园林学院
52	JG2023052	一般项目	“食品仪器分析”线上线下混合式一流课程的建设与实践	陈运娇	食品学院
53	JG2023053	一般项目	课程思政与“五育并举”融合发展的高校足球课程改革与实践研究	陈存志	体育教学研究部
54	JG2023054	一般项目	思政引导的体育课程内容体系改革与实践--以瑜伽课为例	何灵捷	体育教学研究部
55	JG2023055	一般项目	新农科背景下的《遗传学》教学改革与实践	汪文毅	农学院
56	JG2023056	一般项目	“双一流”背景下基于信息技术的实验室安全教育平台优化建设	陈志民	基础实验与实践训练中心
57	JG2023057	一般项目	理论力学多平台混合教学线上线下一体化设计与实践	刘新红	水利与土木工程学院
58	JG2023058	一般项目	《饲料学》线上线下一体化混合教学设计探究	朱勇文	动物科学学院
59	JG2023059	一般项目	新农科背景下基于“双元制”教学设计的《蚕桑概论》课程改革与实践	杨婉莹	动物科学学院
60	JG2023060	一般项目	多模态话语分析理论下《跨文化交际》课程中中国文化对外传播路径研究	杜龙鼎	外国语学院
61	JG2023061	一般项目	人工智能赋能英语语言技能课程教学创新实践研究	李飞武	外国语学院

62	JG2023062	一般项目	基于研究性教学的大学物理实验课程思政实践教学探索	杨小红	电子工程学院(人工智能学院)
63	JG2023063	一般项目	高校思政课问题式教学的探索研究 ——以“思想道德与法治”课程为例	刘智娴	马克思主义学院
64	JG2023064	一般项目	《果蔬栽培学》课程思政设计与实践	苏钻贤	园艺学院
65	JG2023065	一般项目	《园艺种质资源与分类》实验教学内容体系优化的研究与实践	张志珂	园艺学院
66	JG2023066	一般项目	“真人+数字人”的智能教学模式探索与实践--以《数据结构》课程为例	梁云	数学与信息学院、软件学院
67	JG2023067	一般项目	数字化背景下《大学生心理健康教育》课程思政建设与实践	严颖	数学与信息学院、软件学院
68	JG2023068	一般项目	基于创新型人才培养目标对《海洋环境化学》的教学改革研究	公晗	海洋学院
69	JG2023069	一般项目	“四微一体”课堂教学实践与应用-以“水产品质量安全控制”课程为例	周爱国	海洋学院
70	JG2023070	一般项目	数智时代《财务管理》课程教学改革研究	周小春	经济管理学院
71	JG2023071	一般项目	虚-实二维教学模式在环境监测实践课程中的应用	郑芊	资源环境学院
72	JG2023072	一般项目	能源动力类专业课程产教研协同育人模式探索与实践	魏国强	生物质工程研究院
73	JG2023073	一般项目	心理资本视域下心理健康课程改革探索	林媛	党委学生工作部(党委研究生工
74	JG2023074	一般项目	“轻咨询”融入高校心理健康教育课程的实践路径研究——以必修课《大学生心理健康教育》为例	刘桂娥	党委学生工作部(党委研究生工
75	JG2023075	一般项目	基于可持续创新人才培养的《首饰镶嵌工艺》教学改革与实践	潘子广	艺术学院
76	JG2023076	一般项目	“声动岭南”——音乐与表演专业中的新媒体技术创新教学项目研究与实践	冯逸章	艺术学院
77	JG2023077	一般项目	人工智能技术背景下影像艺术课程的内容改革：以乡村影像为例	程蔚新	艺术学院

78	JG2023078	一般项目	《植物学实验》“课程思政”的探索与实践	梁祥修	生命科学学院
79	JG2023079	一般项目	聚焦SCAU特色的遗传学课程思政建设的维度研究与实践	郑少燕	生命科学学院
80	JG2023080	一般项目	教育数字化转型背景下微课赋能古代文学教学应用研究	李桂芹	人文与法学学院
81	JG2023081	一般项目	领军人才培养目标下的《人类行为与社会环境》教学改革与实践	林诚彦	公共管理学院
82	JG2023082	一般项目	行政管理专业法学类课程思政教学改革研究与实践——以《行政法》为例	宋丽娟	公共管理学院
83	JG2023083	一般项目	融合“新工科”视野的工业设计专业--交互设计课程跨学科融合与实践创新	李莎莎	工程学院
84	JG2023084	一般项目	以知识图谱的构建赋能信号类课程的OBE理念教学模式探索	徐梅宣	电子工程学院(人工智能学院)
85	JG2023085	青年项目	基于工程教育认证的食品科学与工程一流专业人才生产实践教学体系重构	黎攀	食品学院
86	JG2023086	青年项目	基于“新工科”背景的《畜产食品工艺学》多模式教学探索与实践	艾民珉	食品学院
87	JG2023087	青年项目	传感检测类课程的混合式教学创新与实践	冯婉媚	电子工程学院(人工智能学院)
88	JG2023088	青年项目	优秀传统文化融入“马克思主义基本原理”课程的路径研究	刘海成	马克思主义学院
89	JG2023089	青年项目	案例教学法在《习近平新时代中国特色社会主义思想概论》课程的创新与运用	朱露	马克思主义学院
90	JG2023090	青年项目	“大历史观”引领《中国近现代史纲要》研究与实践	李征	马克思主义学院
91	JG2023091	青年项目	基于“SPOC+Seminar”的《形势与政策》教学模式改革与实践	聂锐	马克思主义学院
92	JG2023092	青年项目	数字化兽医病例库的构建及其在OBE教学模式下《兽医临床诊断学》教学中的探索与实践	廖建昭	兽医学院
93	**JG2023093**	**青年项目**	**新文科背景下本科生《产业经济学》课程体系建设与教学改革路径优化**	**赵纯凯**	**经济管理学院**

94	JG2023094	青年项目	教育数字化背景下高校管理会计课程改革探析	张京心	经济管理学院
95	JG2023095	青年项目	高校史学理论教学的多元化模式研究——以《史学概论》为例	赖泽冰	人文与法学学院
96	JG2023096	自筹项目	树立文化自信的城市建设史与规划史课程思政探索与实践	王婷	林学与风景园林学院
97	JG2023097	自筹项目	深化产教融合校企合作的育人模式探索与实践--以华南农业大学为例	解加米	林学与风景园林学院
98	JG2023098	自筹项目	以学生为中心的城乡规划专业基础美术课程混合式教学模式探索与实践	吴宝娜	林学与风景园林学院
99	JG2023099	自筹项目	高校体育课程思政教学质量评价指标体系构建研究	姚业戴	体育教学研究部
100	JG2023100	自筹项目	基于翻转课堂的管理通识训练课程混合式教学模式的探索与实践	谢佳佳	基础实验与实践训练中心
101	JG2023101	自筹项目	探究式教学模式在绿色建筑设计与节能技术课程中的实践	崔艳琦	水利与土木工程学院
102	JG2023102	自筹项目	新农科背景下《基因工程原理》线上线下教学一体化教学设计与实践	孙媛	动物科学学院
103	JG2023103	自筹项目	语言智能时代非英语专业学生翻译能力提升路径创新研究	李志英	外国语学院
104	JG2023104	自筹项目	大学英语四级词汇教学改革与实践：比较文　化的视角	黄海翔	外国语学院
105	JG2023105	自筹项目	农业院校统计学专业创新创业教育的研究与实践	徐小红	数学与信息学院、软件学院
106	JG2023106	自筹项目	基于知识图谱的《计算机组成原理》课程智慧化教学研究与实践	杨磊	数学与信息学院、软件学院
107	JG2023107	自筹项目	管理类本科创新创业实践教学生态圈构建	毛小娟	数学与信息学院、软件学院
108	JG2023108	自筹项目	基于新农人创新能力培养的“四位一体”教学方法研究与实践	郭永龙	兽医学院
109	JG2023109	自筹项目	以价值引领、学科融合、全面育人为目标的公选课混合式教学模式探究——以《奇妙的动物》课程为例	曾芳	海洋学院

110	JG2023110	自筹项目	“百千万工程”背景下高校“三大课堂”融合培育爱农兴农人才的教学体系研究与实践	陈洋	经济管理学院
111	JG2023111	自筹项目	双一流建设背景下丁颖班本科人才培养机制深化研究	陈建平	植物保护学院
112	JG2023112	自筹项目	工程测量学实践、实验教学内容体系改革和整体优化	林观土	资源环境学院
113	JG2023113	自筹项目	基于虚拟仿真平台的《测量学》虚实融合的混合式实验教学改革	于红波	资源环境学院
114	JG2023114	自筹项目	多元信息素养教育中线上线下教学一体化研究	刘汉忠	图书馆
115	JG2023115	自筹项目	学科交叉在生物学中的教学实践—以课程《细胞的语言》建设为研究案例	胡宇飞	生命科学学院
116	JG2023116	自筹项目	历史建筑档案在文化遗产研学中的教育教学功能开发策略研究	谢君	文博馆(档案馆、华南农业博物
117	JG2023117	自筹项目	教与评：新文科背景下思政教育融入“中国史学史”课堂教学的探索与实践	翟麦玲	人文与法学学院
118	JG2023118	自筹项目	课程思政背景下国际法专业教学模式创新与实践研究	官隆清	人文与法学学院
119	JG2023119	自筹项目	后疫情时代《大学生心理健康教育》课程教学质量提升研究	石如彬	公共管理学院
120	JG2023120	自筹项目	新文科背景下本科生就业能力评价及提升策略研究——以土地资源管理专业为例	朱庆莹	公共管理学院
121	JG2023121	自筹项目	基于工程实践的工程热力学与传热学课程混合教学设计研究与实践	张烨	工程学院
122	JG2023122	自筹项目	融合AIGC建设工业设计创新性实验课程的研究与实践	曾志雄	工程学院

科研项目

国家自然科学基金资助项目批准通知

（包干制项目）

赵纯凯　先生/女士：

根据《国家自然科学基金条例》、相关项目管理办法规定和专家评审意见，国家自然科学基金委员会（以下简称自然科学基金委）决定资助您申请的项目。项目批准号：72304102，项目名称：从军经历下的老年人福利研究：特征事实、作用机理与保障策略，资助经费：30.00万元，项目起止年月：2024年01月至 2026年 12月，有关项目的评审意见及修改意见附后。

请您尽快登录科学基金网络信息系统（https://grants.nsfc.gov.cn），**认真阅读《国家自然科学基金资助项目计划书填报说明》并按要求填写《国家自然科学基金资助项目计划书》（以下简称计划书）**。对于有修改意见的项目，请您按修改意见及时调整计划书相关内容；如您对修改意见有异议，须在电子版计划书报送截止日期前向相关科学处提出。

请您将电子版计划书通过科学基金网络信息系统（https://grants.nsfc.gov.cn）提交，由依托单位审核后提交至自然科学基金委。自然科学基金委审核未通过者，将退回的电子版计划书修改后再行提交；审核通过者，打印纸质版计划书（一式两份，双面打印）并在项目负责人承诺栏签字，由依托单位在承诺栏加盖依托单位公章，且将申请书纸质签字盖章页订在其中一份计划书之后，一并报送至自然科学基金委项目材料接收工作组。纸质版计划书应当保证与审核通过的电子版计划书内容一致。**自然科学基金委将对申请书纸质签字盖章页进行审核，对存在问题的，允许依托单位进行一次修改或补齐。**

向自然科学基金委提交电子版计划书、报送纸质版计划书并补交申请书纸质签字盖章页截止时间节点如下：

1. **2023年9月7日16点：**提交电子版计划书的截止时间；

2. **2023年9月14日16点：**提交修改后电子版计划书的截止时间；

3. **2023年9月21日：**报送纸质版计划书（一式两份，其中一份包含申请书纸质签字盖章页）的截止时间。

4. **2023年10月7日：**报送修改后的申请书纸质签字盖章页的截止时间。

请按照以上规定及时提交电子版计划书，并报送纸质版计划书和申请书纸质签字盖章页，逾期不报计划书或申请书纸质签字盖章页且未说明理由的，视为自动放弃接受资助；未按要求修改或逾期提交申请书纸质签字盖章页者，将视情况给予暂缓拨付经费等处理。

附件：项目评审意见及修改意见表

国家自然科学基金委员会

2023年8月24日

广东省哲学社会科学规划专项小组

粤社科规专通〔2023〕47号

广东省哲学社会科学规划2023年度学科共建项目立项通知书

赵纯凯同志：

经省哲学社会科学工作领导小组审批，您申报的广东省哲学社会科学规划2023年度学科共建项目《电商下乡对广东农村家庭创业的影响研究》获准立项，批准号：GD23XYJ68。学科共建项目单项经费2万元，由项目负责人所在单位自筹并分期划拨给项目负责人，第一次拨款1.4万元，预留经费0.6万元在项目鉴定结项后拨付。请于9月15日前在广东省哲学社会科学规划项目管理平台认真填写预算。

项目负责人填写预算后，本通知即成为具有约束效力的立项协议。项目负责人所在单位须承担保证责任。项目负责人及所在单位须了解和执行以下规定：

1.课题组须学习并遵守《广东省哲学社会科学规划项目管理办法》。对于出现违规行为的，省哲学社会科学规划专项小组依据《广东省哲学社会科学规划项目管理办法》的相关规定进行处理。

2.立项项目经费不再追加。课题组如不接受，省哲学社会科学规划专项小组将撤销该项目立项。而立项经费一经接受，课题组将不得以资助经费不足为由，擅自变更原设计的最终成果形式和内容。

3.项目研究过程中，如有变更项目负责人、延长完成时间、改变成果内容或形式、变更项目管理单位、变更或增补课题组成员、终止项目或撤销项目等重要事项，项目负责人或所在单位必须按要求填写《广东省哲学社会科学规划项目重要事项变更审批表》，并及时报省哲学社会科学规划专项小组审批。

4.项目研究的阶段性成果发表时，必须注明“广东省哲学社会科学规划项目”字样。

5.项目成果的鉴定结项由省哲学社会科学规划专项小组组织，实行匿名鉴定制度。项目鉴定结项所需材料及装印要求，详见“广东社科规划”网站。

6.成果鉴定等级分为优秀、良好、合格和不合格四个

等次。不合格者即未能通过结项，不予拨付预留经费。项目成果的最终鉴定等级均通过“广东社科规划”网站予以公布，并通知项目负责人所在单位。

7.为提高成果质量，项目的最终成果（专著、研究报告）必须经鉴定结项后方能出版。对于违反规定擅自出版的，视为项目负责人自行终止立项协议，该项目即时终止，省哲学社会科学规划专项小组将不再受理该项目的结项鉴定申请，并通知项目负责人所在单位不再拨付预留经费。

以上规定，项目负责人及所在单位应严格遵守。如有异议，可不接受资助（逾时未在广东省哲学社会科学规划项目管理平台填报预算视为放弃立项），立项协议自行废止。

地址：广州市天河北路618号广东社科中心B座928室

电话：(020)83825078

邮编：510635

广东省哲学社会科学规划专项小组

2023年9月5日

农业农村部2024年度软科学课题
拟立项名单

命题类型及序号		课题名称	主持人	承接单位
开放性命题	1	粮食主产区农业现代化空间格局、分区特征与差异优化路径研究	马　楠	中南民族大学
		以加快农业农村现代化更好推进中国式现代化建设研究	张　辉	农业农村部规划设计研究院
	2	统筹新型城镇化和乡村全面振兴研究	曾　晨	华中农业大学
	3	县域城乡融合发展引领乡村振兴的体制机制研究	高　强	南京林业大学
	4	农业新质生产力发展研究	林乐芬	南京农业大学
	5	行为社会政策视角下增强脱贫地区和脱贫群众内生发展动力的路径研究	李　鹏	武汉工程大学

规定性命题	6	我国多元化食物供给潜力及政策研究	生吉萍	中国人民大学
	7	集成配套推动大面积提高粮食单产路径研究	林文声	北京理工大学
	8	粮食产销区省际横向利益补偿机制研究	空缺	
	9	农资保供稳价应对机制研究	郭金秀	农业农村部农村经济研究中心
	10	防止返贫帮扶政策和农村低收入人口常态化帮扶政策衔接并轨研究	檀学文	中国社会科学院农村发展研究所
	11	调整优化巩固拓展脱贫攻坚成果同乡村振兴有效衔接考核评估指标体系研究	空缺	
	12	智慧农业发展研究	贾　敏	全国农业展览馆
	13	突破糖料蔗机收短板路径研究	空缺	
	14	构建现代农业经营体系发展农业适度规模经营问题研究	刘文勇	中国农业科学院农业经济与发展研究所
	15	完善乡村建设机制研究	谭小建	浙江省之江乡村振兴研究院
	16	提升乡村治理能力策略与路径研究	陈荣卓	华中师范大学
	17	农业产业科技的内涵和组织模式研究	季　晨	浙江大学

规定性命题	18	农户承包地依法自愿有偿退出机制研究	曲　颂	中国农业科学院农业经济与发展研究所
	19	以宅基地制度为基础的农民住房多元化保障机制研究	张孝宇	上海市农业科学院
		以宅基地制度为基础的农民住房多元化保障机制研究	杨璐璐	中国矿业大学
	20	新型农村集体经济发展的实践探索及前景展望研究	罗明忠	华南农业大学
		新型农村集体经济发展的实践探索及前景展望研究	钟　真	中国人民大学
	21	促进农村劳动力高质量充分就业研究	黄开兴	北京大学
	22	面向现代化的农业支持政策体系创新研究	王术坤	中国社会科学院农村发展研究所
自主性命题		大食物观背景下森林食品供给潜力与发展策略研究	孙中孝	中国农业大学
		粮食主产区环境成本测度及省际横向利益补偿机制研究	王嫚嫚	三峡大学
		不同粮食功能区域促农增收关键因素识别与有效举措研究	王振振	农业农村部农村经济研究中心
		中国式现代化背景下易地搬迁集中安置区转型发展的内在逻辑与实践路径研究	丁凡琳	南开大学

自主性命题	粮食安全视域下家庭农场综合竞争力提升路径研究	来晓东	农业农村部农村经济研究中心
	农业转移人口城市融入的策略选择与分类推进路径研究	赵清军	湖州学院
	农村青年婚恋困境的形成机制、社会风险与治理研究	徐依婷	湖北大学
	数字乡村建设的普惠效应、作用机理与优化路径研究	赵纯凯	华南农业大学
	数字技术赋能农产品溯源体系的转型驱动和价值增值研究	周雄勇	福州大学
	“小田并大田”的实现路径与提升策略研究	史鹏飞	农业农村部农村经济研究中心
	脱贫农户生计转型发展的内生动力及长效机制研究	任　立	湖北经济学院
	“双碳”目标下畜禽养殖场户农牧结合促进政策研究	王如玉	天津市农业科学院
	农村人口转变对农业农村发展影响研究	张　琛	中国社会科学院人口与劳动经济研究所

编号：04240063

政府购买服务合同

服务事项：数字乡村建设的普惠效应、作用机理与优化路径研究

购买主体：农业农村部政策与改革司（农业农村部乡村振兴专家咨询委员会办公室，甲方）

承接主体：华南农业大学 （乙方）

签订时间：2024 年 6 月 4 日

农业农村部计划财务司 制

使 用 说 明

一、本合同（范本）作为农业农村部在政府购买服务探索阶段使用的参考文本，双方协商签订合同时可结合项目特点和具体要求对相关条款内容进行适应性调整，或者增补其他条款，但原则上不得删减既定格式条款。

二、本合同（范本）中以脚注形式标注的文字，属于相关条款的补充说明、注意事项或者规范性要求，供甲乙双方在协商订立相关条款的过程中使用，在甲乙双方正式签订合同时，可以删除无关内容。

三、本合同（范本）为农业农村部政府购买服务事项的通用合同样式。确需使用其他合同样式的，应当作为本合同附件或者将本合同作为其附件，或者在其合同正文或者补充协议中明确包含本合同的既定格式条款。

四、本合同（范本）仅涉及购买主体（甲方）与承接主体（乙方），如相关政府购买服务事项涉及有必要在合同中明确权利和义务的其他责任主体，可根据需要增加合同签订方并增补相应条款。

五、承接主体利用本合同向金融机构融资时，应当配合金融机构做好合规性管理，并确保本合同约定的购

买内容和期限等符合政府购买服务有关法规制度规定，且不得含有购买主体对承接主体向金融机构融资的担保条款。

根据《中华人民共和国民法典》、《中华人民共和国政府采购法》、《国务院办公厅关于政府向社会力量购买服务的指导意见》（国办发〔2013〕96 号），以及财政部印发的《政府购买服务管理办法》（财政部令第 102 号）等有关规定，为保证农业农村部政府购买服务事项的规范实施，明确甲乙双方（以及本合同涉及的其他当事方）的权利义务，经双方协商，本着平等互利和诚信的原则，一致同意签订本合同。

第一条 服务的内容及要求

（一）【目录/代码】本合同所涉及的政府购买服务事项为：数字乡村建设的普惠效应、作用机理与优化路径研究，列入《农业农村部政府购买服务指导性目录》中四级目录“农业农村软科学课题研究服务”，政府购买服务代码为 B020103。

（二）【服务采购方式】甲方通过如下第【1】种[1]方式，确定由乙方提供服务:

1. 公开招标;

[1] 属于政府集中采购目录以内或政府采购限额标准以上的，应在 1-5 项中选择一项并填入序号；不属于政府采购范围的可以选择第 6 项“自行采购”并填入序号；方括号内只能填入 1 个选项的序号。

2. 邀请招标;

3. 竞争性谈判;

4. 竞争性磋商;

5. 单一来源采购;

6. 自行采购。

（三）【服务内容】由乙方向农业农村部乡村振兴专家咨询委员会办公室提供软科学课题研究咨询服务，具体内容详见附件二：农业农村部软科学课题研究项目任务书。

（四）【服务期限】本合同履行时间为 2024 年 月 日至 2024 年 12 月 31 日。其中[2]:

1. 乙方应当于 2024 年 8 月 31 日前完成阶段性研究;

2. 乙方应当于 2024 年 12 月 31 日前完成课题研究报告。

确需延长项目执行期限的，应向甲方提交书面申请，经书面答复同意后予以执行。未按时提交最终研究报告的，甲方有权取消乙方三年内申报农业农村部软科学课题的资格。

2. 本款如对相关服务内容的实施进展有阶段性考核要求，或对不同的具体服务内容有不同的时间节点要求，可根据需要分项进一步细化约定；如不涉及阶段性考核要求，也不涉及对不同的具体服务内容有不同的时间节点要求，可删除分项约定。

（五）【服务成果】乙方按照本条第四款约定期限，提供本条第三款所列服务内容，应当提交满足相应质量要求、规范标准和技术指标的如下成果:

1.向甲方提交 3000 字左右的决策参考报告（即阶段性研究成果）;

2.向甲方提交不少于 30000 字的课题最终研究报告和修改完善的决策参考报告。

第二条 服务的验收及评价

（一）【验收申请】乙方在完成第一条约定的服务内容并向甲方提供相应工作成果的基础上，以如下方式向甲方提出验收申请：乙方向甲方提交最终研究报告和决策参考报告，视同提出验收申请。

本合同约定服务事项的验收（☐以☑不以）第三方专业机构出具的审计报告作为验收的前置条件。

（二）【验收方式】乙方的服务成果至少满足以下任意两个条件，视为通过验收:

1.通过农业农村部乡村振兴专家咨询委员会办公室上报，以农业农村部软科学课题成果名义获得党和国家领导人或农业农村部领导批示;

2.农业农村部乡村振兴专家咨询委员会办公室给予肯定;

3.在农业农村部主办的《乡村振兴文稿》发表至少1篇文章。

（三）【服务质保】本合同约定的政府购买服务事项的质量保证期为自验收通过之日起【60】个自然日之内。该期限内乙方提供的工作成果若出现任何问题，乙方均应当在接到甲方（□书面☑口头）通知之日起【30】个自然日内进行修正，包括但不限于重新提供服务、修正服务成果、采取补救措施等形式，并不得向甲方额外收取任何费用。本款约定不排除乙方质保期过后的相关质保义务。

（四）【绩效管理】根据《中共中央国务院关于全面实施预算绩效管理的意见》(中发〔2018〕34 号)以及《政府购买服务管理办法》等有关规定，甲方应当建立健全政府购买服务项目绩效目标设置与审核、绩效执行监控、绩效评价及结果运用的全过程绩效管理机制，不断提高政府购买服务质量和效益。

（五）【绩效目标】就本合同约定的政府购买服务事项等，甲方与乙方协商设置相应绩效目标以及数量、质量、时效、成本、效益和服务对象满意度等绩效指标，以附件（一）形式作为本合同不可分割的一部分。

（六）【绩效监控】在本合同履约期内，乙方应当加强绩效执行监控，对绩效目标及指标实现程度进行跟

踪分析，并将绩效执行监控结果以书面形式报告甲方。实际执行情况与绩效目标偏离的，应当及时予以纠正，确保绩效目标如期实现。甲方原则上按照财政部组织开展相应项目支出预算绩效执行监控的具体要求和时间节点，督促乙方开展相应绩效执行监控工作并提交报告。

（七）【绩效自评】本合同约定期限届满时，乙方应当对本合同约定绩效目标及指标的实现情况开展绩效自评，填报绩效目标完成情况，分析未完成的原因，提出改进措施，并于本合同期限届满之日起【30】个工作日内将绩效自评结果以书面形式报告甲方。在本合同约定期限尚未届满、但预算年度已经终了的情况下，甲方应按照财政部组织开展相应项目支出预算绩效自评的要求，督促乙方开展相应绩效自评工作并提交报告。

（八）【绩效评价】根据《财政部关于推进政府购买服务第三方绩效评价工作的指导意见》（财综〔2018〕42 号），甲方可对本合同约定的政府购买服务事项组织第三方机构开展绩效评价，或者配合财政部门直接组织开展的绩效评价。乙方应当积极配合相应绩效评价工作。

第三条 合同金额及款项支付

（一）【合同金额】本合同服务报酬总金额为人民币（大写）：壹万元（小写：¥10000 元）。乙方提供本

合同约定服务所发生的各项成本费用，以及与履行本合同有关的一切税费，均已包含在以上服务报酬总金额之中，由乙方负担。

（二）【收款账户】乙方开户名称为：华南农业大学；开户银行为：中国工商银行广州五山支行；银行账号为：3602002609000310520。

（三）【付款方式】根据有关资金管理规定和服务类型，选择以下任一种付款方式：甲方以（☑财政拨款 □甲方实有资金账户转账支付）方式付款。

（四）【付款时间】甲方按照以下第【1】项（仅可选择一项）约定的付款时间向乙方支付合同款项[3]：

【事前一次性付款】本合同生效之日起60个工作日内，甲方一次性支付全部服务费。

第四条 甲方的权利和义务

（一）甲方应当加强政府购买服务项目履约管理，督促乙方严格履行合同等约定，有权随时向乙方了解提供服务的实施进度，并要求乙方提供相关资料。

（二）甲方有权按照本合同约定或者有关法律法规、部门规章、规范性文件的规定，对乙方履约情况进行监督和检查，并督促乙方加以整改。但不因甲方行使

3. 根据本合同约定购买服务事项的特点及内容、期限、成果等约定，甲乙双方协商确定分期付款的条件和金额。

该监督和检查权而减轻或者免除乙方应承担的任何义务或者责任。

（三）甲方有权根据国家政策或者法律法规的变动对服务的质量要求、标准规范和技术指标做出相应变动，有权在发现乙方履行合同过程中出现损害或者可能损害公共利益、公共安全情形时终止本合同。

（四）甲方有权将乙方履行合同情况及不符合政府购买服务管理相关规定的情况，向相关部门报告并纳入不良信用记录、年检（报）、评估、执法等监管体系中。

（五）甲方承诺确保本合同约定的政府购买服务事项，符合国家关于政府购买服务的内容要求，并已列入农业农村部政府购买服务指导性目录。

（六）甲方承诺履行本合同所需财政资金，已经列入相应年度的部门预算和中期支出规划安排，并已经按规定编制相应政府购买服务预算及政府采购预算。

（七）甲方应及时向乙方提供或者协调有关方面提供与履行本合同相关的，且为乙方开展工作必备前提的政策、制度、规范、标准、数据等资料。

（八）甲方应当根据乙方履行合同和开展工作的实际需要，积极协助乙方做好与相关政府部门、企事业单位及其他第三方的沟通、协调工作。

（九）甲方应当依照本合同约定的金额、方式、时间、条件等，及时向乙方支付服务报酬。

（十）向乙方下达课题研究任务和开展相应的组织管理工作时，必须遵守《中华人民共和国著作权法》和农业农村部软科学课题管理的有关规定。

（十一）视需要要求乙方提交阶段性研究成果。

（十二）组织乙方参加有关学术活动。

（十三）及时组织对研究成果的结题验收，对验收合格的成果出具结题验收证明。

（十四）因乙方或课题主持人违反本项目任务书约定以及未履行有关义务，给甲方造成损失的，甲方有权取消乙方或课题主持人今后申请课题的资格。

第五条 乙方的权利和义务

（一）乙方有权要求甲方提供或者要求甲方协调有关方面提供与履行本合同相关的，且为乙方开展工作必备前提的政策、制度、规范、标准、数据等资料。

（二）乙方有权根据履行合同和开展工作的实际需要，要求甲方协助做好与相关政府部门、企事业单位及其他第三方的沟通、协调工作。

（三）乙方有权依照本合同约定的金额、方式、时间、条件等向甲方收取服务报酬，在账户信息变更时有

权要求甲方根据乙方出具的相关证明材料向变更后的账户付款。

（四）乙方在不违反本合同约定或者有关法律法规、部门规章、规范性文件的规定的前提下，有权根据提供服务的实际成本费用等，自主安排使用所获服务报酬。

（五）乙方应当配备具有相应资质、能力、经验的工作人员承担本合同约定服务事项的具体实施，保证其具有充裕的工作时间，并为其提供相应条件支持。

（六）乙方应当按照本合同约定的内容、期限、数量、质量、标准等要求，及时提供服务并提交成果，并根据甲方需要报告工作进展情况。乙方不得将本合同约定的服务项目转包给其他主体。

（七）乙方应当按照有关规定建立承接政府购买服务项目台账，妥善保存与本合同履约活动相关的纸质文档、电子文档、影音图像等信息，保存期限为本合同履行完毕或者终止后 15 年。

（八）乙方应当建立健全财务制度，严格遵守相关财务规定，确保政府购买服务资金规范管理和使用，以辅助账等适当形式进行单独核算，相应原始凭证、会计账簿等财务资料依照上款要求妥善保存。

（九）乙方有义务接受和配合甲方或者有关部门对本合同履行情况的监督和检查，按要求完整、准确、如实提供所需业务和财务资料，并对监督检查提出的问题及时做出合理解释或者加以整改。

（十）乙方应当全面履行本合同约定服务事项实施过程中的相关安全管理职责，因乙方未尽到管理职责发生安全事故的，由乙方承担相应的法律责任。

（十一）受甲方委托，按照农业农村部软科学课题管理的有关规定，承担本项目的各项管理工作，包括课题立项、督促指导等。乙方承担甲方下达的课题研究任务，应遵守《中华人民共和国著作权法》有关规定。

（十二）保证课题主持人及主要研究人员完成本课题所需工作时间，并配备人员、办公场所等必要条件。对承担的研究课题要认真组织研究，保证质量，不得中途变更和调整课题项目任务书中的内容，包括项目名称、主持人、主持单位和研究的主要内容。确需变更的，应向甲方提交书面申请，经书面答复同意后才予变更。

（十三）因乙方非不可抗力原因确需延长项目执行期限的，应向甲方提交书面申请，经书面答复同意后予以执行。未按时提交最终研究报告的，甲方有权取消乙方三年内申报农业农村部软科学课题的资格。

（十四）甲方组织开展项目结题验收，并将验收结果通报乙方和课题主持人。通过验收的课题，由甲方出具结题证明。需要修改的课题，课题主持人在60天内可再次提出验收申请。经两次申请仍未通过验收的，主持人三年内不得再次申报农业农村部软科学课题。

（十五）上报的研究成果和资料，不得再用于其他单位委托的研究课题，也不得以过去的研究成果或其他研究成果代替本课题成果。

（十六）乙方在申报成果鉴定、评奖以及在公开或内部媒介上应用研究课题研究成果时，必须事先书面征得甲方同意，并标明农业农村部软科学课题字样。

（十七）协调组织课题主持人参加甲方组织的相关学术活动。完成的研究成果有权参加甲方组织的成果鉴定和评奖。

第六条 违约责任

（一）乙方提供的服务不符合本合同约定或相关法律法规、部门规章、规范性文件之规定的，甲方有权拒收并由乙方按照本合同约定服务报酬总金额的 1 %，向甲方支付违约金。

（二）乙方未能按照本合同约定时间提供服务或提交成果的，从逾期之日起每日按逾期金额的 1 %，向甲

方支付违约金；逾期 60 日以上的，甲方有权终止合同，由此造成的甲方经济损失由乙方承担。

（三）乙方不得将本合同服务项目转包给其他主体。如擅自转包，则乙方应当按照本合同约定服务报酬总金额的 20 %，向甲方支付违约金。

（四）其他违约责任，甲方可以根据本合同约定购买服务事项的具体情况，依据《中华人民共和国民法典》等法律法规的相关规定，在本合同中约定。

（五）经甲乙双方协商一致，同意免除或者部分免除对方违约责任的，双方应当签订《补充协议》或《谅解备忘录》进行约定。

第七条 知识产权事务

（一）【使用甲方知识产权】甲方按照本合同第四条第（七）款的约定向乙方提供的资料，如涉及甲方享有的知识产权，甲方不得以任何名义向乙方另行收取相应知识产权使用费。未经甲方书面许可，乙方不得将相应知识产权用于本合同约定服务事项之外的其他用途，亦不得以任何形式侵占甲方相应知识产权。

（二）【使用乙方知识产权】乙方按照本合同有关约定向甲方提供的资料，如涉及乙方在履行本合同之前已经享有的知识产权（不含乙方因履行本合同而形成的知识产权），乙方不得以任何名义向甲方另行收取相应

知识产权使用费。未经乙方书面许可，甲方不得将相应知识产权用于本合同约定服务事项之外的其他用途，亦不得以任何形式侵占乙方相应知识产权。

（三）【使用第三方知识产权】甲方按照本合同第四条第（七）款的约定，向乙方提供或者协调第三方提供的资料，如涉及第三方享有的知识产权，甲方应当取得有关权利人的合法授权。如发生涉及到专利权、著作权、商标权等争议，甲方负责处理并承担由此引起的全部法律责任。

（四）【使用第三方知识产权】除甲方按照本合同第四条第（七）款的约定，向乙方提供或者协调第三方提供的资料外，乙方在履行本合同过程中使用其他第三方依法享有的知识产权，应当取得有关权利人的合法授权。如发生涉及到专利权、著作权、商标权等争议，乙方负责处理并承担由此引起的全部法律责任。

（五）【乙方形成的知识产权】乙方因履行本合同而形成的研究成果，包括但不限于各种形式的课题报告、技术资料、发明专利、论文著作、影音图像、种质资源、遗传资源等，其知识产权归甲方所有，包括但不限于作品的发表权、署名权、修改权、复制权、发行权、信息网络传播权、汇编权和其他权利。

第八条 保密条款

（一）【甲方提供的涉密信息】甲方按照本合同第四条第（七）款的约定，向乙方提供或者协调第三方提供的资料，如有属于《中华人民共和国保守国家秘密法》规定的国家秘密范围的信息，应当严格按照相关保密规定进行管理，并责成乙方严格按照保密规定进行管理。

（二）【甲方提供的敏感信息】甲方按照本合同第四条第（七）款的约定，向乙方提供或者协调第三方提供的资料，如有属于甲方工作秘密或者内部资料的，应当以编列目录清单的形式如实告知乙方，并责成乙方严格按照相关保密规定和本合同约定进行管理。

（三）【乙方对甲方信息的保密责任】乙方应当对甲方提供或者协调第三方提供的资料负有保密责任。属于本条第（一）款所列国家秘密以及本条第（二）款所列工作秘密或内部资料的，乙方应当严格按照相关保密规定和甲方要求进行管理。其他信息未经甲方书面许可，乙方亦不得提供给其他单位和个人。

（四）【乙方对其他涉密信息的保密责任】乙方在承接本合同约定的政府购买服务事项的过程中，所获得的政府、企事业单位、公民个人等各种信息和资料，如有属于《中华人民共和国保密法》规定的国家秘密范围

的信息，乙方应当严格按照相关保密规定和甲方要求进行管理。

（五）【乙方对其他相关信息的保密责任】乙方在承接本合同约定的政府购买服务事项的过程中，所获得的政府、企事业单位、公民个人等各种信息和资料，不属于《中华人民共和国保密法》规定的国家秘密范围的，未经甲方书面许可，乙方不得提供给其他单位和个人，也不得以乙方名义或者乙方工作人员个人名义做出公开披露和公开发表观点。

（六）【甲方对乙方信息的保密责任】甲方应当保护乙方的商业秘密（含技术秘密）和其他敏感信息，未经乙方书面许可，不得将所知悉的乙方商业秘密（含技术秘密）和其他敏感信息提供给其他单位和个人。但甲方依据相关法律法规规定、法定职责或本合同约定，向社会公开或提供给相关部门的除外。

第九条 争议的解决

本合同在履行过程中发生的任何争议，由甲乙双方友好协商解决。协商未果的，双方均同意以如下第【1】种（仅可选择一项）方式解决:

1. 任何一方均可通过甲方所在地有管辖权的人民法院提起诉讼处理;

2. 将争议提交 ×××××××××××× 仲裁委员会，适用申请仲裁时现行有效的仲裁规则裁决，该仲裁结果是终局的，对双方均具有约束力。

第十条 不可抗力

任何一方由于不可抗力原因不能履行合同时，应在不可抗力事件发生后 3 个工作日内向对方通报，以减轻可能给对方造成的损失，在取得有关机构的不可抗力证明或者对方谅解确认后，双方可以协商延期履行或者修订合同；双方根据具体情况可协商一致，全部或者部分免除受不可抗力影响的一方的违约责任。

第十一条 合同的终止

（一）本合同期满，双方不再续签合同的；

（二）乙方丧失提供服务的能力，致使本合同服务无法正常提供的；

（三）在履行合同过程中，发现乙方已不符合承接主体应具备的条件，造成合同无法履行的；

（四）乙方明确表示不再履行本合同的；

（五）乙方履行本合同不符合约定，在甲方确定的合理期限内仍不能纠正的；

（六）受国家政策或者法律法规变动影响，经双方协商终止本合同的。

第十二条 其他

（一）本合同所有附件均为本合同的有效组成部分，与本合同具有同等法律效力。若合同附件与本合同正文条款存在不一致的，则以本合同正文条款为准。

（二）在履行本合同过程中，所有经双方签署确认的文件（包括但不限于会议纪要、补充协议、往来信函等）即成为本合同的有效组成部分。

（三）如一方名称、地址、负责人、联系人、联系方式或者乙方银行账户信息有变更，应在发生变更之日起 3 个工作日内通知对方，否则应承担未及时通知产生的相应后果。

第十三条 合同生效

（一）本合同订立时间：2024年6月4日。

（二）本合同订立地点：北京市朝阳区农展南里 11 号。

（三）本合同自甲、乙双方法定代表人或者代理人签字并加盖公章之日起生效。

（四）本合同甲方的代理人即作为购买主体的农业农村部相关司局或派出机构的负责人。

（五）本合同一式肆份，甲乙双方各执贰份，双方文本具有同等法律效力。

第十四条 合同附件[4]

4. 政府购买服务绩效目标表是本合同必备附件；经由政府采购程序实施服务采购的，相应采购文件及供应商应答文件，应作为本合同附件；若有其他附件应逐一列示附件名称。

（一）政府购买服务绩效目标表;

（二）农业农村部软科学课题研究项目任务书。

（以下无正文。）

（此表单面打印）

签 章 页

甲方	购买主体	法定代表人或者代理人签字： 单位公章： 2024年6月4日		
	业务负责处室	负责人签字： 2024年6月4日		
	联系人	周厚熹	联系电话	010-59191593
	通讯地址	北京市朝阳区农展南里11号	邮政编码	100125
乙方	承接主体	法定代表人或者代理人签字： 单位公章： 2024年6月9日		
	业务负责部门	负责人签字： 2024年6月9日		
	联系人		联系电话	[redacted]
	通讯地址		邮政编码	510642

科研论文

SCAULIB202519170

检索证明

根据委托人提供的论文材料，委托人华南农业大学经济管理学院 赵纯凯 8 篇论文收录情况如下表。

序号	论文名称	发表刊物及发表的年月卷期/页码等	作者排名	论文等级	作者文中单位	收录情况	影响因子	中科院大类分区
1	Rising housing prices and marriage delays in China: Evidence from the urban land transaction policy	Cities 出版年：2023 卷期：135 页码：- 文献号：104214 文献类型：Article	第一作者	T2 类	华南农业大学 经济管理学院	SSCI	IF2-year=6.0 IF5-year=6.8 （2023）	经济学 1 区 Top 期刊：是 （2023）
2	Is digital finance environmentally friendly in China? Evidence from shared-bike trips	Transport Policy 出版年：2023 卷期：138 页码：129-143 文献号： 文献类型：Article	第一作者	A 类	华南农业大学 经济管理学院	SSCI	IF2-year=6.3 IF5-year=5.9 （2023）	工程技术 2 区 Top 期刊：否 （2023）
3	ICT in education can improve students' achievements in rural China: The role of parents, educators and authorities	Journal of Policy Modeling 出版年：2023 卷期：45 2 页码：320-344 文献号： 文献类型：Article	第一作者	A 类	华南农业大学 经济管理学院	SSCI	IF2-year=3.5 IF5-year=3.3 （2023）	经济学 2 区 Top 期刊：否 （2023）

华南农业大学图书馆 SCAULIB202519170
华南农业大学图书馆 信息检索专用章

4	School nutritious feeding and cognitive abilities of students in poverty: Evidence from the nutrition improvement program in China	Children and Youth Services Review 出版年：2024 卷期：159 页码：- 文献号：107519 文献类型：Article	第一作者	B类	华南农业大学 经济管理学院	SSCI	IF2-year=1.7 IF5-year=2.8 （2024）	社会学 3区 Top期刊：否 （2025）
5	来自军营的印记———军人职业与老年人健康	经济科学 出版年：2023 卷期： 页码：- 文献号： 文献类型：	第一作者	B类	华南农业大学 经济管理学院	CSSCI,北大核心	无	无
6	Military Service and Health Outcomes of the Elderly in China	Armed Forces & Society 出版年：2024 卷期：50/3 页码：781-809 文献号： 文献类型：Article	第一作者	B类	华南农业大学 经济管理学院	SSCI	IF2-year=1.7 IF5-year=1.7 （2024）	社会学 3区 Top期刊：否 （2025）
7	Understanding government pollution control actions and well being: Evidence from micro individuals	AIR QUALITY ATMOSPHERE AND HEALTH 出版年：2024 卷期：17 9 页码：1917-1935 文献号： 文献类型：Article	第一作者	C类	华南农业大学 经济管理学院	SCI	IF2-year=2.9 IF5-year=3.1 （2024）	环境科学与生态学 4区 Top期刊：否 （2025）

华南农业大学图书馆SCAULIB20251910
华南农业大学图书馆 信息检索专用章

8	The effect of digital finance on Residents' happiness: the case of mobile payments in China	Electronic Commerce Research 出版年：2024 卷期：24 1 特刊 SI 页码：69-104 文献号： 文献类型：Article	第一作者	C 类	华南农业大学 经济管理学院	SSCI	IF2-year=4.7 IF5-year=4.6 （2024）	管理学 4 区 Top 期刊：否 （2025）

说明：论文等级和中科院大类分区按《华南农业大学学术论文评价方案（试行）》划分。

报告免责声明:如未盖章,报告无效

检索员：尹银怀

华南农业大学图书馆

2025-07-15

华南农业大学图书馆SCAULIB20251[illegible]

SCAULIB202519175

检索证明

根据委托人提供的论文材料，委托人华南农业大学经济管理学院 赵纯凯 8 篇论文收录情况如下表。

序号	论文名称	发表刊物及发表的年月卷期/页码等	作者排名	论文等级	作者文中单位	收录情况	影响因子	中科院大类分区
1	Official Turnover and Sustainable Development in China	Journal of Chinese Political Science 出版年：2024 卷期：29 4 页码：649-672 文献号： 文献类型：Article	通讯作者	T2 类	华南农业大学 经济管理学院	SSCI	IF2-year=3.5 IF5-year=3.6 （2024）	社会学 1 区 Top 期刊：是 （2025）
2	Online markets, offline happiness: E-commerce development and subjective well-being in rural China	China Economic Review 出版年：2024 卷期：87 页码：- 文献号：102247 文献类型：Article	通讯作者	A 类	华南农业大学 经济管理学院	SSCI	IF2-year=5.5 IF5-year=6.7 （2024）	经济学 2 区 Top 期刊：是 （2025）
3	More is less: Homeroom teachers' administrative duties and students' achievements in China	Teaching and Teacher Education 出版年：2022 卷期：119 页码：- 文献号：103857 文献类型：Article	通讯作者	A 类	华南农业大学 经济管理学院	SSCI	IF2-year=3.9 IF5-year=4.8 （2022）	教育学 2 区 Top 期刊：是 （2022）

华南农业大学图书馆 信息检索专用章
华南农业大学图书馆SCAULIB202519175

4	The long-term effect of childhood left-behind experiences on entrepreneurial activities in adulthood: Empirical evidence from China	Children and Youth Services Review 出版年：2024 卷期：164 页码：- 文献号：107879 文献类型：Article	通讯作者	B 类	华南农业大学 经济管理学院	SSCI	IF2-year=1.7 IF5-year=2.8 （2024）	社会学 3 区 Top 期刊：否 （2025）
5	Can national forest city construction mitigate air pollution in China? Evidence from a quasi natural experiment	Environmental Geochemistry and Health 出版年：2023 卷期：45 6 页码：3003-3024 文献号： 文献类型：	通讯作者	B 类	华南农业大学 经济管理学院	SCI	IF2-year=3.2 IF5-year=4.0 （2023）	环境科学与生态学 3 区 Top 期刊：否 （2023）
6	Did innovative city constructions reduce carbon emissions? A quasi-natural experiment in China	ENVIRONMENT DEVELOPMENT AND SUSTAINABILITY 出版年：2024 卷期：26 3 页码：6315-6340 文献号： 文献类型：Article	通讯作者	B 类	华南农业大学 经济管理学院	SCI	IF2-year=4.2 IF5-year=4.5 （2024）	环境科学与生态学 3 区 Top 期刊：否 （2025）
7	Beyond the scenery: Understanding the impact of rural tourism development on household consumption in China	Review of Development Economics 出版年：2024 卷期：28 3 页码：1152-	通讯作者	B 类	华南农业大学 经济管理学院	SSCI	IF2-year=2.2 IF5-year=2.2 （2024）	经济学 3 区 Top 期刊：否 （2025）

华南农业大学图书馆 SCAULIB20251912

		1191 文献号： 文献类型：Article						
8	Teachers' administrative positions and students' non-cognitive abilities in China: evidence from a quasi-natural experiment	Applied Economics 出版年：2024 卷期：56 50 页码：6091-6108 文献号： 文献类型：Article	通讯作者	C类	华南农业大学 经济管理学院	SSCI	IF2-year=2.1 IF5-year=2.2 （2024）	经济学 4区 Top 期刊：否 （2025）

说明：论文等级和中科院大类分区按《华南农业大学学术论文评价方案（试行）》划分。

报告免责声明:如未盖章,报告无效

检索员：尹银怀

华南农业大学图书馆

2025-07-15

华南农业大学图书馆CAULIB20251975

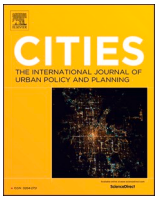
第一作者论文

Cities 135 (2023) 104214

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities



Rising housing prices and marriage delays in China: Evidence from the urban land transaction policy

Chunkai Zhao [a], Boou Chen [b,*], Xing Li [c]

[a] College of Economics and Management, South China Agricultural University, No.483, Wushan Street, Tianhe District, Guangzhou, Guangdong Province, China
[b] School of Economics, Nanjing University of Finance and Economics, No.3 Wenyuan Road, Qixia District, Nanjing, Jiangsu Province, China
[c] College of Economics and Management, Southwest University, No. 2, Tiansheng Road, Beibei District, Chongqing, China

ARTICLE INFO

Keywords:
Rising housing prices
Marriage delays
Age at first marriage
Urban land transaction policy (ULTP)
Fertility
China

ABSTRACT

In recent years, the phenomenon of delayed age at first marriage and declining fertility in China has become increasingly prominent, and high housing prices may be one of the triggers. By exploiting the quasi-natural experiment of the urban land transaction policy (ULTP) since 2002 and using the nationally representative data from the China Labor-force Dynamic Survey (CLDS), we employed a difference-in-differences (DID) method to identify the causal effects of rising housing prices on marriage delays. The results show that rising housing prices from the ULTP significantly delayed the age of first marriage by 0.2280 years old, with males delaying by 0.4372 years old. In addition, we found that rising housing prices raised the cost of first marriage; instead, young people with houses have a younger age at first marriage. Furthermore, the marriage delay effect of rising housing prices is more pronounced for more educated females, males with more siblings, and individuals from large cities. Worse, the effect of delayed marriage caused by rising housing prices further depressed fertility intentions and reduced the actual fertility rate. These findings complemented to the research field on the impact of rising housing prices and provided some insights for refining real estate public policies.

1. Introduction

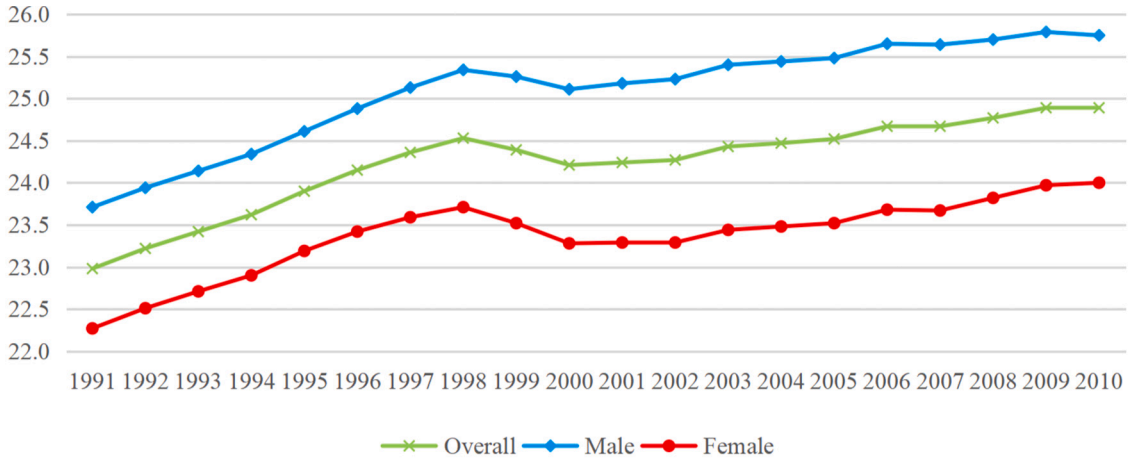
China is a country of universal marriage, yet the postponement of age at first marriage is becoming more and more prominent (Nie, 2020; Yu & Xie, 2015). As shown in Fig. 1(a), statistics from the Seventh National Population Census of China in 2020 show that the average age at first marriage rose from 22.98 in 1991 to 24.89 in 2010. Among them, the age rose from 23.71 to 25.75 for males and from 22.27 to 24.00 for females.[1] Worse, this phenomenon is even more serious in urban areas. In Fig. 1(b), the average age of first marriage of urban residents is consistently higher than rural residents after 1991. Especially since 2002, the age of first marriage of urban residents has increased significantly, while that of rural residents has remained relatively stable. In 2009, the gap between the age of first marriage of urban and rural residents had reached near 2 years. In the long run, excessive and widespread delay in marriage will result in delayed births, fertility declines (Gutiérrez-Domènech, 2008; Wrenn et al., 2019), and exacerbate the population aging (Raymo, 1998; Zeng & Vaupel, 1989). These negative effects on society are undoubtedly pose challenges to sustainable development in developing countries, including China as well.

The phenomenon of marriage delays has attracted considerable attention from scholars. Previous research analyzed the potential reasons from multiple perspectives, such as changes in marriage concepts (Cao et al., 2019; Ji, 2015; Xu & Whyte, 1990), population structure (Bongaarts & Greenhalgh, 1985; Cai, 2010; Huang et al., 2020), gender imbalance (Anderson, 2007; Nie, 2020; Wei & Zhang, 2011), socioeconomic status of women (Ji, 2015; Ye & Zhao, 2018; Zhang et al., 2008), and extension of years of education (Goldstein & Kenney, 2001; Kalmijn, 2013; Stevenson & Wolfers, 2007). In recent years, a growing number of studies have looked closely at the delay in marriage brought about by the rising cost of getting married (Horioka & Terada-Hagiwara, 2017; Wei & Zhang, 2011; Yang et al., 2021). In many areas of China, the bride prices, private cars, and houses provided by males are considered necessities for marriage, especially the house (Cheung et al., 2020; Fang & Tian, 2018;Hu et al., 2021 ; Wrenn et al., 2019). According to BEIKE Research Institute, from January to October 2021, the average age of house purchase in China's first-tier cities was about 36.9 years old, 34.2

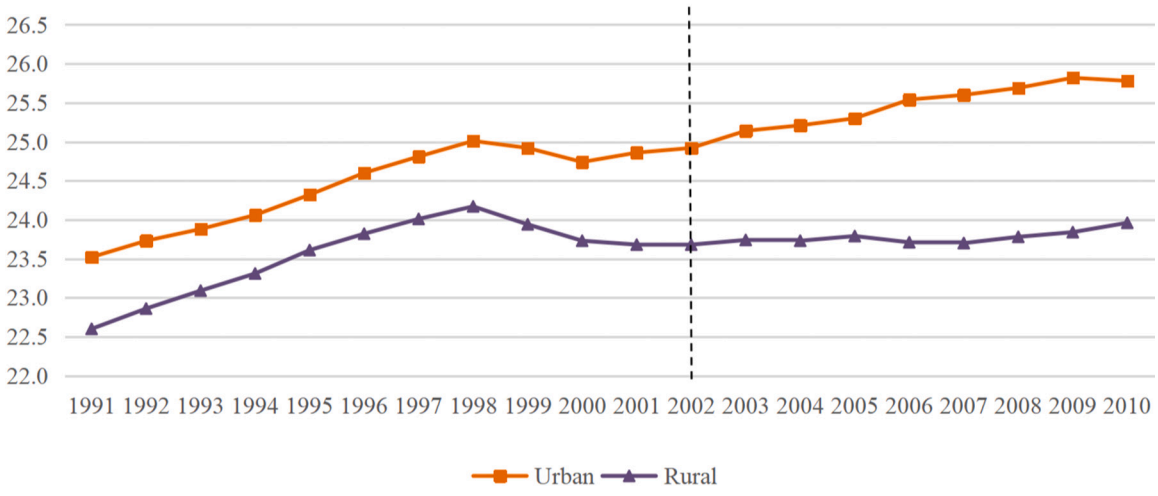
* Corresponding author.
E-mail address: drchen2015@163.com (B. Chen).

[1] Data source: http://www.stats.gov.cn/tjsj/pcsj/rkpc/7rp/zk/indexce.htm.

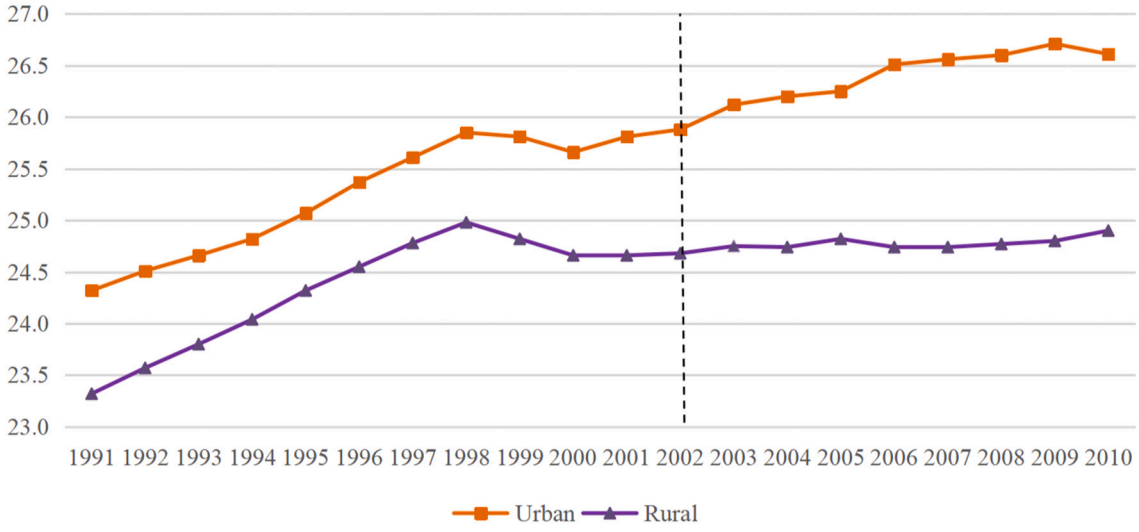
https://doi.org/10.1016/j.cities.2023.104214
Received 18 November 2021; Received in revised form 14 December 2022; Accepted 15 January 2023
Available online 26 January 2023
0264-2751/© 2023 Elsevier Ltd. All rights reserved.



(a)



(b)



(c)

Fig. 1. (a). Average age at first marriage in China: 1991–2010.
(b). Average age at first marriage in urban and rural China: 1991–2010.
(c). Average age at first marriage for males in urban and rural China: 1991–2010.
Source: the Seventh National Census of China in 2020 (http://www.stats.gov.cn/tjsj/pcsj/rkpc/7rp/zk/indexce.htm)

years old in new first-tier cities and 34.8 years old in second-tier cities.[2] From the above survey data, it is clear that the burden of purchasing a house remains high for young people in China. Considering the strong correlation between home ownership and marriage in China, it is likely that the rise in housing prices has not only delayed the age of home ownership for young people, but also postponed their age of first marriage (Hu et al., 2021; Wrenn et al., 2019).

Prior studies explored the impact of rising housing prices on marriage in two main ways. On the one hand, some scholars have emphasized the burden effect of rising housing prices, which led to a decline in marriage rates (Assaad et al., 2021; Hu et al., 2021; Hu & Wang, 2020; Wrenn et al., 2019). For example, Wrenn et al. (2019) found that on average, a 1 % increase in Chinese housing prices is associated with an average 0.31 % decrease in the marriage rate. On the other hand, focusing on the wealth effect, another category of literature concluded that rising housing prices help promote marriage and fertility (Eriksen, 2010; Hu & Wang, 2020; Lovenheim & Mumford, 2013; Pan & Yang, 2022). Given these two effects, Su et al. (2020) found that the impact of rising housing prices on marriage may be nonlinear.

The conflict in previous literature is mainly due to the differences in research perspectives, as there is a very strong link between house purchase and marriage decisions. Some scholars found that getting married or competition in the marriage market is an important driver of housing price inflation (e.g., Cheung et al., 2020; Feijten & Mulder, 2002; Mulder & Wagner, 1998; Yang et al., 2021), especially in countries with gender imbalances and marriage squeeze, like China (Wei et al., 2017; Wei & Zhang, 2011). Thus, there may be a two-way causal relationship between marriage delay and rising housing prices. In addition, rising housing prices in a given city is influenced by many factors, and some of these confounders are difficult to fully control for, such as infrastructure, historical situation, surrounding environment, and quality of education (Engle et al., 1985; Wen et al., 2019; Yuan et al., 2020). Therefore, potential omitted variable bias may also lead to inaccuracy in estimates.

To examine the causal effect of housing prices on the age at first marriage, we used the urban land transaction policy (ULTP) implemented in 2002 as a quasi-natural experiment to identify the causal relationship between rising housing prices and marriage delays of young people, as Chinese housing prices are closely related to land policies (Deng et al., 2012; Ding, 2007; Tian et al., 2020; Wang & Hou, 2021). Specifically, the Chinese government introduced more flexible market mechanisms into urban land transactions since 2002, allowing firms and institutions to acquire access to land through public bidding, auctions, and offers. As an important component of local government land finance policies, this land transaction policy led to a sharp increase in housing and land prices after implementation (Hu et al., 2021). More importantly, this policy is only implemented in urban areas, while land and housing prices in rural areas are hardly affected (Hu et al., 2021; Wang et al., 2018), which sets the stage for us to use a difference-in-differences (DID) framework to examine the effect of housing prices on the age at first marriage in China. In addition, as mentioned in Fig. 1, we observed that the gap between the age at first marriage of young people in urban and rural China began to widen around 2002, which provides a realistic basis for analyzing the effect of rising housing prices on marriage delays from the perspective of ULTP.

This paper has several interesting findings and contributes to the existing literature in the following ways. First, by using the nationally representative data from the China Labor-force Dynamic Survey (CLDS), we found that the effect of rising housing prices from ULTP delayed the age at first marriage for Chinese young people. From the perspective of marriage, our findings support the idea that the burden effect of rising housing prices on Chinese young people is higher than the wealth effect (Waxman et al., 2020; Wrenn et al., 2019). Thus, our study further enriches two types of studies on the adverse effects of high housing prices on Chinese families (e.g., Chen et al., 2021; Hu et al., 2021; Hu & Wang, 2020; Wang & Wen, 2012; Waxman et al., 2020; Wrenn et al., 2019), and causes for marriage delays among Chinese young people (e.g., Cai, 2010; Huang et al., 2020; Ji, 2015; Ye & Zhao, 2018). In particular, previous studies have mainly emphasized the effects of rising housing prices on household formation and marriage rates (Hu et al., 2021; Wrenn et al., 2019). In a country of universal marriage like China, it is more relevant to explore the effect of rising housing prices on the age at first marriage.

Second, our findings show that after ULTP implementations, the more rapidly the housing prices rose, the more serious the marriage delay phenomenon. In addition, rising housing prices also caused more educated females and males with more siblings to delay their marriage. Our findings help explain the household economic behaviors of different groups of marriages in the Chinese context of rising housing prices (e.g., Hu et al., 2021; Hu & Sun, 2021; Wei & Zhang, 2011), and also provide insights for the development of differentiated and targeted public policies related to real estate.

Third, our estimates suggested that a further consequence of rising housing prices from the ULTP is a decline in fertility intentions and actual fertility. Therefore, this study provides some evidence to further explain China's low fertility phenomenon in the context of China's fertility incentives (Cai, 2010; Jiang & Liu, 2016; Zheng, 2019). This paper also complements related research fields on housing prices and fertility (e.g., Atalay et al., 2021; Clark et al., 2020; Lovenheim & Mumford, 2013; Pan & Yang, 2022). From the perspective of stimulating marriage and fertility, our analysis provides some potential directions for housing prices regulation in China and may have some references for other countries as well.

The remainder of this study is organized as follows. Section 2 introduces the institutional background of the land transaction policy in China. Section 3 outlines the theoretical framework. Section 4 explains the estimation strategy, data, and variables. Section 5 presents the main empirical results and robustness checks. Section 6 is the further studies. Section 7 concludes.

2. The land transaction policy in China

In 1986, the Chinese government formally adopted the Land Management Law of the People's Republic of China,[3] which established the system of land leasing with a fee and was the first major attempt to manage urban land by market-oriented approaches. Subsequently, cities such as Shenzhen and Shanghai achieved success in the pilot reform of land leasing with a fee, and this system was gradually being extended nationwide. As an emerging land transaction policy, the land leasing system has given rise to the real estate market (Hu et al., 2021; Wang et al., 2018). However, the lack of openness and market competition in the transaction process resulted in several problems such as poor utilization of land resources, corruption, hindering the equilibrium of land transaction markets, and destruction of state-owned assets (Deng et al., 2012; Gao, 2019; Liu & Zhang, 2020).

To solve these problems, the government has made a series of institutional innovations in land supply methods, resulting in the coexistence of state-owned land allocation, granting, leasing, authorized operation of land use rights, and capital contribution of land use rights for equity (He et al., 2014; Tian & Ma, 2009; Wang et al., 2018), which has created momentum for urban economic development (Li, 2014; Mo,

[2] Data source: BEIKE Research Institute, 2021 Annual Report on Consumption Trends of Residential Customers. BEIKE is one of the major real estate trading platforms in China, and BEIKE Research Institute is the research department of the company that analyzes data related to real estate transactions in China.

[3] Information source: http://www.npc.gov.cn/wxzl/gongbao/2000-12/06/content_5004471.htm.

2018). One representative land transaction policy is the "Regulations on the Sale of State-owned Land Use Rights by Bidding, Auction, and Listing" announced in 2002, which requires all urban commercial, tourism, entertainment, and commercial residential land to be sold by bidding, auction, or listing.[4] Moreover, the bidding process and bidding information were further transparent, and a series of information including the grantor, the land to be granted, the tenure, the use, the planning and design requirements, and the method of bidding qualifications were required to be announced. This ULTP marked a transformation from a negotiation-based to a market-based approach to land transactions in China (Deng, 2005), which put the land resource allocation system and the local government's land finance system on the right track (Chen et al., 2019; Hu et al., 2021).

The ULTP has made the transaction process of urban land more transparent, fair and openly informative, with a greater impact on both land prices and housing prices in China's cities (Wang et al., 2018). Several previous studies pointed out that the market competition introduced by the reform in the way of land transactions led to rapid increases in land transaction prices and housing prices (Chen et al., 2019; Hu et al., 2021; Wang et al., 2018). They also argued that the public auction process affects housing prices mainly through the starvation land supply effect and the market shuffling effect. Additionally, based on data from the National Bureau of Statistics of China, we found that the housing prices started to rise after the policy was implemented in 2002. As shown in Figs. 2 and 3, housing prices remained largely stable before 2002, and then rose sharply with an average annual rate of 11.97 %, approaching 3700 yuan/m^2 in 2008. Housing prices at the provincial level also skyrocketed after 2002, especially in Beijing, Shanghai, Guangdong, Zhejiang, Jiangsu, and Fujian. As housing prices rose, the housing burden on Chinese households became heavier.

3. Theoretical framework

In traditional Chinese culture, a house is considered to be a place for settling down, a symbol of home, and important financial security. Especially in marriage, "buying houses before getting married" has become a social custom in China (Fang & Tian, 2018; Horioka & Terada-Hagiwara, 2017; Wrenn et al., 2019). In the marriage market, men and women have no common social division of labor in the household (Becker & Murphy, 1992; Brines, 1994), and men's social status, economic conditions, and social relationships are valued more (Fisman et al., 2006; Hu & Wang, 2020; Wei & Zhang, 2011). As economically rational people, women's objective function in marriage matching is to select a high-quality spouse to maximize the expected benefits of the family after marriage (Becker, 1974). The house, as an important aspect reflecting the individual's wealth and socioeconomic conditions, is more valued by females (Chu et al., 2020; Hu & Wang, 2020; Wei et al., 2017). In addition, given the gender imbalance in China, a large number of men face a marriage squeeze (Guilmoto, 2012; Liu et al., 2014). Previous research also showed that to enhance their sons' advantage in the marriage market, Chinese parents had to increase their savings for buying houses (Wei & Zhang, 2011).

As mentioned earlier, rising housing prices affect young people's marriages through two main mechanisms. First, in terms of burden effects, the increased cost of marriage brought about by rising housing prices may force young people to make trade-offs between work and marriage. Rising housing prices largely reflect the cost of living in urban areas, leading to uncertainty in young people's expectations for the future (Clark et al., 2020; Ermisch, 1999). The down payment on a house is usually huge for a family, and it will directly squeeze out the short-term consumption of people who want to buy a house. In addition, in the long term, rising housing prices will lead to higher future expenses and put more financial pressure on them, especially for young people, people with low income or high debt (Sung & Qiu, 2020; Wrenn et al., 2019; Zhang & Zhang, 2019). Therefore, considering the limited budget constraints and their career development, young people are likely to prefer to continue working rather than getting married, potentially causing delays in marriage.

Second, from the perspective of wealth effects, the increase in household wealth from rising housing prices may also raise the value of future nominal wealth (Lovenheim & Mumford, 2013; Phang, 2001; Zhang & Zhang, 2019), which may offset the negative impact of increased marriage costs; thus, the wealth effect of rising housing prices may boost young people's marriage intentions. However, the wealth effect is predicated on homeownership (Dvornak & Kohler, 2007; Wang & Zhang, 2021; Wei et al., 2017). For most Chinese households, the proportion of having an additional house for marriage is not high, especially in urban areas (Clark et al., 2020). In contrast to previous studies that emphasized the positive impact of the wealth effect from rising housing prices (e.g., Atalay et al., 2021; Cheng et al., 2020; Dvornak & Kohler, 2007; Lovenheim & Mumford, 2013), for most Chinese young people, they may not accumulate enough wealth to purchase a new house and often need parental support (Hu et al., 2021; Liu et al., 2013). Thus, the burden effect of rising housing prices should be greater from the perspective of marriage decisions in China.

In addition to these two channels, the social resources associated with homeownership, the reflection of social status, and other benefits that homeowners receive may also have an impact on marriages (Hu et al., 2021; Hu & Wang, 2020). Nevertheless, as housing prices rise, it is becoming increasingly difficult to own a home, especially for young people. Since ULTP implementations in 2002, urban housing prices have risen rapidly in China (Hu et al., 2021; Wang et al., 2018). This indicated a substantial increase in housing expenditure and the cost of marriage, which may eventually lead young people to delay their marriage.

4. Methodology and data

4.1. Data source

The microdata involved in our empirical analysis came from the CLDS, the first nationally representative tracking survey in China targeting the labor force, initiated by Sun Yat-sen University since 2012. The CLDS conducted biennial surveys around communities (villages), households, and individual laborers to systematically monitor the social structure of communities (villages) and the changes and interactions of households and individual laborers (Liu et al., 2021; Wang et al., 2017). It adopted a multi-stage and multi-level probability sampling method proportional to the size of the labor force and a rotating sample tracking method. These measures can be better adapted to China's rapidly changing socioeconomic environment and provide high-quality basic data for empirical analysis and policy research (Liu & Liu, 2020; Wang & Zhang, 2021).

The main reason why we used the CLDS is that it takes young and middle-aged laborers as primary survey respondents, which allows for a considerable number of observations that are likely to be married around 2002. More importantly, the CLDS provides the year in which respondents first married. We studied the effect of past housing prices increases on respondents' age at first marriage in a retrospective manner, and the fact that the respondents got married occurred in the past, which makes it inappropriate to observe the same respondent multiple times. Therefore, we used a mixed cross-sectional data, even though the CLDS is a tracking survey.[5] We screened the sample as follows. First, given that our aim is to explore the reasons why Chinese young people delay marriage rather than the possibility of getting married, we removed

[4] Information source: http://www.gov.cn/gongbao/content/2003/content_62586.htm.

[5] That is, only one observation was retained for all respondents, although they may come from different survey years in 2012, 2014, 2016, and 2018.

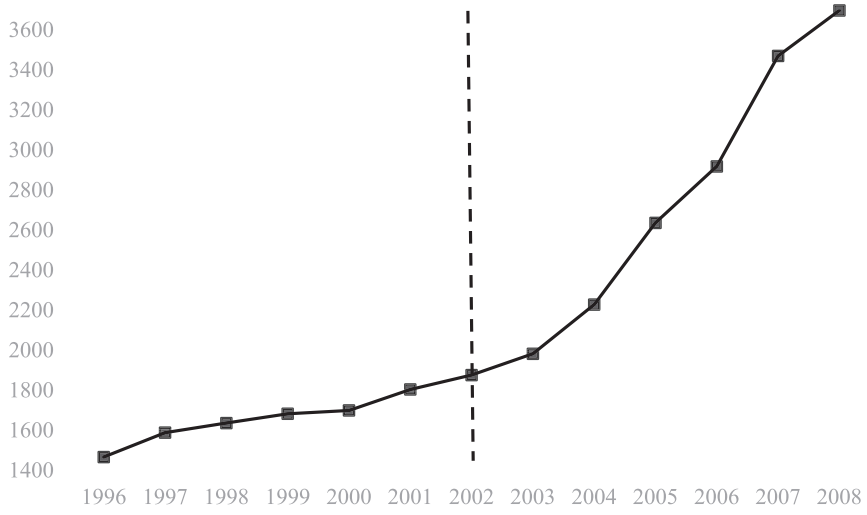


Fig. 2. Average housing prices (yuan/m^2) in China: 1996–2008.
Source: National Bureau of Statistics of China

Fig. 3. Changes in average housing prices (yuan/m^2) in some provinces/municipalities.
Source: National Bureau of Statistics of China

observations that were still unmarried in the survey year.[6] Second, we strictly limited the age of the respondents to 20–40 years old to avoid overestimation from excessive age heterogeneity in the treated and control groups (Hu et al., 2021). Third, we excluded respondents who had not reached the legal age of marriage[7] but got married.[8] Fourth, as previously highlighted, we used ULTP as an identification, making it necessary for us to strictly distinguish between urban and rural samples. Therefore, we further excluded the floating population and respondents whose *hukou* had changed. This practice allowed respondents to keep their birthplace, current *hukou*, and place of residence as consistent as possible.

4.2. Empirical strategy and variables

We aimed to investigate the causal relationship between rising housing prices and marriage delay by using a quasi-natural experiment from the ULTP. As noted above, in China, rural land ownership is collective and houses in rural areas can only be traded within the collective. Therefore, rural housing prices are not affected by the land transaction policies. In contrast, young people living in urban areas are hit by rising housing prices due to this land policy. Moreover, considering that the ULTP has been implemented since 2002, those who were already married before 2002 could not have delayed their marriage because of this policy, as their age at first marriage was already determined before the ULTP implementations. Therefore, these samples can be considered as an eligible control group.

More specifically, as listed in Table 1, the full sample contained four groups: two could be treated as treated groups, one as a control group, and the last one was removed due to our study objectives. In both Group 1 and Group 2, as long as we observed that the respondents were

[6] The unmarriedness does not include divorce and widowhood.
[7] Based on the Marriage Law of China issued in 1980, marriage is not allowed if a man is under 22 or a woman is under 20.
[8] Forming a marriage relationship without reaching the legal age of marriage is unlawful cohabitation in China. Therefore, we excluded this special case.

Table 1
Difference groups in the sample.

Groups	In 2002	In the survey year	Were marriage decisions influenced by the ULTP	Treated/control groups
Group1	Unmarried, under the legal age of marriage	Married	Yes	Treated group
Group2	Unmarried, reaching the legal age of marriage	Married	Yes	Treated group
Group3	Married	Married	No	Control group
Group4	Unmarried	Unmarried	Not clear	Deleted, no observable age at first marriage

married in the survey year, and regardless of whether they had reached the legal age of marriage in 2002, their marriage decisions were influenced by the effect of rising housing prices due to the ULTP. Therefore, both groups should be considered as treated groups. For Group 4, respondents remain unmarried in 2002 and the survey year. Since the explained variable is the age at first marriage, i.e., we need to ensure that all respondents were married (with the age at first marriage) in the survey year, this group are removed. The Group 3, in line with previous studies (Chen et al., 2020; Moser & Voena, 2012), was considered as a control group, as their marital decisions could not be influenced by the ULTP.

In the DID design framework, the first dimension of the difference comes from urban versus rural areas and the second dimension of the difference is whether marriage decisions were strictly influenced by the ULTP. To the best of our knowledge, there are no surveys specifically targeting the marriageable population in China. Therefore, we drew on the idea of cohort DID (Chen et al., 2020; Chen & Zhou, 2007; Duflo, 2001)[9] and employed current surveys to trace whether respondents were affected by the ULTP and their marital status in the past. By comparing the differences between these two dimensions, we could better observe the causal effect of housing prices increase on age at first marriage.

However, a major concern is that marital decisions and age at first marriage are linked, however, our main specification relies on whether individuals' marriage decisions are influenced by the ULTP. That is, a reverse causation problem warrants our vigilance: the age at first marriage may in turn affect individual marriage decisions in 2002. We made efforts to overcome this endogeneity issue as much as possible in two ways. First, in terms of the main pathways through which individual marriage decisions affect the age at first marriage, age, city characteristics, education, and the year characteristics of marriage may play a vital role. Therefore, in the main specification, we strictly controlled for first marriage year fixed effect (FE), city FE, 5-year age FE,[10] and educational attainment. In addition, we added a high-dimensional FE, province-marriage year FE, in the regressions, which helps to capture the province-specific marriage year trends (Chen et al., 2020; Zhao & Guo, 2022). We considered that controlling for these major exogenous variables that influence marital decisions may alleviate to a greater extent the endogeneity caused by reverse causality problems.

Second, we categorized the treated groups according to the exogenous criterion of whether the respondents were under the legal age for marriage in 2002 (i.e., Group 1 and Group 2 in Table 1). Clearly, the attainment of the legal age of marriage is determined by the exogenous variable of the year of birth. More important, this identification strategy has also been widely used in the previous literature (e.g., Chen et al., 2020; Duflo, 2001; Xiao et al., 2017). For instance, in the DID framework, Chen et al. (2020) determined whether individuals' elementary education is influenced by the Send-down Movement of young talent based on the year of birth. Thus, we considered that these measures help to better identify the marriage postponement effect of rising housing prices by using the quasi-natural experiment of ULTP implementation. To begin, the basic estimation model is set as follows:

$$FMA_{its} = \alpha_0 + \alpha_1 Urban_s + \alpha_2 Urban_s * Treated_i + \alpha_3 X_{its} + \varphi(A_5) + \rho_\tau + \theta_t + \delta_c + \omega_{p\tau} + \varepsilon_{its} \quad (1)$$

where subscript i is individual, t represents survey year, and s denotes residence of the respondent (urban or rural). FMA_{its} is the age at first marriage for individual i from residence s in survey year c. $Urban_s$ is dummy variable, which equals 1 for individuals in urban areas and 0 for rural areas, as the ULTP is only implemented in urban China. $Treated_i$ refers to an indicator for the treated group, which is assigned to 1 if individual i is from the treated group listed in Table 1, and 0 for control group, as their age of first marriage is unlikely to be affected by the ULTP.[11] The estimated coefficient α_2 on the interaction term $Urban_s$ * $Treated_i$ is the DID estimator that we closely focused on.

Additionally, X_{its} represents a series of exogenous control variables, including *Gender*, *Education*, *Education of father*, *Education of mother*, *Sibling*, and *Ethnicity*. Specifically, we defined the female and ethnic Han as 1, and 0 otherwise. *Education* is an ordered variable whose assigned integers 1–6 indicate being unschooled, elementary school, junior high school, senior high school (including vocational high school), college or bachelor's degree, and master's degree or Ph.D., respectively. Two other education-rated variables, *Education of father* and *Education of mother*, are also assigned according to the same rules. *Sibling* indicates the number of siblings. Furthermore, in Eq. (1), $\varphi(A_5)$ is the 5-year age FE, ρ_τ and θ_t are marriage year FE and survey year FE, respectively, as mixed cross-sectional data are used in this study. δ_c denotes the city FE and ω_{pc} is the province-marriage year FE. Controlling for these FEs are beneficial to further mitigate potential endogeneity issues. Last, ε_{its} represents the random error term.

At this point, what needs to be tested is to determine whether the treated group and control group had parallel trends in their age at first marriage prior to the ULTP implementations, which is the prerequisite for employing the DID model. Housing prices in urban China remained relatively stable before the land transaction policy, but began to soar henceforth (as shown in Fig. 2). We hypothesized that, before the implementation of the policy, there is no significant difference in the change of age at first marriage for young people in urban and rural areas. With the implementation of the ULTP since 2002, housing prices have risen, and the age at first marriage of young people has also increased. To test the hypothesis, we conducted the following regressions for parallel trend tests:

$$FMA_{its} = \alpha_0 + \alpha_1 Urban_s + \sum_\kappa \zeta_\kappa Urban_s * Treated_i + \alpha_3 X_{its} + \varphi(A_5) + \rho_\tau + \theta_t + \delta_c + \omega_{p\tau} + \varepsilon_{its} \quad (2)$$

where κ denotes the relative time from being implemented the ULTP. If κ

[9] One difference between our empirical strategy and traditional cohort DID is that individuals do not make marital decisions at a fixed age, so it is difficult to identify control and control groups at the cohort level. However, similarly, previous literature using cohort DID has typically regarded groups that are not (or less) affected by policies or events as control groups (Chen et al., 2020; Chen & Zhou, 2007).

[10] To avoid completely collinearity, we could not control for the age FE; therefore, a 5-year age FE was added as a proxy. Specifically, the grouping criteria for age were 20–25 years old, 26–30 years old, 31–35 years old, and 36–40 years old.

[11] Since we controlled for the marriage year FE, the variable *Treated* does not require additional control in the models.

equals to 0, it means that the year the ULTP was implemented coincides with the year of first marriage. Thus, a negative (positive) κ indicates that the year of first marriage is greater (less) than the year the ULTP was implemented in 2002. The significance of coefficients ζ_κ indicates whether there is a difference between the treated group and the control group. The parallel trend is satisfied if ζ_κ is insignificant when κ is negative.

Furthermore, as discussed earlier, based on whether respondents were of legal age to marry at the time of ULTP implementation, we further splitted the full treated group into two subgroups to mitigate potential endogeneity. The specific model is shown in Eq. (3):

$$FMA_{its} = \beta_0 + \beta_1 Urban_s + \beta_2 Urban_s * Treated1_i + \beta_3 Urban_s * Treated2_i + \beta_4 X_{its} + \varphi(A_5) + \rho_\tau + \theta_t + \delta_c + \omega_{p\tau} + \varepsilon_{its} \quad (3)$$

Specifically, $Treated1_i$ and $Treated2_i$ denote the dummy variables when Group 1 and Group 2 are considered as treated groups, respectively. Similarly, coefficients β_2 and β_3 on two interactions, $Urban_s$ * $Treated1_i$ and $Urban_s$ * $Treated2_i$ are the DID estimators in this model. The significantly positive estimates could further indicate that the age of first marriage among urban young people was delayed after ULTP implementations.

4.3. Summary statistics

The summary statistics of variables are shown in Table 2. Our final sample includes 4619 married observations in survey year, of which 2439 respondents are female and 2180 are male. In the full sample, we found that the average age at first marriage was 24.2464 years old, 59.64 % of observations as the treated group, and 24.92 % lived in the urban areas and had an urban *hukou*. The average age at first marriage for females is 23.1599 years old, while this age for males is greater at 25.4619 years old. In Panel A of Table 3, we compared the average age at first marriage from two perspectives: treated group and urban group. It is clear that respondents in the treated and urban groups had a higher age at first marriage, both in the full sample and in the gender-specific subsamples. More importantly, the *t*-test statistics for these differences are all significant at the 1 % level. In Panel B of Table 3, based on the *t*-tests, we found that after ULTP implementations, the average housing prices based on national level statistics increased by an average of 1004 yuan/m^2 and those at the provincial level increased by 934 yuan/m^2,[12] which is similar with the information provided in Figs. 2 and 3.

5. Results

5.1. The effect of rising housing prices on the age at first marriage

As mentioned in the empirical strategy, the ULTP implemented in 2002 led to a substantial increase in housing prices in urban areas, while rural areas have not been significantly affected. Based on this quasi-natural experiment, we used the DID method to explore the effect of rising housing prices on the age at first marriage in China. Table 4 presents the baseline results, and all standard errors have been cluster-corrected at the individual level. In column (1), we controlled for a series of FEs, including 5-year age, marriage year, survey year, city, and province-marriage year FEs, finding the estimated coefficient on *Urban* * *Treated* is significantly positive.[13] Moreover, all control variables are included in column (2), and the estimates of the core explanatory variable remain significantly positive at the 5 % level, indicating that the effect of rising housing prices from the ULTP did lead to a drop in the age at first marriage for urban youths. In the preferred specification, column (2), we found that the coefficient on *Urban* * *Treated* is 0.2280, suggesting that compared to young people in rural areas, rising housing prices have delayed the age at first marriage for urban young people by an average of 0.2280 years old, equivalent to 2.736 months and 0.0258 standard deviations. According to Table 3, the ULTP implementation is assumed to result in a 1000 yuan increase in the average housing prices, without considering other factors. That is, for each thousand yuan increase in the average housing prices in China, the effect on the delay in the age at first marriage is about 2.736 months.

In China, before getting married, the bride prices, private cars, and houses provided by males are usually considered as necessities (e.g., Cheung et al., 2020; Fang & Tian, 2018; Wei & Zhang, 2011; Wrenn et al., 2019), especially the house. Therefore, due to this marriage culture, the importance and demand of housing for males is higher than that of females. Considering the above differences, we examined the effect of rising housing prices on the age at first marriage for males and females, respectively. As shown in columns (3) of Table 4, we found that the coefficient on *Urban* * *Treated is* insignificant; however, results from column (4) indicate that the rising housing prices due to ULTP significantly delayed the age at first marriage for urban males by 0.4372 years old, equivalent to 5.2464 months and 0.0537 standard deviations. As expected, these results suggest that the delayed effect of rising housing prices on the first marriage among young males is even more pronounced in China.

The baseline results support that the burden effect of rising housing prices exceeds the wealth effect in China, implying an overall crowding out effect of rising housing prices on the marriages of young people. Rising housing prices put considerable pressure on buying a house before marriage. Owning a house almost determines the attractiveness of young males in the marriage market (Chu et al., 2020; Hu et al., 2021; Wei & Zhang, 2011). Thus, rising housing prices delayed the age at first marriage for males, as many families are unable to own a house.

5.2. Parallel trend tests

DID estimations are based on the premise that the parallel trend assumption must be satisfied. The estimation results of Eq. (3) are shown in Fig. 4. Specifically, in Fig. 4(a), we found that all coefficients are insignificant before ULTP implementations, indicating that the parallel trend in the DID estimates is statistically satisfied for the full sample (column (2) in Table 4). Correspondingly, in the fourth period after the ULTP implementations, the estimated coefficient becomes positive and significant at the 5 % level.

In Fig. 4(b), the estimated coefficients are insignificant before and after the policy implementations, suggesting that women's age at first marriage may not be significantly affected by the increase in housing prices from the ULTP implementations. Additionally, Fig. 4(c) shows the results of parallel trends in DID estimation for the male sample. Similarly, there was no significant difference between the treated and control groups before the ULTP was implemented, suggesting that the parallel trend assumption is met for the male subsample (column (4) in Table 4). In contrast, after policy implementations, the statistically significant difference is highlighted, indicating that the effect of rising housing prices from the ULTP led to a later age of first marriage for males.

[12] Although city-level housing prices data are available as far as 1999, only provincial capitals are included. Respondents in provincial capitals accounted for less than 10 % of our sample. We therefore did not show city-level differences in housing prices before and after ULTP implementations. In addition, the earliest officially published provincial-level housing prices data is 2000. Thus, the years prior to the ULTP cover only two years, 2000 and 2001.

[13] The observations in Table 4 are inconsistent with those in the descriptive statistics because we controlled for some FEs and used the "reghdfe" command, which lost some degrees of freedom (Baker et al., 2019; Wee et al., 2018).

Table 2
Summary statistics of variables.

Variables	Full sample			Female			Male		
	N	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.
Marriage age	4619	24.2464	3.1681	2439	23.1599	2.8038	2180	25.4619	3.1106
Urban * Treated	4619	0.1509	0.3580	2439	0.1271	0.3332	2180	0.1775	0.3822
Treated	4619	0.5964	0.4907	2439	0.5305	0.4992	2180	0.6702	0.4703
Urban	4619	0.2492	0.4326	2439	0.2366	0.4251	2180	0.2633	0.4405
Gender	4619	0.5280	0.4993	2439			2180		
Education of father	4619	1.9667	0.9864	2439	1.9569	0.9827	2180	1.9775	0.9905
Education of mother	4619	2.4031	1.0688	2439	2.3907	1.0672	2180	2.4170	1.0707
Education	4619	3.2158	1.1412	2439	3.1128	1.2036	2180	3.3312	1.0553
Sibling	4619	2.3308	1.7044	2439	2.4723	1.6821	2180	2.1725	1.7156
Ethnicity	4619	0.8868	0.3169	2439	0.8840	0.3203	2180	0.8899	0.3131

Table 3
Statistics for age at first marriage and housing prices in different dimensions.

Panel A. Means of age at first marriage

	Treated = 0	Treated = 1	Diff. (*t*-test)	Urban = 0	Urban = 1	Diff. (*t*-test)
Full sample	22.6760	25.3089	−2.6329***	23.7497	25.7428	−1.9931***
Female	22.0210	24.1677	−2.1467***	22.7035	24.6326	−1.9290***
Male	23.7191	26.3196	−2.6006***	24.9626	26.8589	−1.8962***

Panel B. Average housing prices

	Treated = 0	Treated = 1	Diff. (*t*-test)
At the national level	1642.8	2686.8	−1004.0***
At the provincial level	1751.5	2685.5	−934.0***

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively.

Table 4
The effects of rising housing prices on the age at first marriage.

	Full sample		Female	Male
	(1)	(2)	(3)	(4)
Urban * Treated	0.2063*	0.2280**	0.1726	0.4372***
	(0.1089)	(0.1061)	(0.1670)	(0.1609)
Urban	0.2446***	0.2888***	0.3873***	0.2489*
	(0.0816)	(0.0882)	(0.1367)	(0.1329)
Gender		−0.5289***		
		(0.0435)		
Education of father		0.0070	−0.0003	−0.0317
		(0.0308)	(0.0462)	(0.0500)
Education of mother		−0.0208	0.0067	−0.0471
		(0.0259)	(0.0395)	(0.0416)
Education		0.0599**	0.0730**	0.0528
		(0.0275)	(0.0315)	(0.0368)
Sibling		0.0683***	0.0592***	0.0774***
		(0.0147)	(0.0218)	(0.0246)
Ethnicity		0.0497	0.1767	0.1433
		(0.1162)	(0.1769)	(0.1929)
5-Year age FE	Yes	Yes	Yes	Yes
Marriage year FE	Yes	Yes	Yes	Yes
Survey year FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Province-marriage year FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.8366	0.8429	0.8115	0.8228
N	4531	4531	2320	2062

Notes: ***, **, * indicate significance at the levels of 1 %, 5 %, and 10 %, respectively. Standard errors clustered at the individual level are reported in parentheses.

5.3. The effects of different treated groups

As discussed in Table 1, the treated group can be divided into two groups based on whether respondents reached the legal age of marriage in 2002, which helps to weaken the association between marriage decision and age at first marriage to reduce endogeneity. Now we further explored the effects of different treated groups, and the results are represented in Table 5. In Panel A, we regarded Group 1 as the only treated group, finding that the DID estimates are significantly positive in columns (1) and (3). These results further emphasize the marriage postponement effect of the housing prices increase effect due to the ULTP, especially for males. The results in Panel B show that all coefficients on *Urban * Treated2* are insignificant by treating Group 2 as the only treated group. Similar results were found in Panel C, and both Groups 1 and 2 are considered as treated groups.

Clearly, as shown in Figs. 2 and 3, housing prices in China maintained their upward trend during our study period. Compared to Group 2, young people in Group 1 have to bear higher housing prices when they get married, resulting in greater burden effects (Assaad et al., 2021; Wrenn et al., 2019). By distinguishing the different treated groups, we found the effect of rising housing prices from the ULTP mainly hindered the marriage of respondents in Group 2, who had not yet reached the legal age of first marriage in 2002. Thus, these estimates further confirm the marriage postponement effect of rising housing prices, as highlighted in Table 4. Furthermore, our results also imply that from the perspective of housing prices, the phenomenon of delayed marriage or late marriage and fewer children in China may persist for a long time, and some later-born individuals are more likely to delay marriage or even not get married because of such high housing prices in urban areas (Hu et al., 2021).

5.4. Rising housing prices and the cost of first marriage

As discussed in Section 3, we considered that one of the main mechanisms by which rising housing prices delay marriage comes with the elevated cost of marriage in China, especially for males (Fang & Tian, 2018; Wei & Zhang, 2011). In the CLDS database, some respondents reported the spending on their first marriage (10,000 yuan, RMB), which helps us to further validate this mechanism. In Table 6, we

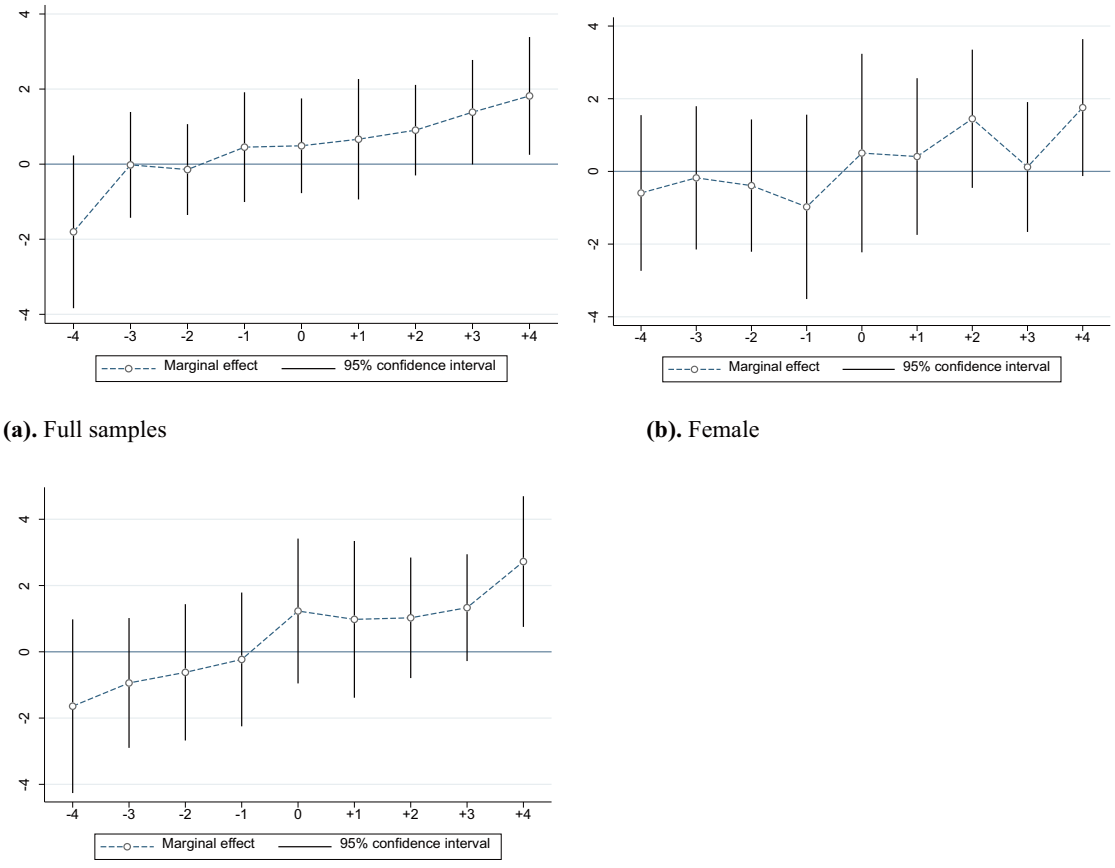


Fig. 4. Parallel trend tests.

Table 5
Different treated groups.

	(1)	(2)	(3)
	Full sample	Female	Male
Panel A. Group1 is the treated group			
Urban * Treated1	0.2618*	0.0577	0.5784**
	(0.1438)	(0.2356)	(0.2261)
Urban	0.4173***	0.4097***	0.3605**
	(0.0964)	(0.1465)	(0.1491)
N	3275	1693	1428
Panel B. Group2 is the treated group			
Urban * Treated2	0.1181	0.1016	0.1880
	(0.1287)	(0.2081)	(0.2037)
Urban	0.4583***	0.5306***	0.4039***
	(0.0952)	(0.1482)	(0.1483)
N	3010	1650	1206
Panel C. Group1 and group2 are treated groups			
Urban * Treated1	0.2286*	0.0412	0.4331**
	(0.1309)	(0.2092)	(0.1982)
Urban * Treated2	0.1318	0.2003	0.0635
	(0.1201)	(0.1898)	(0.1819)
Urban	0.3966***	0.4633***	0.4025***
	(0.0873)	(0.1349)	(0.1349)
N	4531	2320	2062

Notes: ***, **, * indicate significance at the levels of 1 %, 5 %, and 10 %, respectively. Baseline controls and FEs are added in all regressions in three Panels. Standard errors clustered at the individual level are reported in parentheses.

Table 6
Rising housing prices and the cost of first marriage.

	(1)	(2)
	Full sample	Male
Urban * Treated	6.8834*	7.5285*
	(4.1401)	(4.5585)
Control variables	Yes	Yes
5-Year age FE	Yes	Yes
Marriage year FE	Yes	Yes
Survey year FE	Yes	Yes
City FE	Yes	Yes
Province-marriage year FE	Yes	Yes
N	1235	1161

Notes: ***, **, * indicate significance at the levels of 1 %, 5 %, and 10 %, respectively. Standard errors clustered at the individual level are reported in parentheses.

found that the estimated coefficients on *Urban * Treated* are significantly positive in both columns (1) and (2), despite the fact that the observations are only a quarter of the baseline regressions.[14] More specifically, the results show that compared to rural men, rising housing prices have increased the average first marriage cost for urban men by about 75,285 yuan. Taken together, these estimates suggest that the burden effect of rising housing prices has increased the cost of marriage substantially, which is consistent with the findings in previous studies (Cheung et al., 2020; Hu et al., 2021; Wei & Zhang, 2011). Thus, we provided evidence

[14] Only 74 women reported spending on their first marriage.

that by increasing the cost of marriage, rising housing prices delay the age of first marriage for young people in China.

5.5. Robustness checks

Although the baseline results suggested that rising housing prices significantly delay the age at first marriage by using the exogenous shock from the ULTP, these results may be challenged by some confounders. In this subsection, we conducted several robustness tests to verify the reliability of our main results.

5.5.1. Placebo tests

Placebo tests are widely used in robustness checks of DID estimations (e.g., Chen et al., 2020; Liao et al., 2021; Zheng et al., 2020). Here we performed placebo tests by constructing counterfactuals. A "pseudo-policy variable" was constructed based on the number of respondents in the sample that were affected by the ULTP. Fig. 5 shows the kdensity of estimates of "pseudo-policy variable" and corresponding *p*-values by randomly simulating 500 times. In Fig. 5(a), estimated coefficients on "pseudo policy" are mostly concentrated near zero, and most of the estimated p-values are greater than 0.05 (insignificant at the 95 % level). In the male subsamples, similar results can be obtained from Fig. 5(b). These results indicate that our baseline estimates are unlikely to be obtained by chance.

5.5.2. Eliminating potential disruptive policy

We considered three main disruptive policies: college enrollment expansion policy (CEEP), one-child policy, and accession to the WTO in 2001.

First, we tried to eliminate the effect of CEEP on baseline results. Since 1999, China has drastically expanded the scale of higher education. With the implementation of CEEP, the average education level of the young population has rapidly increased (Li et al., 2017). Such a large-scale educational expansion inevitably impacts the marriage market. Previous literature also confirmed that receiving higher education has a significant impact on the individual's decision to marriage delays (e.g., Goldstein & Kenney, 2001; Kalmijn, 2013; Stevenson & Wolfers, 2007). To weed out the potential effects of CEEP, we excluded the samples that were born after 1982 (18 years old in 1999) and were admitted to universities or colleges, that is, these samples affected by the CEEP. In addition, in line with Li et al. (2019), we controlled for the number of universities and the proportion of college students in cities. Moreover, considering the spillover effects of human capital (Gennaioli et al., 2013; Liu, 2007), the CEEP could benefit some individuals who did not attend colleges, making them delay marriage (Che & Zhang, 2018). Therefore, we excluded highly skilled laborers based on the type of employment provided by the CLDS,[15] as they are more likely to benefit from the spillover effects of human capital (Moretti, 2004). In Table 7, these results in three Panels are basically consistent with our main specifications, with significantly positive coefficients of *Urban* * *Treated* in columns (1) and (3). Thus, after considering the impact of CEEP, the effect of rising housing prices on marriage delays is still significant.

Second, we considered the impact of the one-child policy. Numerous studies were devoted to examining the relationship between population policies and marriage delays (e.g., Bongaarts & Greenhalgh, 1985; Cai, 2010; Huang et al., 2020). In China, the one-child policy was formally implemented in 1982 in response to explosive population growth, and had a profound impact on the population structure (Liu, 2014; Zhong, 2014). It should be emphasized that the ULTP examined in this paper was implemented in 2002, and at this time, the newborns affected by the one-child policy (born after 1982) are just about to or have entered the marriageable age. To eliminate potential interference of one-child policy, we only kept samples that were not affected by the one-child policy, that is, those born before 1982. In Panel A of Table 8, we found that the DID estimates remained significantly positive in columns (1) and (3), while there is an increase in magnitude. In addition, given the impact of the one-child policy on the number of siblings and the fact that minority populations are hardly affected by the one-child policy (Bulte et al., 2011; Cai, 2010; Zhong, 2014), in Panels B and C, we controlled for the interaction effects of one-child policy[16] with ethnicity and siblings, respectively. Similarly, these estimates are also consistent with the baseline results in Table 4, implying that the effect of marriage delays on housing prices increases due to ULTP was not challenged by the one-child policy.

Third, another potentially disturbing event comes from China's accession to the WTO in 2001. WTO accession affected productivity and firms' markups by expanding international trade markets, increasing competition, and reducing input costs (Brandt et al., 2019; Lu & Yu, 2015). This shock may increase the opportunity costs of marriage, which also contributes to delaying marriage through channels other than rising housing prices. The following efforts are used to mitigate the potential interference from the WTO accession. First, coastal regions, are known to be more affected by the WTO accession. Thus, we excluded all respondents living in coastal provinces and municipalities. In Panel A of Table 9, we found that the DID estimates remain significantly positive in columns (1) and (3). Second, we controlled for the interaction between reaching the legal age of marriage and WTO accession,[17] considering the time difference between the marriage decision and the WTO accession. Third, we excluded all respondents who worked in international trade industries, given that the CLDS provided industry sectors to which laborers' job belongs. In Panels B and C of Table 9, not surprisingly, the DID estimates show that our main results are satisfactory for capturing the influence of China's accession to the WTO in 2001.

5.5.3. Excluding special samples

In the main specification, although we excluded all floating population and *hukou* switchers to ensure comparability between urban and rural observations. However, there is still a special category of the sample that may interfere with our empirical results, namely, rural people who may move to urban areas to buy houses in the future. Fortunately, the CLDS questionnaire addresses whether rural respondents have a desire to buy a home in the urban areas in the next five years. As presented in Panel A of Table 10, we excluded these special samples, finding that this change has little effect on the main results.

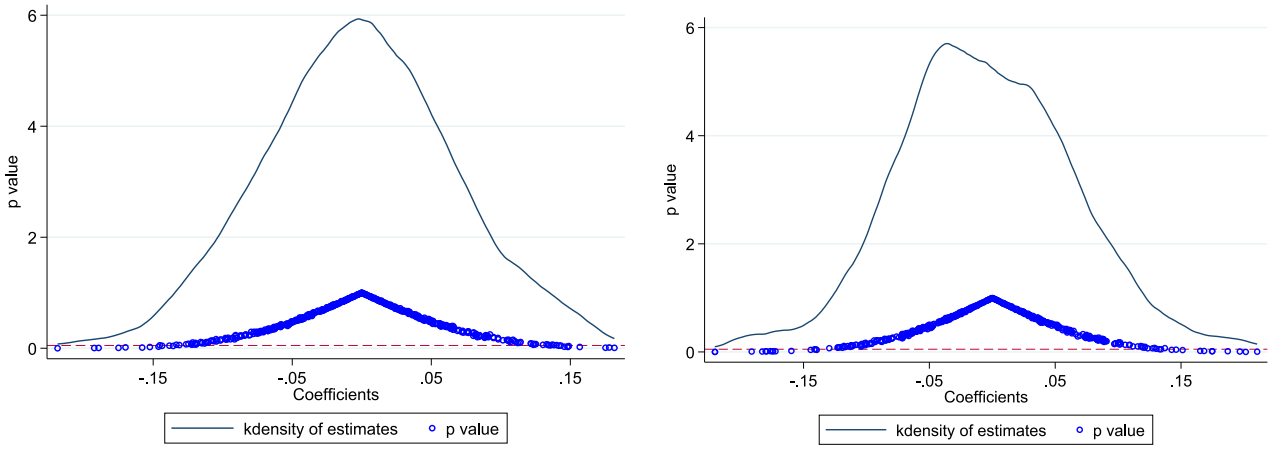
Moreover, prior to the marketization of real estate in China, most of the public sector employers, such as government departments, public institutions, and state-owned enterprises, usually enjoyed the benefits of being allocated a house or buying a house below market prices (Wang, 2000). Even now, some local governments continued to offer real estate incentives to public sector employees and senior talents.[18] For these reasons, if individuals or their parents are public sector employees, rising housing prices may not have a significant impact on their behavior

[15] We considered some technology or capital-intensive industries in which laborers are employed as high-skilled employment, such as finance, social services, education, law, and science and technology.

[16] Respondents who were affected by the one-child policy were set to 1, and 0 otherwise.

[17] As before, respondents who reached the legal age of marriage in 2001 are assigned to 1, and 0 otherwise. For WTO accession, all coastal provinces and municipalities are set to 1, and 0 otherwise.

[18] By querying the website of local governments and the working documents of the Human Resources and Social Security Bureau of cities in China, these preferential policies can be summarized into three categories: (1) directly donating a house (including private property rights); (2) providing a subsidy for the purchase of a house (usually a large one-time subsidy); (3) providing new housing options in some specific communities and purchasing houses at a price lower than the market price.



(a). Full samples **(b).** Male

Fig. 5. Kdensity of estimates of "pseudo policy" and corresponding *p*-values (random 500 times).

Table 7
College enrollment expansion.

	(1)	(2)	(3)
	Full sample	Female	Male
Panel A. Excluding the sample born after 1982 and attending university			
Urban * Treated	0.2772**	0.2527	0.4815***
	(0.1166)	(0.1870)	(0.1724)
N	4169	2128	1895
Panel B. Controlling for the number of universities and the proportion of college students in cities			
Urban * Treated	0.2398**	0.3135	0.4431**
	(0.1175)	(0.1948)	(0.1799)
N	3692	1795	1744
Panel C. Excluding highly skilled labors			
Urban * Treated	0.2065*	0.1593	0.4062**
	(0.1077)	(0.1687)	(0.1639)
N	4374	2207	2008

Notes: ***, **, * indicate significance at the levels of 1 %, 5 %, and 10 %, respectively. Baseline controls and FEs are added in all regressions in three Panels. Standard errors clustered at the individual level are reported in parentheses.

and decision-making, especially when it comes to marriage. That is, given the housing benefits of public sector employment, they are not affected by the marriage delay effect of rising housing prices. Therefore, we excluded respondents employed in the public sectors or whose parents are employed in the public sectors.[19] We expected that the DID estimates may be improved by excluding this particular sample. In Panels B and C of Table 10, not surprisingly, the positive impact of rising housing prices from the ULTP on the age at first marriage is still significant among the full sample and the male subsample. More importantly, compared to the main results in Table 4, the coefficients of *Urban * Treated* do show somewhat larger in magnitude.

6. Further studies

6.1. Housing ownership and marriage delays

In the main identification, we employed a quasi-natural experiment of ULTP implementations as an exogenous proxy for rising urban housing prices, however ignoring the direct evidence of housing ownership. However, the CLDS database does not contain much information on respondents' housing, so we used another database, the China Household Finance Survey (CHFS), for further analysis. The CHFS is a continuous tracking survey[20] of Chinese households and individuals that records a wealth of household economic and financial information in addition to providing some demographic characteristics (Zhao et al., 2022).

In this subsection, our empirical strategy is to focus on the unmarried group and to compare their future marriage decisions and age at first marriage for those who have purchased a house with those who have not. Specifically, we took a sample of unmarried individuals[21] in the baseline wave of the CHFS. Then, we observed their marital status in the subsequent four waves. We constructed two explained variables, *Married* and *Marriage age*. The former is a binary variable, set to 1 if the respondent was married in later waves, and 0 otherwise; the latter is a continuous variable involving the respondent's age at first marriage.[22] The results are shown in Table 11.[23] We found that housing ownership significantly increases the likelihood of individuals marrying and

[19] In the CLDS questionnaire, the public sectors include party and government organs, people's organizations, state-owned and collective institutions, state-run enterprises, and collective enterprises.

[20] Currently, the CHFS has published data for 5 waves of 2011, 2013, 2015, 2017 and 2019.

[21] As with the previous sample selection rules, we limited the age of respondents to be in the 20–40 age range and excluded all samples that were not within the legal age of marriage. Additionally, divorce and widowhood are not part of the unmarried sample here.

[22] It is important to note that there are many missing values in the CHFS for the year of marriage, we therefore calculated the age at first marriage based on marriage information and year of birth, as we only included a sample that was unmarried in the baseline wave in 2011. However, because the CHFS is a biennial survey, there is a risk that the calculation of age at first marriage may be skewed. For example, we observed that respondents were married in the second wave in 2013, but we could not determine whether respondents were married in 2012 or 2013. In this case, we used 2013 as the year of first marriage.

[23] Control variables in Table 11 included gender, *hukou*, education, self-rated health, household income, and household debt. In addition, five-year age FE and province FE are also added in all regressions.

Table 8
One-child policy.

	(1)	(2)	(3)
	Full sample	Female	Female
Panel A. Keeping samples that are not affected by one-child policy (born before 1982)			
Urban * Treated	0.2115**	0.1754	0.4150**
	(0.1074)	(0.1699)	(0.1621)
N	3002	1533	1328
Panel B. Controlling for the interaction between one-child policy and ethnicity			
Urban * Treated	0.2377*	0.1501	0.4549**
	(0.1349)	(0.1260)	(0.2086)
N	4531	2320	2062
Panel C. Controlling for the interaction between one-child policy and sibling			
Urban * Treated	0.2367**	0.1646	0.4542***
	(0.1045)	(0.1366)	(0.1679)
N	4531	2320	2062

Notes: ***, **, * indicate significance at the levels of 1 %, 5 %, and 10 %, respectively. Baseline controls and FEs are added in all regressions in three Panels. Standard errors clustered at the individual level are reported in parentheses.

Table 9
Accession to the WTO in 2001.

	(1)	(2)	(3)
	Full sample	Female	Male
Panel A. Deleting all coastal provinces and municipalities			
Urban * Treated	0.2143*	0.1078	0.4072*
	(0.1221)	(0.2639)	(0.2383)
N	2110	1059	956
Panel B. Controlling for the interaction between reaching the legal age of marriage and WTO accession			
Urban * Treated	0.1778*	0.1032	0.3674**
	(0.1059)	(0.1660)	(0.1645)
N	4531	2320	2062
Panel C. Excluding samples engaged in foreign trade industry			
Urban * Treated	0.1948*	0.0669	0.4057**
	(0.1117)	(0.1774)	(0.1680)
N	4210	2101	1955

Notes: ***, **, * indicate significance at the levels of 1 %, 5 %, and 10 %, respectively. Baseline controls and FEs are added in all regressions in three Panels. Standard errors clustered at the individual level are reported in parentheses.

decreases the age of first marriage, which is in line with prior research (Hu & Wang, 2020). Therefore, using continuous tracking data from the CHFS, we provided evidence that housing ownership affects marriage decisions and age at first marriage, indirectly illustrating the effect of rising housing prices on marriage postponement in China.

6.2. Moderating effects

We explored some moderating effects of rising housing prices on marriage delays from four perspectives: housing prices, city level, education, and siblings.

6.2.1. Housing prices

In Section 5, we exploited a quasi-natural experiment of ULTP to identify the effect of rising housing prices on marriage delays among young people. However, limited by the availability of housing prices data, the net effect of the ULTP on housing prices increases is not clear (Hu et al., 2021). Therefore, based on provincial housing prices data, we further analyzed the moderating effects by housing prices.

Table 10
Excluding special samples.

	(1)	(2)	(3)
	Full sample	Female	Male
Panel A. Rural people who will buy houses in the city in the future			
Urban * Treated	0.2425**	0.1968	0.4516***
	(0.1118)	(0.1827)	(0.1703)
N	3873	1837	1869
Panel B. Public sector employers			
Urban * Treated	0.2644**	0.1987	0.5838***
	(0.1096)	(0.1790)	(0.1670)
N	4072	2006	1892
Panel C. Parents in the public sectors			
Urban * Treated	0.2369**	0.1914	0.4646***
	(0.1073)	(0.1677)	(0.1619)
N	4442	2249	2035

Notes: ***, **, * indicate significance at the levels of 1 %, 5 %, and 10 %, respectively. Baseline controls and FEs are added in all regressions in three Panels. Standard errors clustered at the individual level are reported in parentheses.

More specifically, we used the differences and growth rate of provincial housing prices before and after ULTP implementations as moderating variables, respectively. The change and growth in housing prices may reflect differences in the effects produced by the ULTP. Such differences in outcomes caused by events or policy implementations are also widely used in DID models to identify the causality (e.g., Che & Zhang, 2018; Chen & Zhou, 2007; Chen et al., 2020). The results are reported in Panel A of Table 12. In Panel A1, we found that the coefficient on *Urban * Treated * Price1* is significantly positive in column (3). In addition, in Panel A2, the estimated coefficients on *Urban * Treated * Price2* are both significantly positive in columns (1) and (3), suggesting that the greater the housing prices driving effect from the ULTP, the more pronounced the delayed effect on marriage. These estimates further confirm our conclusion that for first marriages of Chinese young people, especially for men, the burden effect from rising housing prices is higher than the wealth effect.

6.2.2. City level

In our previous analysis, we compared the difference between urban and rural youths' age at first marriage due to their different housing burden. However, housing price growth is not consistent across Chinese cities. Generally, housing prices are higher in big cities and people are more burdened with housing (Su et al., 2020; Yuan et al., 2020), which may lead to more severe marriage postponement. For instance, according to the Shanghai Statistical Report on Marriage Registration in 2018, the age of first marriage for men and women reached 30.65 and 28.81 respectively.[24] Moreover, for urban populations, those living in large and small cities cannot be strictly distinguished in our sample. Housing price increases are much smaller in small cities or counties, which may be closer to the rural areas. Therefore, it is necessary to test the moderating effects brought by the city level.

China's large cities specifically include three levels of cities: provincial capital, province-level municipality, and sub-provincial cities. We defined the dummy variable *Large city*[25] and constructed an interaction term *Urban * Treated * Large city*. The results are shown in Panel C of Table 12. We found that the coefficients of this interaction term are significantly positive in all columns. Inconsistent with our main results, the females in large cities also significantly delayed their age of first

[24] Shanghai is one of the cities with the highest housing prices in China.
[25] The variable *Large city* is set to 1 if the respondents lived in a large city, and 0 otherwise.

Table 11
Housing ownership and marriage delays by using the CHFS data.

	(1)	(2)	(3)	(4)	(5)	(6)
	Married			Marriage age		
	Full sample	Female	Male	Full sample	Female	Male
Housing ownership	0.1450***	0.1322**	0.1506***	−0.8558**	−0.2298	−1.4173***
	(0.0398)	(0.0540)	(0.0521)	(0.3605)	(0.4322)	(0.5168)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
5-Year age FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2084	1003	1081	1427	678	749

Notes: ***, **, * indicate significance at the levels of 1 %, 5 %, and 10 %, respectively. Standard errors clustered at the individual level are reported in parentheses.

Table 12
Moderating effects by housing prices and city level.

	(1)	(2)	(3)
	Full sample	Female	Male
Panel A. Housing prices			
Panel A1: the difference in average housing prices before and after the ULTP			
Urban * Treated * Price1	0.0909	0.1043	0.1110*
	(0.0889)	(0.1785)	(0.0669)
Urban * Treated	0.2543	0.2064	0.5737*
	(0.2033)	(0.4273)	(0.3232)
N	4531	2320	2062
Panel A2: the growth rate of average housing price before and after ULTP			
Urban * Treated * Price2	0.0484**	0.0019	0.1194***
	(0.0210)	(0.0371)	(0.0400)
Urban * Treated	0.1457	0.1538	0.3616**
	(0.2184)	(0.2583)	(0.1767)
N	4531	2320	2062
Panel B. City level			
Urban * Treated * Large city	0.0372*	0.0427**	0.1384***
	(0.0207)	(0.0205)	(0.0486)
Urban * Treated	0.1520	0.0725	0.4456**
	(0.1472)	(0.2287)	(0.2026)
N	4531	2320	2062

Notes: ***, **, * indicate significance at the levels of 1 %, 5 %, and 10 %, respectively. Baseline controls and FEs are added in all regressions. Standard errors clustered at the individual level are reported in parentheses.

Table 13
Moderating effects by higher education and number of siblings.

	Higher education			Number of siblings		
	(1)	(2)	(3)	(4)	(5)	(6)
	Full sample	Female	Male	Full sample	Female	Male
Urban * Treated * Higher education	0.4328	0.3988**	0.3450			
	(0.2918)	(0.1866)	(0.2732)			
Urban * Treated * Sibling				0.0281	0.0689	0.1173*
				(0.0206)	(0.0853)	(0.0632)
Urban * Treated	0.3163***	0.3171	0.3690**	0.3150**	0.2559	0.3521**
	(0.1205)	(0.2078)	(0.1556)	(0.1243)	(0.1886)	(0.1737)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
5-Year age FE	Yes	Yes	Yes	Yes	Yes	Yes
Marriage year FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey year FE	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-marriage year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	4531	2320	2062	4531	2320	2062

Notes: ***, **, * indicate significance at the levels of 1 %, 5 %, and 10 %, respectively. Standard errors clustered at the individual level are reported in parentheses.

Table 14
The effect of rising housing prices on fertility.

	Number of children		Fertility intentions	
	(1)	(2)	(3)	(4)
	Full sample	Female	Full sample	Female
Urban * Treated	−0.1857**	−0.1804**	−0.0680**	−0.0654**
	(0.0816)	(0.0817)	(0.0297)	(0.0277)
Control variables	Yes	Yes	Yes	Yes
5-Year age FE	Yes	Yes	Yes	Yes
Marriage year FE	Yes	Yes	Yes	Yes
Survey year FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Province-marriage year FE	Yes	Yes	Yes	Yes
N	2260	2237	1476	1360

Notes: ***, **, * indicate significance at the levels of 1 %, 5 %, and 10 %, respectively. Standard errors clustered at the individual level are reported in parentheses.

marriage due to rising housing prices. The possible reasons are as follows. On the one hand, although the females are not required to own a house when getting married under the traditional marriage culture, rising housing prices made males owning a house a scarce resource. The females need to spend more time and search costs in the marriage market (Becker, 1974; Lichter et al., 1991), especially in big cities. On the other hand, because of the high prices of housing in big cities, some women need to take on additional housing burdens for marriage, such as sharing the down payment and mortgage repayment, which also delays the age at first marriage for females. All in all, these estimates provide evidence that compared with small and medium-sized cities, the increase in housing prices has a stronger effect on the postponement of age at first marriage in large cities.

6.2.3. Education and siblings

Considering the impact of education on marriage delay, especially higher education (Goldstein & Kenney, 2001; Kalmijn, 2013; Stevenson & Wolfers, 2007). We examined the moderating effect of education and the results are shown in Table 13. We found that the coefficient on *Urban * Treated * Higher education* is significantly positive only in column (2) and insignificant in the other two columns, indicating that as housing prices rise, the more educated females are, the more likely they are to delay their age at first marriage. As mentioned earlier, in the Chinese marriage market, there is a gradient of mate selection, with women preferring to marry men with higher overall strengths than themselves, especially in economic strength (Gui, 2017; Wei & Zhang, 2011). For females with higher education levels, their choice of spouse is more demanding. As housing prices rise, there are fewer men to match with, thus further delaying their age of first marriage (Hu & Sun, 2021; Lichter et al., 1991).

Additionally, we examined the moderating effect of siblings. From the theoretical perspective of household economics, in a family, as the number of siblings increases, the economic resources that individuals can obtain from their parents will be diluted (Angrist et al., 2010; Becker & Lewis, 1973; Hanushek, 1992; Liu, 2014). Moreover, the influence of an individual's economic resources and status on marriage and age at first marriage has been widely recognized (Fisman et al., 2006; Hu & Wang, 2020; Wei & Zhang, 2011). For Chinese males in particular, they may not afford to spend on marriage (including buying a house) without the support of their parents (Hu et al., 2021; Liu et al., 2013). Thus, the effect of rising housing prices on marriage is likely to vary with the number of their siblings. In columns (4) to (6) of Table 13, we added the interaction term between siblings and the ULTP. As expected, we found a significant effect of this interaction on marriage delay for males, implying that as housing prices rise, males with more siblings have a further delay in the age at first marriage.

6.3. The effect of rising housing prices on fertility

In this subsection, we further explored two related issues regarding the long-term impacts of rising housing prices from the ULTP. The married sample we used facilitates our examination of the effect of rising housing prices on fertility. In columns (1) and (2) of Table 14, we found that rising housing prices have a significant negative impact on the number of children, which means that the rise in housing prices may limit the family's fertility decision. A vital reason for above results may be the marriage delay caused by rising housing prices. Delays in the first marriage will obviously cause delays in the birth of children. More specifically, the effect of ULTP-induced housing prices increases resulted in an average decrease in the number of children of 0.1857. For fertility intentions, as shown by the results in columns (3) and (4) of Table 14, we also found that rising housing prices have significantly reduced fertility intentions.

The results in Table 14 are similar to previous studies that emphasized the inhibitory effect of housing burden on fertility (Clark et al., 2020; Wrenn et al., 2019), confirming that rising housing prices led to lower fertility in China. Obviously, the long-term impacts of rising housing prices on fertility choices and intentions need to be taken seriously, as demographic structure and population growth are now the key concerns of the Chinese government. On the one hand, the rise in housing prices directly increases the expected cost of raising children, since the house is often considered as necessities for marriage (Cheung et al., 2020; Fang & Tian, 2018; Wrenn et al., 2019). As this expected cost rises, fertility intentions will also decline. On the other hand, through the effect of marriage delays, rising housing prices further postpone the childbirth and lower the fertility intentions. Therefore, controlling the excessive rise in housing prices to avoid excessive delays in first marriages may be an effective way to increase the birth rate in China.

7. Conclusions and Discussion

7.1. Conclusions

The impact of rising housing prices on Chinese families has been widely discussed (e.g., Clark et al., 2020; Hu et al., 2019; Li & Wu, 2014; Pan & Yang, 2022; Su et al., 2020; Wang & Hou, 2021), but little literature addressed the causal relationship between rising housing prices and marriage delays. Some previous studies showed that rising housing prices brought about a decline in the likelihood of young people getting married (Hu et al., 2021; Wrenn et al., 2019); yet as a country with universal marriage, marriage postponements and the age of first marriage are more relevant to the new population and fertility policies that deserve more attention from all walks of life. In 2002, the Chinese government improved the urban land policy, making the transaction process more market-oriented and open, but leading to a rapid increase in housing prices (Hu et al., 2021; Wang et al., 2018). By using this quasi-natural experiment of ULTP in China, we adopted the DID method to identify the causal relationship between rising housing prices and the age at first marriage among young people.

By using a nationally representative data from the CLDS, we found that rising housing prices from the ULTP significantly delayed the age of first marriage by 0.2280 years old, with males delaying by 0.4372 years old, indicating that for young people in urban China, the burden effect of rising housing prices is higher than the wealth effect. We divided the treated groups into two groups to reduce endogeneity according to whether respondents reached the legal age of marriage and obtained similar findings. In addition, our main results are supported by a series of robustness tests, such as placebo tests, excluding potential interference from the CEEP, one-child policy, and accession to the WTO in 2001, and removing special samples. Moreover, we found that the ULTP-induced rising housing prices led to higher spending of the first marriage; instead, young people who have houses are more likely to get

married and get married at a younger age. We further analyzed some moderating effects, finding that after the ULTP implementations, the higher the provincial housing prices, the more severe the delay in marriage among young people; rising housings prices have a greater negative impact on the age of first marriage for more educated females, males with more siblings, and individuals from large cities. Furthermore, our estimates suggest that the effect of delayed first marriage caused by rising housing prices further depresses fertility intentions and reduces the actual fertility rate.

7.2. Discussion

Our findings contribute to the study of the adverse effects of rising housing prices on Chinese families. Using the DID approach, we found that rising housing prices are responsible for the marriage delay and the increase in "leftovers" among young people, which supports the view that high housing prices lead to higher marriage costs and discourage marriage (e.g., Assaad et al., 2021; Hu et al., 2021; Hu & Wang, 2020; Wei & Zhang, 2011; Wrenn et al., 2019). In addition, in the debate on the wealth effect and burden effect of rising housing prices, our results favored the burden effect, emphasizing the negative impact of rising housing prices on the welfare of Chinese households. More importantly, given the strong association between marriage postponement and fertility (Cai, 2010), our study provided some evidence to further explain the low fertility phenomenon in China. A series of negative effects caused by the booming real estate market and soaring housing prices deserve further academic attention.

Currently, population problems have become quite serious in China. The large number of "leftover" females and males in the marriage market, the low fertility rate, and the aging population are all closely related to the sustainable socioeconomic development (Clark et al., 2020; Jiang & Liu, 2016; Wei et al., 2017; Zhan et al., 2021; Zheng, 2019). As our findings show rising housing prices, as a big mountain over the heads of Chinese young people, lead them to postpone marriage. Thus, it is necessary to develop appropriate and relevant public policies to regulate the real estate market, and further alleviate the problems of delayed marriage and low fertility.

Balancing the development of real estate markets with changes in demographic structure is quite important, which avoids low marriage rates and low fertility rates caused by rising housing prices in advance. More importantly, marriage delay may only be a symptom, and what needs more attention may be the excessive economic and social pressures on young people from high housing prices. Therefore, in a society with a high cost of living and low marriage rates, the government can try to strengthen economic and psychological support for young people through tax cuts or subsidies. In addition, these measures may also be of interest to other countries facing high housing prices and rapidly developing real estate markets. Taken together, this paper suggested that reducing housing prices and people's housing burden could be important and effective strategies to alleviate the problems of late marriage and low fertility.

Some limitations need to be addressed in subsequent studies. First, since city-level housing prices data around 2002 are not available and we employed a quasi-natural experiment from the ULTP, we are unable to precisely capture the net effect of ULTP on rising housing prices in urban China. Despite our efforts through descriptive analysis and moderating effects analysis, it is still difficult to obtain the magnitude of specific elasticity of the effect of rising housing prices on marriage delays among Chinese young people. Second, since data were not available when the ULTP was implemented in 2002, it may raise reverse causality issues, even though we were able to observe respondents' marital status and the age at first marriage. In our study, we tried to overcome this endogeneity problem as much as possible by adding the main FEs and core exogenous control variables and by grouping them according to the exogenous threshold of legal age at marriage. However, other exogenous characteristics that influence respondents' past marriage decisions may not be fully captured by the limitations of the CLDS questionnaire. Therefore, this issue is expected to be addressed in a follow-up study with a survey specifically targeting unmarried young adults. Third, we simply divided the sample into two categories: urban and rural observations. In fact, there are also areas such as suburbs and urban-rural junctions, where people living there may be influenced by ULTP as well as rural land rules (Hu et al., 2021). Subject to the limitations of the database, these observations are indistinguishable in this study. Last, although we tested the mechanism of marriage cost, i.e., we found that ULTP implementations led to an increase in the spending of first marriage, but some rest potential mechanisms need to be explored and examined in follow-up research, such as marriage intentions and the expected increase in the cost of living after marriage.

CRediT authorship contribution statement

Chunkai Zhao: Conceptualization, Formal analysis, Software, Methodology, Writing – original draft, Writing – review & editing, Funding acquisition. **Boou Chen:** Conceptualization, Methodology, Validation, Writing – review & editing, Funding acquisition. **Xing Li:** Data curation, Formal analysis, Software, Resources.

Conflict of interest

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Data availability

Data will be made available on request.

Acknowledgments

The authors acknowledge the support from the Fundamental Research Funds for the Central Universities (No. QCDC-2020-10; No. QCDC-2020-21), and the Tsinghua Rural Studies PhD Scholarship (202110).

References

Anderson, S. (2007). Why the marriage squeeze cannot cause dowry inflation. *Journal of Economic Theory, 137*(1), 140–152.
Angrist, J., Lavy, V., & Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics, 28*(4), 773–824.
Assaad, R., Krafft, C., & Rolando, D. J. (2021). Evaluating the impact of housing market liberalization on the timing of marriage: Evidence from Egypt. *Population Studies, 75*(3), 343–361.
Atalay, K., Li, A., & Whelan, S. (2021). Housing wealth, fertility intentions and fertility. *Journal of Housing Economics, 54*, Article 101787.
Baker, M., Halberstam, Y., Kroft, K., Mas, A., & Messacar, D. (2019). *Pay transparency and the gender gap*. National Bureau of Economic Research, No. w25834.
Bongaarts, J., & Greenhalgh, S. (1985). An alternative to the one-child policy in China. *Population and Development Review, 11*(4), 585–617.
Becker, G. S. (1974). A theory of marriage: Part II. *Journal of Political Economy, 82*(2), S11–S26.
Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy, 81*(2), S279–S288.
Becker, G. S., & Murphy, K. M. (1992). The division of labor, coordination costs, and knowledge. *Quarterly Journal of Economics, 107*(4), 1137–1160.
Brandt, L., Van Biesebroeck, J., Wang, L., & Zhang, Y. (2019). WTO accession and performance of chinese manufacturing firms: Corrigendum. *American Economic Review, 109*(4), 1616–1621.
Brines, J. (1994). Economic dependency, gender, and the division of labor at home. *American Journal of Sociology, 100*(3), 652–688.
Bulte, E., Heerink, N., & Zhang, X. (2011). China's one-child policy and "the mystery of missing women": Ethnic minorities and male-biased sex ratios. *Oxford Bulletin of Economics and Statistics, 73*(1), 21–39.
Cai, Y. (2010). China's below-replacement fertility: Government policy or socioeconomic development? *Population and Development Review, 36*(3), 419–440.
Cao, H., Fine, M., Fang, X., & Zhou, N. (2019). Chinese adult children's perceived parents' satisfaction with adult children's marriage, in-law relationship quality, and

67

adult children's marital satisfaction. *Journal of Social & Personal Relationships, 36*(4), 1098–1122.
Che, Y., & Zhang, L. (2018). Human capital, technology adoption and firm performance: Impacts of China's higher education expansion in the late 1990s. *Economic Journal, 128*(614), 2282–2320.
Chen, J., Hu, M., & Lin, Z. (2019). Does housing unaffordability crowd out elites in Chinese superstar cities? *Journal of Housing Economics, 45*, Article 101571.
Chen, J., Hu, M., & Lin, Z. (2021). China's housing reform and labor market participation. *Journal of Real Estate Finance and Economics*. https://doi.org/10.1007/s11146-021-09827-3
Chen, Y., Fan, Z., Gu, X., & Zhou, L. A. (2020). Arrival of young talent: The send-down movement and rural education in China. *American Economic Review, 110*(11), 3393–3430.
Chen, Y., & Zhou, L. A. (2007). The long-term health and economic consequences of the 1959–1961 famine in China. *Journal of Health Economics, 26*(4), 659–681.
Cheng, Z., Prakash, K., Smyth, R., & Wang, H. (2020). Housing wealth and happiness in urban China. *Cities, 96*, Article 102470.
Cheung, W. K. S., Chan, J. T. K., & Monkkonen, P. (2020). Marriage-induced homeownership as a driver of housing booms: Evidence from Hong Kong. *Housing Studies, 35*(4), 720–742.
Chu, C. C., Lin, J. C., & Tsay, W. J. (2020). Males' housing wealth and their marriage market advantage. *Journal of Population Economics, 33*(3), 1005–1023.
Clark, W. A., Yi, D., & Zhang, X. (2020). Do house prices affect fertility behavior in China? An empirical examination. *International Regional Science Review, 43*(5), 423–449.
Deng, F. F. (2005). Public land leasing and the changing roles of local government in urban China. *Annals of Regional Science, 39*(2), 353–373.
Deng, Y., Gyourko, J., & Wu, J. (2012). *Land and house price measurement in China. NBER working paper, No. 18403*.
Ding, C. (2007). Policy and praxis of land acquisition in China. *Land Use Policy, 24*(1), 1–13.
Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review, 91*(4), 795–813.
Dvornak, N., & Kohler, M. (2007). Housing wealth, stock market wealth and consumption: A panel analysis for Australia. *Economic Record, 83*(261), 117–130.
Engle, R. F., Lilien, D. M., & Watson, M. (1985). A dymimic model of housing price determination. *Journal of Econometrics, 28*(3), 307–326.
Ermisch, J. (1999). Prices, parents, and young people's household formation. *Journal of Urban Economics, 45*(1), 47–71.
Eriksen, M. D. (2010). Homeownership subsidies and the marriage decisions of low-income households. *Regional Science and Urban Economics, 40*(6), 490–497.
Fang, L., & Tian, C. (2018). Housing and marital matching: A signaling perspective. *China Economic Review, 47*, 27–46.
Feijten, P., & Mulder, C. H. (2002). The timing of household events and housing events in the Netherlands: A longitudinal perspective. *Housing Studies, 17*(5), 773–792.
Fisman, R., Iyengar, S. S., & Simonson, K. I. (2006). Gender differences in mate selection: Evidence from a speed dating experiment. *Quarterly Journal of Economics, 121*(2), 673–697.
Gao, H. (2019). Public land leasing, public productive spending and economic growth in Chinese cities. *Land Use Policy, 88*, Article 104076.
Gennaioli, N., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2013). Human capital and regional development. *Quarterly Journal of Economics, 128*(1), 105–164.
Goldstein, J. R., & Kenney, C. T. (2001). Marriage delayed or marriage forgone? New cohort forecasts of first marriage for US women. *American Sociological Review, 66*, 506–519.
Gui, T. (2017). "Devalued" daughters versus "appreciated" sons: Gender inequality in China's parent-organized matchmaking markets. *Journal of Family Issues, 38*(13), 1923–1948.
Guilmoto, C. Z. (2012). Skewed sex ratios at birth and future marriage squeeze in China and India, 2005–2100. *Demography, 49*(1), 77–100.
Gutiérrez-Domènech, M. (2008). The impact of the labour market on the timing of marriage and births in Spain. *Journal of Population Economics, 21*(1), 83–110.
Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy, 100*(1), 84–117.
He, C., Huang, Z., & Wang, R. (2014). Land use change and economic growth in urban China: A structural equation analysis. *Urban Studies, 51*(13), 2880–2898.
Horioka, C. Y., & Terada-Hagiwara, A. (2017). The impact of sex ratios before marriage on household saving in two asian countries: The competitive saving motive revisited. *Review of Economics of the Household, 15*(3), 739–757.
Hu, M., Su, Y., & Ye, W. (2019). Promoting or inhibiting: The role of housing price in entrepreneurship. *Technological Forecasting and Social Change, 148*, Article 119732.
Hu, M., & Sun, Z. (2021). Too educated to be married? An investigation into the relationship between education and marriage in China. *Journal of Family Studies*. https://doi.org/10.1080/13229400.2021.1927152
Hu, M., & Wang, X. (2020). Homeownership and household formation: No homeownership, no marriage? *Journal of Housing and the Built Environment, 35*(3), 763–781.
Hu, M., Wu, L., Xiang, G., & Zhong, S. (2021). Housing prices and the probability of marriage among the young: Evidence from land reform in China. *International Journal of Emerging Markets*. https://doi.org/10.1108/IJOEM-09-2020-1116
Huang, W., Lei, X., & Sun, A. (2020). Fertility restrictions and life-cycle outcomes: Evidence from the one-child policy in China. *Review of Economics and Statistics, 103*(4), 1–17.
Ji, Y. (2015). Between tradition and modernity: "Leftover" women in Shanghai. *Journal of Marriage and Family, 77*(5), 1057–1073.
Jiang, Q., & Liu, Y. (2016). Low fertility and concurrent birth control policy in China. *History of the Family, 21*(4), 551–577.
Kalmijn, M. (2013). The educational gradient in marriage: A comparison of 25 european countries. *Demography, 50*(4), 1499–1520.
Li, H., Ma, Y., Meng, L., Qiao, X., & Shi, X. (2017). Skill complementarities and returns to higher education: Evidence from college enrollment expansion in China. *China Economic Review, 46*, 10–26.
Li, J. (2014). Land sale venue and economic growth path: Evidence from China's urban land market. *Habitat International, 41*, 307–313.
Li, J., Lu, Y., Song, H., & Xie, H. (2019). Long-term impact of trade liberalization on human capital formation. *Journal of Comparative Economics, 47*(4), 946–961.
Li, L., & Wu, X. (2014). Housing price and entrepreneurship in China. *Journal of Comparative Economics, 42*(2), 436–449.
Liao, L., Zhao, C., Li, X., & Qin, J. (2021). Towards low carbon development: The role of forest city constructions in China. *Ecological Indicators, 131*, Article 108199.
Liu, H. (2014). The quality-quantity trade-off: Evidence from the relaxation of China's one-child policy. *Journal of Population Economics, 27*(2), 565–602.
Liu, H., Li, S., & Feldman, M. W. (2013). Gender in marriage and life satisfaction under gender imbalance in China: The role of intergenerational support and SES. *Social Indicators Research, 114*(3), 915–933.
Liu, J., & Liu, T. (2020). Two-child policy, gender income and fertility choice in China. *International Review of Economics & Finance, 69*, 1071–1081.
Liu, L., Jin, X., Brown, M. J., & Feldman, M. W. (2014). Male marriage squeeze and inter-provincial marriage in central China: Evidence from Anhui. *Journal of Contemporary China, 23*(86), 351–371.
Liu, S., & Zhang, Y. (2020). Cities without slums? China's land regime and dual-track urbanization. *Cities, 101*, Article 102652.
Liu, Y., Liu, Y., & Lin, Y. (2021). Upward or downward comparison? Migrants' socioeconomic status and subjective wellbeing in Chinese cities. *Urban Studies, 58*(12), 2490–2513.
Liu, Z. (2007). The external returns to education: Evidence from Chinese cities. *Journal of Urban Economics, 61*(3), 542–564.
Lichter, D. T., LeClere, F. B., & McLaughlin, D. K. (1991). Local marriage markets and the marital behavior of black and white women. *American Journal of Sociology, 96*(4), 843–867.
Lovenheim, M. F., & Mumford, K. J. (2013). Do family wealth shocks affect fertility choices? Evidence from the housing market. *Review of Economics and Statistics, 95*(2), 464–475.
Lu, Y., & Yu, L. (2015). Trade liberalization and markup dispersion: Evidence from China's WTO accession. *American Economic Journal: Applied Economics, 7*(4), 221–253.
Mo, J. (2018). Land financing and economic growth: Evidence from chinese counties. *China Economic Review, 50*, 218–239.
Moretti, E. (2004). Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *American Economic Review, 94*(3), 656–690.
Moser, P., & Voena, A. (2012). Compulsory licensing: Evidence from the trading with the enemy act. *American Economic Review, 102*(1), 396–427.
Mulder, C. H., & Wagner, M. (1998). First-time home-ownership in the family life course: A west german-dutch comparison. *Urban Studies, 35*(4), 687–713.
Nie, G. (2020). Marriage squeeze, marriage age and the household savings rate in China. *Journal of Development Economics, 147*, Article 102558.
Pan, Y., & Yang, H. (2022). Impacts of housing booms on fertility in China: A perspective from homeownership. *International Regional Science Review*. https://doi.org/10.1177/01600176211066472
Phang, S. Y. (2001). Housing policy, wealth formation and the Singapore economy. *Housing Studies, 16*(4), 443–459.
Raymo, J. M. (1998). Later marriages or fewer? Changes in the marital behavior of Japanese women. *Journal of Marriage and the Family, 60*(4), 1023–1034.
Stevenson, B., & Wolfers, J. (2007). Marriage and divorce: Changes and their driving forces. *Journal of Economic Perspectives, 21*(2), 27–52.
Su, C. W., Khan, K., Hao, L. N., Tao, R., & Peculea, A. D. (2020). Do house prices squeeze marriages in China? *Economic Research-Ekonomska Istraživanja, 33*(1), 1419–1440.
Sung, J., & Qiu, Q. (2020). The impact of housing prices on health in the United States before, during, and after the great recession. *Southern Economic Journal, 86*(3), 910–940.
Tian, L., & Ma, W. (2009). Government intervention in city development of China: A tool of land supply. *Land Use Policy, 26*(3), 599–609.
Tian, L., Yan, Y., Wu, Y., & Shao, L. (2020). Breaking the land monopoly: Can collective land reform alleviate the housing shortage in China's mega-cities? *Cities, 106*, Article 102878.
Xiao, Y., Li, L., & Zhao, L. (2017). Education on the cheap: The long-run effects of a free compulsory education reform in rural China. *Journal of Comparative Economics, 45*(3), 544–562.
Wang, F., & Zhang, C. (2021). Housing differentiation and subjective social status of Chinese urban homeowners: Evidence from CLDS. *Housing Studies, 36*(4), 567–591.
Wang, H., Smyth, R., & Cheng, Z. (2017). The economic returns to proficiency in english in China. *China Economic Review, 43*, 91–104.
Wang, R., & Hou, J. (2021). Land finance, land attracting investment and housing price fluctuations in China. *International Review of Economics & Finance, 72*, 690–699.
Wang, X., & Wen, Y. (2012). Housing prices and the high chinese saving rate puzzle. *China Economic Review, 23*(2), 265–283.
Wang, Y., Yu, Y., & Su, Y. (2018). Does the tender, auction and listing system in land promote higher housing prices in China? *Housing Studies, 33*(4), 613–634.
Wang, Y. P. (2000). Housing reform and its impacts on the urban poor in China. *Housing Studies, 15*(6), 845–864.

Waxman, A., Liang, Y., Li, S., Barwick, P. J., & Zhao, M. (2020). Tightening belts to buy a home: Consumption responses to rising housing prices in urban China. *Journal of Urban Economics, 115*, Article 103190.
Wee, S., Coffman, M., & La Croix, S. (2018). Do electric vehicle incentives matter? Evidence from the 50 US states. *Research Policy, 47*(9), 1601–1610.
Wei, S. J., Zhang, X., & Liu, Y. (2017). Home ownership as status competition: Some theory and evidence. *Journal of Development Economics, 127*, 169–186.
Wei, S., & Zhang, X. (2011). The competitive saving motive: Evidence from rising sex ratios and savings rates in China. *Journal of Political Economy, 119*(3), 511–564.
Wen, H., Xiao, Y., & Hui, E. (2019). Quantile effect of educational facilities on housing price: Do homebuyers of higher-priced housing pay more for educational resources? *Cities, 90*, 100–112.
Wrenn, D. H., Yi, J., & Zhang, B. (2019). House prices and marriage entry in China. *Regional Science and Urban Economics, 74*, 118–130.
Xu, X., & Whyte, M. K. (1990). Love matches and arranged marriages: A chinese replication. *Journal of Marriage and the Family, 52*(3), 709–722.
Yang, S., Guariglia, A., & Horsewood, N. (2021). To what extent is the chinese housing boom driven by competition in the marriage market? *Journal of Housing and the Built Environment, 36*(1), 47–67.
Ye, B., & Zhao, Y. (2018). Women hold up half the sky? Gender identity and the wife's labor market performance in China. *China Economic Review, 47*, 116–141.
Yu, J., & Xie, Y. (2015). Changes in the determinants of marriage entry in post-reform urban China. *Demography, 52*(6), 1869–1892.
Yuan, F., Wei, Y. D., & Wu, J. (2020). Amenity effects of urban facilities on housing prices in China: Accessibility, scarcity, and urban spaces. *Cities, 96*, Article 102433.
Zeng, Y., & Vaupel, J. W. (1989). The impact of urbanization and delayed childbearing on population growth and aging in China. *Population and Development Review, 15*(3), 425–445.
Zhan, P., Ma, X., & Li, S. (2021). Migration, population aging, and income inequality in China. *Journal of Asian Economics, 76*, Article 101351.
Zhang, J., Han, J., Liu, P., & Zhao, Y. (2008). Trends in the gender earnings differential in urban China, 1988–2004. *Industrial & Labor Relations Review, 61*(2), 224–243.
Zhang, C., & Zhang, F. (2019). Effects of housing wealth on subjective well-being in urban China. *Journal of Housing and the Built Environment, 34*(4), 965–985.
Zhao, C., & Guo, J. (2022). Are veterans happy? Long-term military service and the life satisfaction of elderly individuals in China. *Journal of Happiness Studies, 23*(2), 477–508.
Zhao, C., Wu, Y., & Guo, J. (2022). Mobile payment and Chinese rural household consumption. *China Economic Review, 71*, Article 101719.
Zheng, X., Yuan, Z. Q., & Zhang, X. (2020). Does happiness dwell in an owner-occupied house? Homeownership and subjective well-being in urban China. *Cities, 96*, Article 102404.
Zheng, Z. (2019). Reproductive behaviour and determinants in a low-fertility era in China. *Asian Population Studies, 15*(2), 127–130.
Zhong, H. (2014). The effect of sibling size on children's health: A regression discontinuity design approach based on China's one-child policy. *China Economic Review, 31*, 156–165.

Transport Policy 138 (2023) 129–143

Contents lists available at ScienceDirect

Transport Policy

journal homepage: www.elsevier.com/locate/tranpol

Is digital finance environmentally friendly in China? Evidence from shared-bike trips

Chunkai Zhao [a], Yuhang Wang [b], Zhenyu Ge [b,*]

[a] College of Economics and Management, South China Agricultural University, No.483, Wushan Street, Tianhe District, Guangzhou, Guangdong Province, 510642, China
[b] Business School, Xiangtan University, Yanggutang Street, Yuhu District, Xiangtan, Hunan Province, 411105, China

ARTICLE INFO

JEL classification:
G23
Q51
R41

Keywords:
Digital finance
Shared bike
Green travel behavior
Payment convenience
China

ABSTRACT

As a green travel method, shared bikes have penetrated all aspects of residents' daily lives; however, few studies have examined the role of digital finance in shared-bike trips, particularly in China, where digital finance tools are widely adopted. Using nationally representative data from the China Household Finance Survey (CHFS), we find that digital finance can significantly promote shared-bike trips of Chinese residents, indicating that digital finance is indeed environmentally friendly in terms of green travel behavior. In particular, the results suggest that a one standard deviation increase in digital finance increases the probability of riding shared bikes by 0.1468 standard deviations. The baseline results remain robust after using the instrumental variable approach to control potential endogeneity. Furthermore, we provide evidence that this positive effect of digital finance is likely to be explained by payment convenience, credit services, and social interactions. By stimulating shared-bike trips, digital finance further reduces the possibility of people using private cars. Moreover, the heterogeneous analysis shows inclusive attributes of digital finance. We find that digital finance has a significant positive effect on shared-bike trips for some socially disadvantaged groups, such as low-educated and low-income households and those living in cities with low levels of economic and financial development and digitalization. This study confirms the environment-friendly characteristics of digital finance from the particular perspective of shared bikes in China, supplementing the literature on influencing factors of shared bikes and the benefits of digital finance. Our evidence from the Chinese context provided insights for other countries to develop public policies and promote digital finance and shared bikes.

1. Introduction

Environmental issues have attracted widespread attention worldwide; excessive carbon emissions and air pollution, have become one of the main factors restricting socioeconomic development and threatening human health (Brunekreef and Holgate, 2002; Xu et al., 2013; Zhao et al., 2021). Previous studies have confirmed that transportation contributes about a quarter of global greenhouse gas and produces significant air pollutant emissions in cities (e.g., Creutzig et al., 2015; Tongwane et al., 2015). Private vehicles, in particular, consume substantial fossil energy and emit polluting gases; they are widely recognized as one of the causes of degrading urban air quality and impeding sustainable development (Cao and Shen, 2019; Chen et al., 2020; Zhang and Mi, 2018). From a combined environmental, social, and economic perspective, bicycles are considered the most sustainable mode of travel among all modes of transportation (Buekers et al., 2015; Lu et al., 2018).

In 1965, to alleviate traffic congestion and air pollution in Amsterdam, a Dutch environmental organization started the world's first public bicycle program called "White Bicycle"; however, this program went bankrupt shortly after its inception due to the misappropriation of bicycles. Subsequently, in 1991, Denmark developed a second generation of coin-operated or deposit-based public bicycles, but it was also not sustainable owing to bicycle loss and damage. In the late 1990s, the management capabilities of shared-bike programs were dramatically improved as technology developed. Then, high technologies, such as sensor technology, communication technology, data processing technology, network technology, automatic control technology, and information distribution technology, were gradually incorporated into public bicycle programs (Gao et al., 2019; Yang et al., 2021). France established a new generation of shared-bike systems relying on electronic information technology, which gradually spread to Spain, Belgium, Ireland, the United States, Japan, and other countries, which

* Corresponding author.
E-mail addresses: ckz_zhao@163.com (C. Zhao), wangyuhang325@163.com (Y. Wang), 775195543@xtu.edu.cn (Z. Ge).

https://doi.org/10.1016/j.tranpol.2023.05.007
Received 15 May 2022; Received in revised form 14 April 2023; Accepted 15 May 2023
Available online 16 May 2023
0967-070X/© 2023 Elsevier Ltd. All rights reserved.

successfully achieved commercial transformation.

Given the environmental benefits, the factors influencing people's adoption of shared bikes have attracted widespread academic interest (e.g., Campbell et al., 2016; Jie et al., 2021; Li C et al., 2021; Teixeira et al., 2022); however, scholars overlooked the potential impact of digital finance on shared-bike trips, particularly in China, where digital finance and bike sharing are very popular. Digital finance refers to a new generation of Internet-based financial services and information technology combined with traditional financial services format; in China, these mainly comprise mobile payments, Internet loans, online investment, wealth management, and Internet crowd funding (Li et al., 2020; Zhao et al., 2022). With the payment convenience and credit services provided by digital finance, the sharedbike has quickly exploded in China (Li et al., 2019; Yao et al., 2019). To date, China has the largest number of shared bicycles globally. Shared bikes provide residents with the "first mile" and "last mile" transportation convenience and offer a greener and low-carbon way to travel.

Using the nationally representative sample from the China Household Finance Survey (CHFS) in 2017, we adopt econometric methods to explore the environmentally friendly features of digital finance from the perspective of stimulating shared bike-trips. With the quasi-natural experiment with shared bikes entering different Chinese cities for the first time (Huang et al., 2022), we firstly use difference-in-differences (DID) methods to show that shared bicycles help reduce air pollution and confirm the environment-friendly characteristics of shared bikes in China. Then, we further examine the impact of digital finance on shared-bike trips. Considering the potential endogeneity issue, we strictly control for the district/county fixed effect and many city characteristic variables, with correction with city-level cluster standard errors. More importantly, instrumental variable (IV) estimates also imply the robustness of our results. This research has important practical implications as the inclusiveness of digital finance benefits more socially vulnerable groups and makes up for their short-distance trip options with cheaper and more convenient shared bikes. Supported by digital finance, shared bikes are expected to alleviate air pollution and reduce carbon emissions in China. The results of this study offer some insights for policymakers to promote digital finance further, advocate shared-bike trips, and develop green transportation, which can improve environmental quality and accelerate sustainable urban development.

This study contributes to the existing literature in three aspects. First, it is the first to employ nationally representative sampling survey data to evaluate the impact of digital finance on shared-bike trips. Our study confirms that digital finance development in China is environmentally friendly from the particular perspective of the shared bike as green travel, complementing the relevant literature on the benefits of digital finance (e.g., Chen and Zhao, 2021; He and Li, 2020; Ozili, 2018). Second, we analyze and verify in detail how digital finance promotes people to use shared bikes, examining the mechanisms behind payment convenience, credit services, and social interactions. Moreover, we provide evidence that this positive effect reduces people to adopting polluting modes of travel, which complements the research on the environmental benefits of shared bikes (e.g., Cao and Shen, 2019; Ma et al., 2018; Teixeira et al., 2021). Finally, our heterogeneity focuses on the inclusive financial attributes of digital finance from macro and micro aspects. We find that digital finance is more helpful for some socially disadvantaged groups, confirming the positive effect of inclusive finance on residents' welfare from the perspective of shared bikes (Lai et al., 2020; Zhao et al., 2022).

The remainder of this study is organized as follows. Section 2 summarizes the relevant literature and briefly describes the institutional background. Section 3 outlines the mechanisms of digital finance on shared-bike trips, and Section 4 introduces the data, variables, and methodology. Section 5 presents and explains the empirical results, and Section 6 presents the conclusions and discussion.

2. Literature review and institutional background

2.1. Literature review

Three strands of literature are related to our study. First, we focus on the environmental impacts of shared bikes in prior research. Second, a class of studies on the factors influencing shared-bike trips is included in this subsection as we explore the effect of digital finance on shared-bike trips. Finally, we briefly investigate the literature related to the benefits of digital finance.

2.1.1. Environmental impact of shared bikes

Given the critical role of bicycle travel in reducing traffic congestion and improving the environment, many scholars have explored the environmental benefits of bicycle sharing. Studies in China and other countries have found that the development of bike sharing is associated with reductions in greenhouse gases and air pollutant emissions (e.g., Cao and Shen, 2019; Chen and Lu, 2016; Chen et al., 2020; Lu et al., 2018; Teixeira et al., 2021; Zhang and Mi, 2018). For instance, the Bixi bike-sharing program in Montreal reported that in less than 1 year, the implementation of the program reduced greenhouse gases by 3 million pounds; not coincidentally, the bike-share program in Lyon, France, reduced CO_2 by more than 18 million pounds (Lu et al., 2018; Yang et al., 2021). In Chinese cases, Zhang and Mi (2018) found that shared bikes in Shanghai decreased CO_2 emissions by 25,240 tons in 2016; the evidence from Cao and Shen (2019) suggested that bike-sharing registration, distance traveled, and usage rates significantly reduced carbon emissions in Beijing.

Furthermore, several studies explored whether bicycle sharing could replace people's use of some polluting modes of transportation. Fuller et al. (2013) and Lu et al. (2018) found that few residents were willing to switch from private cars to bicycle sharing, whereas more owned private bicycles or frequently walked or took the bus. Nonetheless, some scholars also pointed out that shared bicycles have improved people's overall cognition of cycling (Goodman and Cheshire, 2014) and the travel behavior of low-income individuals is more likely to be influenced by bike-sharing programs towing to the universality and accessibility of shared bikes (Murphy and Usher, 2015).

Conversely, a category of literature pointed out the potential negative environmental impact of shared bikes (Bonilla-Alicea et al., 2020; Luo et al., 2019). From a life-cycle perspective, environmental burdens may be caused during the production, operation, management, disposal, and recycling stages of shared bicycles (Cheng et al., 2022; Luo et al., 2019, 2020; Mao et al., 2021). In particular, rubber and aluminum products in shared bikes may produce greenhouse gases and pollutants during production and recycling (Mao et al., 2021). Moreover, China's bike-sharing market has an overcapacity problem in the competition of capital chasing, resulting in a supply far outweighing the demand (Han, 2020; Luo et al., 2020); thus, some abandoned shared bikes have become urban solid waste (Mao et al., 2021). Nevertheless, after carefully considering the negative environmental impacts of bike sharing, some studies also confirmed the positive role of bike sharing in improving the environment and promoting sustainable development in China (Cheng et al., 2022; Luo et al., 2020). Sun and Ertz (2021) found that shared bikes led to lower consumption of aluminum, steel, plastic, and rubber in Beijing, highlighting the positive effect of shared bicycles on resource conservation.

2.1.2. Factors influencing of shared-bike trips

In addition to exploring the environmental impacts of shared bikes, previous literature focused on the driving factors for using shared bicycles. First, among the macro factors, scholars found that economic development, higher population density and employment, good transportation infrastructure, and flat topography are positively associated with the frequency of riding shared bikes (e.g., Ebrahimi et al., 2022; Li, C. et al., 2021; Li et al., 2019; Ricci, 2015; Winters et al., 2016; Yao et al.,

2019; Zhao et al., 2014). In particular, several studies highlighted the positive effect of information technology on promoting bike sharing (e.g., Gao et al., 2019; Li C et al., 2021; Ma et al., 2018; Yang et al., 2020). Furthermore, other research was interested in the impact of weather on shared-bike use, finding that severe weather conditions, such as cold temperatures, heavy rain, high humidity, and stormy weather reduced bike-sharing activity, whereas comfortable temperatures (around 90 °F) increased shared bicycle trips (Campbell et al., 2016; Godavarthy and Taleqani, 2017; Zhang et al., 2017). Moreover, some studies emphasized the negative impact of air pollution on shared-bike trips (Apparicio et al., 2016; Hong et al., 2022; Zhao P et al., 2018).

Second, in terms of the bike-sharing system, other studies focused on pricing mechanisms (Papanikolaou, 2021; Waserhole and Jost, 2016), space layout and parking convenience of bikes (Campbell and Brakewood, 2017; Faghih-Imani et al., 2014), incentive mechanisms (Fricker and Gast, 2016), and pickup and delivery options (Hernández-Pérez and Salazar-González, 2004). Since bike sharing is free or relatively cheap in some cities, pricing mechanisms are not a critical factor influencing bike-sharing trips (Chu et al., 2021). Instead, residents are more concerned about the convenience of shared bikes, such as the distribution of bike sharing, ease of docking, and interchange with other modes of transportation (Campbell and Brakewood, 2017; Faghih-Imani et al., 2014; Teixeira et al., 2022).

Finally, from the perspective of individual characteristics, scholars concluded that age, gender, education, and income are important factors in using shared bikes (Buck et al., 2013; Efthymiou et al., 2013; Fishman et al., 2014; Li C et al., 2021; Jie et al., 2021; Orvin and Fatmi, 2021; Picasso et al., 2020). Some studies have found that young and middle-aged residents were more active in bike-sharing trips (Buck et al., 2013; Orvin and Fatmi, 2021) towing to their dependence on bike sharing to get to work or school (Barbour et al., 2019; Vogel et al., 2014). Furthermore, regarding gender, prior research confirmed that men were more willing to use bike sharing than women (Barbour et al., 2019; Jie et al., 2021; Vogel et al., 2014). Additionally, several studies provided evidence that bike sharing was closely related to higher income and education of users (Fishman et al., 2014; Rixey, 2013).

2.1.3. *Benefits of digital finance*

As a representative of inclusive finance, the benefits of digital finance have recently received close attention from scholars, particularly the positive impact on micro-welfare. First, with the Internet and digital technology, digital finance can make up for the shortcomings of traditional finance and reduce the high transaction costs caused by the backwardness of financial infrastructure (Shen et al., 2021; Zhao et al., 2022), enabling more people to access financial services, particularly socially vulnerable groups (Chen and Zhao, 2021; He and Li, 2020; Lai et al., 2020). Second, digital finance has a two-way function for information transfer and financial services (Li et al., 2020; Ozili, 2018), which helps increase household income and alleviate poverty by reducing financial exclusion, alleviating information asymmetry, and easing credit constraints (Chen and Zhao, 2021; Liu et al., 2021; Ozili, 2018; Zhao et al., 2022). Third, the role of digital finance in providing microfinance and widening social networks can effectively stimulate entrepreneurship (Luo and Zeng, 2020; Yin et al., 2019). Moreover, digital finance may smooth consumption and reduce household risks, further improving household welfare (Lai et al., 2020; Li et al., 2020). Finally, some scholars have also referred to the positive effects of digital finance on business management and operations (Chen and Zhang, 2021; Wang et al., 2020).

2.2. *Shared bikes in China*

In anticipation of the Beijing Olympics and in response to severe traffic congestion, the Beijing government established one of the first public bicycle systems in China in 2007. Soon, government-led docked public bicycle systems were piloted in some major Chinese cities such as Shanghai, Guangzhou, Hangzhou, and Wuhan. As a part of public transportation, these docked public bicycles positively relieve congestion and facilitate short-distance trips for residents.

In 2015, dockless public bicycles built on Internet of Things (IoT) and Internet technologies first appeared at Peking University. With GPS positioning, smartphones, and QR code payment, the new generation of public bicycles, both in terms of use and management, has become increasingly convenient (Zhao N et al., 2018). In 2016, commercial shared bikes, represented by Ofo and Mobike, officially entered some of China's major cities and became the most dominant bike-sharing brands. According to the China Shared Bicycle Market Research Report in 2017, these two companies occupied more than 96% of the market share. Subsequently, driven by venture capital, commercial shared-bike market in China expanded rapidly (Chen et al., 2020). Fig. 1 shows that by 2020, there were 434 shared-bike companies in China, with an average annual growth rate of 47.57% from 2016 to 2020. The annual number of newly registered shared-bike companies from 2017 to 2020 reached 80, 50, 77, and 164, respectively. Currently, the leading bike-sharing companies include Meituan (formerly Mobike), Hellobike, and Didi (Qingju Bike).

Similarly, the number of bike-sharing users and market placement has also proliferated. Fig. 2 shows that from 2016 to 2017, the number of shared-bike users in China grew from 28 million to 205 million. In 2019, China's shared-bike users reached 260 million, equivalent to 18.44% of the total population and 30.64% of the urban population. Furthermore, bike-sharing launches grew rapidly from 3.5 million in 2016 to 16.1 million in 2017. The number of shared bicycles in just four first-tier cities (Beijing, Shanghai, Guangzhou, and Shenzhen) in China reached 5 million in 2017, and shared bikes are actively pushing into second- and third-tier cities. For example, Hellobike has been present in more than 100 cities.[1] The 3rd Smart Transportation Summit in 2019 noted that the average daily use of shared bikes in China exceeded 40 million,[2] indicating that shared bikes have become one essential means of transportation for Chinese residents in their daily travels.

3. Theoretical analysis

In this section, we discuss the potential mechanisms by which digital finance promotes residents to use shared bikes as a greener travel mode. We classify possible mechanisms into the following three categories: payment convenience, credit services, and social interactions.

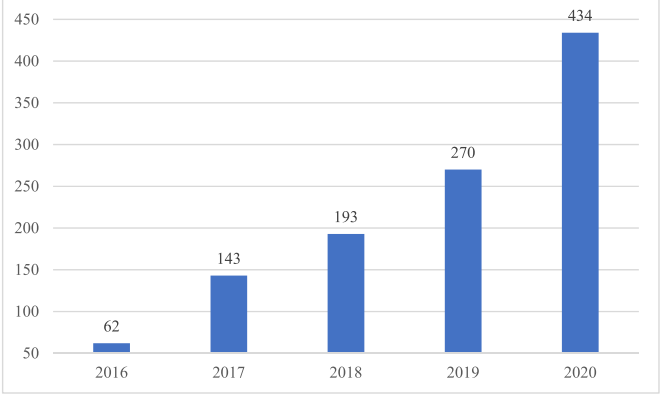


Fig. 1. Number of bike-sharing companies in China (2015–2020).
Source: Qichacha Tec Co., Ltd. (www.qcc.com)

[1] Data source: https://www.thepaper.cn/newsDetail_forward_1789848.
[2] Data source: https://ishare.ifeng.com/c/s/7qJ8Gz9HYEB.

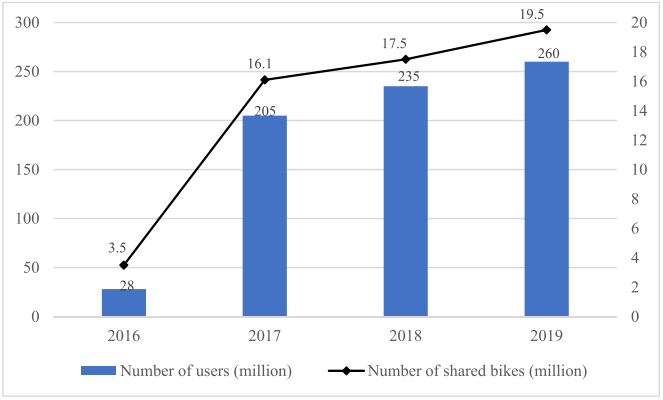


Fig. 2. The scale of shared bikes in China (2015–2019).
Source: Ministry of Transport of the People's Republic of China and Foresight Industrial Research Institute

3.1. Payment convenience

As mentioned earlier, the shared-bike platform in China has evolved from the traditional government-led docked public bicycle; however, as a supplement to urban public transportation (Yang et al., 2021), most docked bikes rely on the traditional payment system, such as cash, coins, or public transportation cards (Zhao N et al., 2018). For example, the Beijing Public Bike requires a designated location to apply for a bus card, open the public bike function, and pay a deposit of 200 Chinese Yuan (CNY). Moreover, first-time users of public bikes must register on-site with credentials.[3] In subsequent uses, if the balance on the bus card is insufficient, users must add funds to ride docked public bikes (Shaheen et al., 2011). Therefore, government-led public bikes rely on traditional payment methods; this inconvenience in payments leads to increased transaction costs, reduced user experience, and significant customer loss.

The commercial shared bike is becoming increasingly popular in China owing to the payment convenience offered by digital financial technology. As an essential part of digital finance, mobile payments make electronic money transactions through electronic products, such as smartphones (Huang et al., 2020; Yin et al., 2019). Through scanning the QR code, mobile payments avoid these traditional payment methods, which significantly save transaction costs in the payment process (Zhao et al., 2022). Additionally, for bike-sharing deposits, users can pay through third-party digital financial platforms, such as mobile payments, eliminating the time costs from on-site registration and handling shared-bike functions at designated places. Digital payment methods are now the preferred way to pay for bike sharing in China. Furthermore, China is one of the most developed countries for mobile payments in the world, and the user base of mobile payments is extensive (Yin et al., 2019). With the help of digital finance and mobile payment platforms, some bike-sharing companies have launched a series of promotional measures to attract more customers (Zhao N et al., 2018), such as one-way coupons, monthly rental discounts, and deposit-free rides. In summation, the payment convenience provided by digital finance creates conditions for developing shared bikes in China, potentially increasing the probability of residents using shared bikes as a green travel behavior.

3.2. Credit services

Essentially, a shared bike is a credit rental service that requires support from a sound social credit system (Li C et al., 2021). In 2017, the National Development and Reform Commission signed credit information-sharing agreements with 10 major Chinese shared bicycle enterprises, including Mobike and Ofo.[4] By strengthening the sharing of credit information and establishing a mechanism of joint incentives for trustworthiness and punishment for dishonesty, users who have violated the law and shown dishonesty will be restricted from using the service. In contrast, people with good credit can access priority use or convenience services, such as deposit-free, to build a standard development system for shared bicycles with credit as the core (Chen and Lu, 2016; Eren and Uz, 2020; Yao et al., 2019). Additionally, to reduce the potential financial risks associated with deposit pools and improve the convenience of bike-sharing trips, many Chinese shared-bike companies have begun waiving deposits, increasing the demand for credit services (Fishman et al., 2013).

Nonetheless, China's credit system is not perfect. In addition to credit information from government departments, credit services provided by digital finance are also crucial (Li et al., 2020; Yin et al., 2019). Digital financial tools, such as mobile payment, Internet lending, and online investment, are being adopted more frequently by Chinese residents; thus, digital credit systems are gradually being built (Li J et al., 2021). Many large digital financial platforms in China have digital credit services, such as Sesame Credit in Alipay and JD Credit. Additionally, small and microloans provided by WeChat Pay and Baidu Finance also rely on a personal digital credit system. In terms of shared bikes, these credit services from digital finance can help further reduce operating costs and attract newer bike-sharing users (Zhao N et al., 2018), increasing the likelihood of people using this green travel method. Several shared-bike companies have already been actively connected to digital credit in China. For example, since 2017, Hellobike and Ofo users can ride without a deposit if their Sesame Credit Score is over 650 and Funbike is free of deposit for users in Shenzhen with a Sesame Credit Score of less than 700. Therefore, with the development of digital finance and the improvement of the digital credit system, residents can enjoy more and more credit services while using shared bicycles.

3.3. Social interactions

Compared with traditional finance, digital finance may enhance the level of social interaction of households. Digital finance promotes using various software for digital payments, Internet lending, and online investment. Based on smartphones, tablets, and personal computers, the software performs its financial functions and frequent online information interactions (He and Li, 2020). For example, combining the custom of WeChat and Alipay red packets with traditional Chinese characteristics enhances residents' online social interaction experience (Zhao et al., 2022). Furthermore, digital finance has made it easier for people to participate in online entertainment. Some entertainment requires payment or membership for improved services, especially watching movies and TV series, listening to music, playing online games, etc. Undoubtedly, the convenience of digital finance gives residents more options to enjoy online fun easily.

In addition to online activities, digital finance makes it easier for offline social interactions (Chen and Zhao, 2021). Adopting digital financial tools is widespread in these offline scenarios in China. Regarding transportation, people can use digital finance tools to take buses, subways, and taxis, and there are discounts on highway payments. For offline shopping, residents can make purchases through mobile payments. Regarding dining out and hotel accommodations, reservations can be made through several digital financial platforms. In communication, digital finance tools can improve the service of quick inquiry and paying phone and Internet bills. Additionally, previous studies addressed the positive impact of social interactions on people's

[3] Source: https://xw.qq.com/cmsid/20210528A0CH6P00.

[4] Source: http://m.ce.cn/cj/gd/201704/27/t20170427224150 95.shtml?from=toutiao.

use of shared bikes (Jie et al., 2021; Li C et al., 2021; Shaheen et al., 2011); therefore, we consider that digital finance may enhance the likelihood of riding shared bikes by facilitating social interactions.

Furthermore, since digital finance and bike sharing are new technologies that have emerged in recent years, promoting them leads to some users testing them first. Under the influence of the peer effect, residents' acceptance and usage of digital finance and shared bikes may gradually increase. Given the close relationship between social interactions and peer effects (Bramoullé et al., 2009; Calabrese et al., 2006), we believe that the peer effect may also be a key channel in the social interaction effects from digital finance to promote shared-bike use.

4. Data and methodology

4.1. Data source

This study uses two primary data sources. First, the microdata are from the 2017 wave of the CHFS, which was released by Southwest University of Finance and Economics.[5] As a nationally representative household database, this round of the CHFS covers 27,279 urban householders from 29 provincial-level administrative regions in mainland China. The CHFS focuses on households' economic and financial status, financial tool usage, demographic characteristics, and employment. More importantly, the CHFS provides some information about shared-bike trips for residents, including whether they used shared bikes in the past week, the frequency of use and usual ride duration, the brand of shared bikes, and the factors considered while riding.[6]

We remove individuals living in rural areas, as China's bicycle sharing is mainly in urban areas. Furthermore, we deleted minor householders under the age of 18. Previous studies on adopting shared bikes mostly used small-scope and small-scale surveys (e.g., Chen and Huang, 2021; Guo et al., 2017; Ma et al., 2018), creating challenges in ensuring the randomness of the sample and the authenticity of the data. We employ a micro-database based on different cities, which can fully ensure the heterogeneity of the sample in terms of geography, income, educational level, and family background and enhance the reliability of empirical results and the generalizability of the conclusions.

Second, macro-data on digital finance at the prefecture level are from the Digital Finance Inclusion Index (DFII) compiled by the Institute of Digital Finance of Peking University in collaboration with Ali Finance. This index shows the development dynamics of digital finance at the provincial, city, and county levels in China from 2011 to 2018, which was widely used in previous studies (e.g., Chen and Zhao, 2021; Lai et al., 2020; Li et al., 2020; Liu et al., 2021). Unfortunately, the CHFS does not publish district/county information for respondents; therefore, we employ city-level digital finance indicators to match the microdata. Additionally, in the 2017 wave of the CHFS, respondents were mainly asked about the previous year (e.g., household income); therefore, the 2016 digital finance index was used in this study (Chen and Zhao, 2021). In sum, the CHFS and DFII provide accessible and reliable data sources for us to explore the effects of digital finance on residents' green travel behaviors from the perspective of bike-sharing trips.

4.2. Empirical models and variables

This section proceeds in two parts. Our first aim is to examine the impact of shared-bike entry on air pollution reduction to clarify the environmentally friendly characteristics of shared bikes. Subsequently, we further explore how digital finance affects residents' bike-sharing trips in China.

4.2.1. Shared bike entry and air pollution

As mentioned earlier, based on the quasi-natural experiment with shared bikes entering different Chinese cities for the first time, we first examine the effect of shared bicycles on air pollution to provide a priori evidence. Thus, the DID model is set as follows:

$$AP_{ct} = \alpha_0 + \alpha_1 SBE_{ct} + \alpha_2 W_{ct} + \lambda_c + \delta_t + \mu_{ct} \quad (1)$$

where AP_{ct} represents the air pollution variable set, including AQI, $PM_{2.5}$, PM_{10}, SO_2, CO, NO_2, and O_3. SBE_{ct} is a binary variable indicating whether shared bikes have entered the city c in year t. As per previous studies (Acemoglu et al., 2016; Lipp, 2007), considering the time lag effects, the entry of shared bikes into a city in the first half of the year is considered effective in the current year; the entry in the second half of the year is considered to be effective in the following year. Additionally, the data range here is 2013–2019 because, starting from 2016, shared bikes started to enter various Chinese cities (concentrated in 2017), and this time range can fully observe the changes in air pollution before and after the shared bike entry. Furthermore, after 2013, the Ministry of Ecology and Environment of China modified the statistical caliber of air pollutants, focusing on pollutants, such as $PM_{2.5}$ (Xiao et al., 2020).

W_{ct} in Equation (1) represents a set of time-varying covariates (e.g., Huang et al., 2022; Li and Zhao, 2022; Xiao et al., 2020), including *GDP per capita*, *Fiscal expenditure*, *Fiscal revenue*, *Road infrastructure*, *FDI*, *Enterprises*, *Financial development*, *Average temperature*, *Relative humidity*, and *Average rainfall*. λ_c and δ_t are the vectors of city and year dummy variables that aim to control the city fixed effect and year fixed effect. μ_{ct} indicates the standard error clustered at the city level. The coefficient α_1 is the DID estimate representing the impact of shared-bike entry on air pollution reduction in China. The details of the above variables are represented in Table A1 in the Appendix.

4.2.2. Digital finance and shared-bike trips

To investigate the effect of digital finance on shared-bike trips in China, we estimate the following Probit model:

$$SB^*_{ic} = \beta_0 + \beta_1 DF_c + \beta_2 X_{ic} + \theta_d + \varepsilon_{ic} \quad (2)$$

$$SB_{ic} = \begin{cases} 1 \; if \; SB^*_{ic} > 0; \\ 0 \; otherwise \end{cases} \quad (3)$$

In Equations (2) and (3), SB_{ic} is a dummy variable indicating whether individual i from city c uses shared bikes, and SB^*_{ic} is a latent variable for the shared bike. We set the explained variable according to the answers to the question, "During the past week, have you ever used the shared bike?" The question has three options: "Yes," "No," and "Have not heard of shared bikes." We construct a binary variable to measure whether residents used shared bikes, with a value of 1 if the answer is "yes" and 0 otherwise. As mentioned above, the CHFS also provides other information about residents riding shared bicycles; therefore, in the robustness tests, we adopt the usual cycling time in shared bikes as a proxy.[7]

DF_c denotes the digital finance index of city c. It is an aggregated index comprising indicators at three levels. The variance coefficient assignment method based on hierarchical analysis is used to determine the index weight, and then it is aggregated by a bottom-up weighted average through a dimensionless method. The secondary indicators include the digitization level, coverage breadth, and usage depth. Additionally, the DFII discloses only some of the tertiary indicators (e.g., Payment), whereas data for tertiary indicators of digitalization level (e.g., Mobility) are not available. The measurement criteria for digital

[5] The CHFS has released five waves of data since 2011, but surveys on using shared bikes are only available in the 2017 wave.

[6] To our knowledge, the CHFS is the first large micro-database to investigate the use of shared bikes by Chinese residents.

[7] Other variables related to shared bike use, such as cycling frequency, were not adopted due to too many missing values.

finance are shown in Fig. 3. We are interested in the coefficient β_1, and if it is significantly positive, it indicates that digital finance has a positive effect on shared-bike trips.

X_{ic} is a set of control variables, including the individual householder, household, and city levels. In householder characteristics, we choose some demographic variables, such as *Age*, *Age-squared, Hukou*, *Gender*, *Education*, *Marriage*, *Self-rated health* and *Employment*, as younger or older, more educated, and healthier people have a higher propensity to use shared bikes (Campbell et al., 2016; Efthymiou et al., 2013; Jie et al., 2021; Li, C. et al., 2021; Vogel et al., 2014). Household economic characteristics, such as *Household income*, *Household consumption, Transportation expenses*, *Household savings*, and *House ownership*, are considered (Campbell et al., 2016; Eren and Uz, 2020; Jie et al., 2021; Rixey, 2013). We also include *Car ownership* for alternative modes of travel (Fishman et al., 2014; Lu et al., 2018; Molina-García et al., 2015). At the city level, economic variables include *GDP per capita*, *GDP growth rate*, *Salary per capita*, *Loans*, *R&D expenditure*, and *Digitalization*. These variables, especially those measuring the levels of financial development and informatization of cities, are closely related to digital finance and potentially directly impact people's use of shared bikes (Chen and Zhao, 2021); thus, controlling for these variables significantly reduces the challenge from omitted variable bias in our estimates (Li C et al., 2021; Yao et al., 2019). Moreover, we further control for urban transportation infrastructure towing to its significant influence on shared-bike trips (Ebrahimi et al., 2022; Godavarthy and Taleqani, 2017; Zhang et al., 2017). These variables are the *Buses per capita*, *Taxis per capita*, and *Scope* (Giallouros et al., 2020). Consistent with digital finance, we adopt one-period lagged control variables[8] to mitigate potential endogeneity. The definitions of these variables are shown in Appendix Table A2.

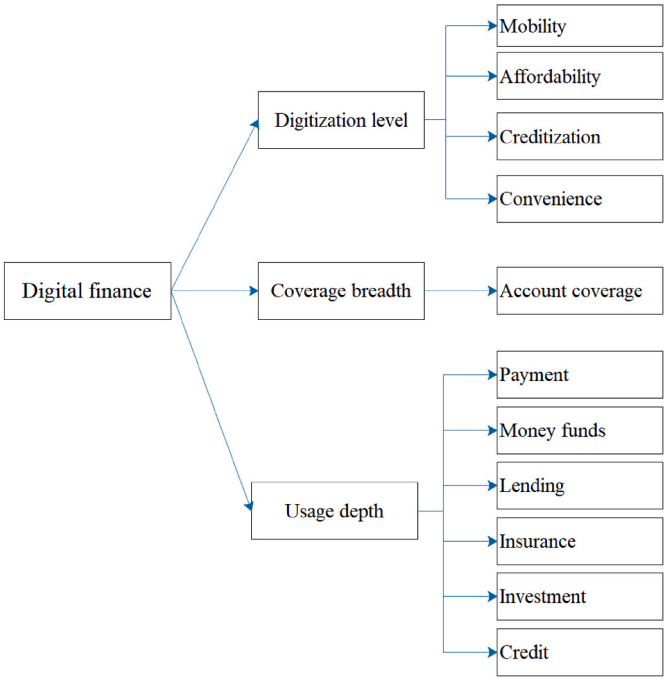


Fig. 3. The measurements of digital finance in China.
Source: The DFII from the Institute of Digital Finance of Peking University (https://idf.pku.edu.cn/yjcg/zsbg/index.htm)

θ_d is a vector of district/county-level dummies, capturing the district/county fixed effect.[9] ε_{ic} refers to the random disturbance term. Furthermore, although we include a considerable number of control variables in the estimated model, further consideration of potential endogeneity issues is theoretically required to accurately identify the causal relationship between digital finance and bike-sharing trips. On the one hand, since the behavior of micro-level individuals riding shared bikes has little effect on the macro-level city digital finance index, the two-way causality problem does not need to be explicitly considered; however, the digital financial development in a city is influenced by some confounders. For example, cities with higher financial foundations are more advantageous in developing digital finance (Lai et al., 2020; Li et al., 2020), and traditional finance may also affect residents' behavior in using bike sharing (Li C et al., 2021). However, some financial development information is not captured in our control variable set. Therefore, potential omitted variable bias may cause our estimates to reflect the impacts of digital finance on bike-sharing use ineffectively.

We attempt to mitigate the potential endogeneity problems from omitted variable bias in the following ways. First, as emphasized earlier, we control for a series of city-level variables. Second, we more tightly control for the district/county fixed effect, capturing district/county-level characteristics that do not change over time. Third, to control for serial correlation and heterogeneity of variables, we cluster the standard errors to the city level, which also helps to mitigate the endogeneity problem (Chen and Zhao, 2021). Finally, we employ the IV approach to identify the effect of digital finance on bike-sharing trips.

4.3. Descriptive statistics

The descriptive statistics for the above variables are presented in Tables A1 and A2—Table A1 presents 1599 observations in the city-year panel data. We can see that air pollution varies considerably from city to city or year in China. For example, the mean of annual SO_2 is 26.7543, with a standard deviation of about 23.3487. Moreover, the percentage of cities with bike-sharing entries from 2013 to 2019 was 10.63%. In Table A2, our final total sample included 9434 residents from 295 Chinese cities. We find that 1397 householders are shared-bike users in 2017, accounting for 14.81%. Furthermore, we find that the mean of the digital finance index is 240.85 in these Chinese cities.

5. Results

5.1. A priori evidence: shared bikes and air pollution reduction

As a green mode of travel, the environmental benefits of shared-bike trips have received confirmation from the previous literature (e.g., Huang et al., 2022; Lu et al., 2018; Monazzam et al., 2021; Zhang and Li, 2023); however, this study analyzes it from the unique perspective of residents' bike-sharing use behavior, which seems to be a bit distant from the environment. Therefore, we provide a priori evidence of the relationship between shared-bike trips and air pollution reduction.

The difficulty in identifying the effects of shared bikes on air pollution lies in the two-way causality problem; air pollution also affects the share bicycle use (Apparicio et al., 2016; Zhao P et al., 2018). Therefore, we resort to the approach of Huang et al. (2022) and adopt shared bikes entering different Chinese cities for the first time as a quasi-natural experiment. With the DID method, we can effectively identify this effect. The twofold differences come from the time when Ofo and Mobike[10] entered different cities. By matching air pollution data and city characteristics data, we can examine the impact of bike-sharing

[8] The data on the average slope in cities are from the ASTER GDEM. Other macrovariables were obtained from the China City Statistical Yearbook 2017.

[9] The CHFS has published district/county dummy codes. Although we cannot know the exact name, we can control the district/county fixed effect.

[10] As mentioned in Section 2, they are the first and main bike-sharing brands to enter the market in China.

entry on air pollution.[11]

The results using the DID methods are represented in Table 1, where we find that the estimated coefficients of *Shared-bike entry* are significantly negative in columns (2)–(4), indicating that shared bicycle entry significantly reduces the concentrations of the following three primary air pollutants in Chinese cities: $PM_{2.5}$, PM_{10}, and SO_2. As expected, shared bikes may partially replace private cars, reducing air pollutant emissions at the source of transportation (Huang et al., 2022; Zhang and Mi, 2018); therefore, our estimates are similar to some prior research (Fan and Zheng, 2020; Huang et al., 2022; Lu et al., 2018; Monazzam et al., 2021), which further emphasizes the positive environmental effects of shared bikes as a green mode of travel.

5.2. The effect of digital finance on shared-bike trips

We further examine the effect of digital finance on shared-bike trips, a green travel behavior. The baseline results are presented in Table 2. Considering that the explained variable is binary, we consider the Probit model as shown in Equations (2) and (3) and employ the linear probability model (LPM) as a reference. In columns (1) and (3), we control for individual and household characteristics, some city-level variables, and the district/county fixed effect, finding significantly positive coefficients on our core explanatory variable, *Digital finance*. Additionally, we add other vital macro control variables, namely *Loans*, *R&D expenditure*, and *Digitalization*, which are closely related to digital finance. Similarly, coefficients in *Digital finance* remain positive in columns (2) and (4).

Moreover, for both Probit and LPM estimations, we find that the significant effect of digital finance on shared-bike trips is statistical and economical. Our preferred specification, column (4), suggests that for every unit (100 points) increase in average digital finance, the probability of residents using shared bicycles increases by 22.37 percentage points. Furthermore, we can convert the magnitude using standard deviations. The estimates indicate that a one standard deviation increase in digital finance increases the probability of riding shared bikes by 0.1468 standard deviations. In sum, these results confirm that digital finance is environmentally friendly in China by promoting the adoption of a greener travel method, shared bikes.

5.3. IV estimations

In the main specification, we strictly control for a variety of city-level macro variables and the district/county fixed effect and adopt city-level cluster standard errors. Next, we employ the IV approach to exclude potential endogeneity. In particular, we selected two IVs: historical Internet penetration at the city level and historical average digital finance within the same province.

On the one hand, digital finance development relies on Internet platforms (Chen and Zhao, 2021; Liu et al., 2021), thus residents living in areas with better Internet infrastructure are likely to be the first group to use digital financial tools in China (Chen and Zhang, 2021; He and Li, 2020). Thus, the positive correlation between Internet penetration and digital finance is indisputable. Moreover, the state of digital finance development in a city is closely related to other cities in the same province towing to the spillover effects on digital finance (Shen et al., 2021; Wang et al., 2020); therefore, in theory, both IVs satisfy the correlation assumption.

On the other hand, as mentioned in Section 2, information and Internet technologies are essential factors in shared-bike promotions (e.g., Gao et al., 2019; Li C et al., 2021; Ma et al., 2018; Yang et al., 2020). Therefore, to ensure the exogeneity of IVs, we use historical Internet penetration and historical average digital finance within the same province, as we strictly control for district/county fixed effect and a range of city characteristics. In particular, for historical Internet penetration, our control variable set includes the Internet penetration in 2016 (*Digitalization*), which helps rule out potential endogenous factors; therefore, we believe these two IVs theoretically satisfy the exclusion restriction hypothesis. More specifically, we use the Internet penetration rate in 2002 and the digital finance aggregation index in 2011, considering the availability and comprehensiveness of the data.[12]

The results from the IV estimates are represented in Table 3. In the first two columns, we attempt to check IV availability by regressions. We find that both IVs are insignificant in column (1) but significant in column (2), suggesting that two IVs may only affect shared-bike trips through the explanatory variable, *Digital finance*; thus, these results confirm that the two IVs we select are quite statistically exogenous. The last three columns show the IV estimation results by employing the two-stage least square (2SLS) and IV-probit methods, of which the first-stage results are reported in column (3). We can see that two IVs are strongly positively correlated with digital finance, indicating that the IVs meet the correlation restriction. More importantly, the first-stage F value is greater than the critical value of 10, which avoids the weak IV problem (Stock and Yogo, 2005).

In columns (4) and (5), the second-stage results suggest that the coefficients on *Digital finance* remain significantly positive. This finding suggests that by controlling for endogeneity and adopting the IV estimations, we obtain a strict causal relationship between digital finance, green travel behavior, and shared-bike trips. Conversely, coefficient values obtained from IV estimations are larger, which may be influenced by the "LATE" property of IV approaches (e.g., Angrist and Pischke, 2008; Zhao and Guo, 2022; Zhao et al., 2022). Moreover, the Anderson–Rubin Wald test for 2SLS estimates further indicates that our IVs are quite strong; the Hansen J statistic implies that there is no over-identification problem by using two IVs. The IV estimates support our conclusion that digital finance can encourage urban residents to use shared bikes after eliminating potential endogeneity.

5.4. Mechanism tests

5.4.1. Payment convenience

As mentioned above, digital finance may facilitate shared-bike trips by providing more convenient payment methods. The CHFS questionnaire records information on whether respondents are mobile payment users. Mobile payments, an essential component of digital finance in China, can provide people with a more convenient and accessible digital payment (Zhao et al., 2022). We construct a binary mechanism variable, *Mobile payment*, which is set to 1 if the householder uses mobile payments and 0 otherwise. In column (1) of Table 4, we find that the coefficient of *Digital finance* is significantly positive, indicating that digital finance will likely encourage residents to adopt mobile payments effectively. Moreover, the estimation results in column (2) show that mobile payments increase the probability of using shared bikes. These results imply that digital finance encourages people to ride shared bikes through mobile payments.

In addition, Fig. 3 shows that the indicator system for digital finance includes several sub-indicators that measure digital payments in cities. First, the coverage breadth is mainly expressed in terms of the coverage of Alipay accounts, one of the most dominant mobile payment platforms in China. Second, a tertiary indicator of usage depth, Payment, directly

[11] In this subsection, the air pollution data are from the official statistics of the Ministry of Ecology and Environment of China; the information on the time of entry of Ofo and Mobike into each city comes from the official websites and weibo and media reports (Huang et al., 2022). In the control variable set, the city characteristics data are from the China City Statistical Yearbook and climate data are from the China Meteorological Administration.

[12] The China City Statistical Yearbook has been publishing data on city Internet users since 2001, but we adopt the data from 2002 due to too many missing values in 2001. In addition, the earliest publicly available DFII data was 2011.

Table 1
A priori study: The effect of shared-bike entry on air pollution (DID estimates).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	AQI	$PM_{2.5}$	PM_{10}	SO_2	CO	NO_2	O_3
Shared-bike entry	−0.0308	−3.5977**	−3.9856**	−3.8667**	−0.0083	−0.3204	0.4697
	(1.8091)	(1.6506)	(1.9893)	(1.7633)	(0.0305)	(0.8047)	(1.6455)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1599	1599	1599	1599	1599	1599	1599
Adj-R^2	0.7492	0.7366	0.7793	0.7481	0.7041	0.7868	0.5893

Notes: The control variables include *GDP per capita*, *Fiscal expenditure*, *Fiscal revenue*, *Road infrastructure*, *FDI*, *Enterprises*, *Financial development*, *Average temperature*, *Relative humidity*, and *Average rainfall*. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the city level are reported in parentheses.

describes the digital payment development at the macro level. In columns (3) and (4) of Table 4, not surprisingly, the estimated coefficients on *Payment* and *Coverage breadth* are both significantly positive, indicating that the payment convenience from digital finance is an important reason for promoting residents adopting a greener travel method, shared bikes.

5.4.2. Credit services

As discussed in Section 3, with the help of Internet platforms and digital technology, digital finance may promote shared-bike use by providing credit services to more residents. The CHFS database does not provide information on the credit services for shared bikes, so we use whether residents use Internet lending as a proxy mechanism variable. It is assigned to 1 if the household uses Internet lending and 0 otherwise. Internet lending is an integral part of digital finance in China, where only credit-qualified customers may receive loans (Wei, 2015). The credit qualifications required for Internet lending are more stringent than those for shared bikes. The estimates in the first two columns of Table 5 show that digital finance significantly boosts the use of Internet lending and that this credit service effect further increases the probability of people riding shared bicycles.

Consistent with Table 4, we construct variables to investigate the impact of the credit service function of digital finance on shared-bike trips. We consider that the digital finance indicator system contains some sub-indicators of digital credit services, such as credit in usage depth and creditization in digitization level; however, as mentioned in Section 4, the data for creditization are not publicly available, so we use the secondary indicator, digitization level, as a proxy mechanism variable. The estimates in columns (3) and (4) in Table 5 suggest a significant positive relationship between digital credit and shared-bike trips, while the coefficient on *Digitization level* is insignificant. These results further suggest that digital finance promotes shared-bike adoption by providing credit services. In addition, the results in column (4) are suggestive, as the measures of digitization level in the DFII are too broad to accurately capture credit services from digital finance.

5.4.3. Social interactions

Section 3 highlighted that social interactions are the last explanation for digital finance to promote residents to ride shared bikes. To fully reflect the social interactions, we construct two mechanism variables from both online and offline (Chen and Zhao, 2021; He and Li, 2020). Regarding online interactions, respondents were asked whether they used the Internet for shopping, learning information, socialization (e.g., using WeChat), entertainment (e.g., playing computer games), and selling products and services. We sum the score across these dimensions to obtain a variable in the range of integers 1 to 5, with higher scores indicating higher levels in online interactions. Moreover, given the complexity of offline interactions, we regard the offline consumption per capita (10,000 CNY) as a proxy (Chung, 2013; Kim et al., 2019).

In columns (1) and (3) of Table 6, the estimation results from the OLS models show that the coefficients on *Digital finance* are significantly positive, indicating that digital finance plays a vital role in more active social interactions among Chinese urban residents. Furthermore, in columns (2) and (4), we find a positive association between online and offline social interactions and shared-bike trips; thus, these results provide supportive evidence that digital finance is likely to promote the adoption of shared bikes as a green travel method by facilitating social interactions. Moreover, as discussed in Section 3, considering the role of the peer effect in the promotion of digital financial tools and shared bikes, we measure the peer effect in terms of the ratio of shared-bike trips of other residents within the same communities, following the previous studies (Calabrese et al., 2006; Zhao et al., 2022). In column (5) of Table 6, as expected, the OLS estimate of digital finance is positive and significant. These estimates mean that as the vital content of social interactions, peer effects also play an active role in digital finance, facilitating shared-bike trips.

5.5. Further studies: do digital finance and shared bikes replace polluting modes of travel?

As shown in Table 1, we believe that the main way bike sharing can help improve the environment is by replacing other polluting modes of travel. This subsection aims to test this effect to provide more evidence that digital finance is environmentally friendly by facilitating shared-bike trips. The CHFS records the primary means of transportation to work for respondents with jobs, including private cars, buses and subways, cabs, electric cars and motorcycles, and bicycles. Private cars are the most dominant polluting mode of travel and a significant cause of urban air pollution in China (He and Jiang, 2021; Jie et al., 2021); therefore, we construct a dummy variable that equals 1 if the respondent usually goes to work via private car.

The results in column (1) of Table 7 indicate that digital finance significantly reduces the likelihood that residents take cars to work. Furthermore, considering that some households do not have private cars, we use a subsample consisting of households with private cars (Teixeira et al., 2022); the results are shown in column (2). We find that the coefficient of *Digital finance* remains significantly negative, suggesting that for households with private cars, digital finance still reduces the probability of going to work by private cars. In column (3), we further explore the relationship between taking private cars to work and using shared bikes, finding a significant negative correlation between the two. These results are similar to the findings in some previous literature (Fuller et al., 2013; Lu et al., 2018), suggesting that digital finance may influence some Chinese residents with private cars to be shared-bike users.

The potential reasons are easy to understand. Shared bikes have been fully integrated into the urban transportation system in most Chinese cities and can effectively complement short-distance trips (Chen et al., 2020; Ma et al., 2018). In particular, the willingness of residents to choose green public transportation has increased through the seamless integration of metros or buses (Chu et al., 2021; Fan and Zheng, 2020); thus, the convenience provided by digital finance for shared bikes

Table 2
The effect of digital finance on shared-bike trips.

	(1)	(2)	(3)	(4)
	LPM	LPM	Probit	Probit
Digital finance	0.2328***	0.2735***	0.1735***	0.2237**
	(0.0686)	(0.0858)	(0.0584)	(0.0878)
Age	−0.0110***	−0.0110***	−0.0089***	−0.0089***
	(0.0019)	(0.0019)	(0.0022)	(0.0022)
Age-squared	0.0066***	0.0066***	0.0043*	0.0043*
	(0.0018)	(0.0018)	(0.0023)	(0.0023)
Hukou	−0.0242***	−0.0241***	−0.0415***	−0.0411***
	(0.0088)	(0.0088)	(0.0141)	(0.0141)
Gender	−0.0027	−0.0026	−0.0016	−0.0017
	(0.0068)	(0.0068)	(0.0085)	(0.0085)
Married	−0.0495***	−0.0491***	−0.0537***	−0.0535***
	(0.0103)	(0.0104)	(0.0119)	(0.0120)
Education	0.0218***	0.0216***	0.0268***	0.0265***
	(0.0045)	(0.0045)	(0.0055)	(0.0055)
Self-rated health	0.0030	0.0030	0.0095**	0.0095**
	(0.0031)	(0.0031)	(0.0044)	(0.0044)
Employment	0.0333***	0.0335***	0.0452***	0.0457***
	(0.0091)	(0.0091)	(0.0115)	(0.0115)
Household income	−0.0024	−0.0025	−0.0038	−0.0039
	(0.0020)	(0.0020)	(0.0024)	(0.0024)
Household consumption	0.0057***	0.0057***	0.0033*	0.0033*
	(0.0020)	(0.0020)	(0.0019)	(0.0019)
Transportation expenses	0.0076***	0.0076***	0.0132***	0.0132***
	(0.0015)	(0.0015)	(0.0020)	(0.0020)
Household savings	0.0008***	0.0008***	0.0006**	0.0006**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Car ownership	0.0138	0.0140	0.0220*	0.0223*
	(0.0099)	(0.0099)	(0.0116)	(0.0116)
House ownership	−0.0027	−0.0021	−0.0012	−0.0004
	(0.0118)	(0.0119)	(0.0141)	(0.0142)
GDP per capita	−0.0054	−0.0098**	−0.0072*	−0.0106**
	(0.0045)	(0.0050)	(0.0038)	(0.0043)
GDP growth rate	0.0035	0.0035	0.0034	0.0034
	(0.0026)	(0.0027)	(0.0022)	(0.0024)
Buses per capita	0.0311	0.0452	0.0302	0.0460
	(0.0242)	(0.0312)	(0.0199)	(0.0295)
Taxis per capita	−0.0101	0.0092	−0.0030	0.0114
	(0.0156)	(0.0190)	(0.0148)	(0.0193)
Scope	0.0012*	0.0013**	0.0009*	0.0010**
	(0.0006)	(0.0006)	(0.0005)	(0.0005)
Salary per capita	−0.0319***	−0.0238*	−0.0158*	−0.0099
	(0.0106)	(0.0129)	(0.0091)	(0.0119)
Loans		−0.0417*		−0.0335*
		(0.0224)		(0.0198)
R&D expenditure		0.0068		0.0054
		(0.0146)		(0.0138)
Digitalization		0.0056		−0.0277
		(0.1347)		(0.1429)
District/county fixed effect	Yes	Yes	Yes	Yes
N	9434	9434	9434	9434

Notes: Columns (3) and (4) report the marginal effects by using the Probit model. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the city level are reported in parentheses.

Table 3
Digital finance and shared-bike trips: IV estimations.

	(1)	(2)	(3)	(4)	(5)
	Probit		IV estimations		
			First stage	Second stage for 2SLS	Second stage for IV-probit
	Shared bike	Shared bike	Digital finance	Shared bike	Shared bike
Digital finance	0.2739***			1.4170**	0.8943*
	(0.0901)			(0.6673)	(0.4731)
Internet penetration in 2002	−0.0115	0.1633**	0.0898***		
	(0.0255)	(0.0712)	(0.0149)		
Average digital finance in 2011	0.1928	0.2832**	0.1696**		
	(0.1205)	(0.1267)	(0.0789)		
First stage F value			19.4328		
Anderson-Rubin Wald test				5.8432**	
				[0.0161]	
Hansen J statistic				1.3012	
				[0.2561]	
Baseline controls	Yes	Yes	Yes	Yes	Yes
District/county fixed effect	Yes	Yes	Yes	Yes	Yes
N	9318	9318	9318	9318	9318

Notes: Columns (1), (2), and (5) report the marginal effects. Baseline controls remain consistent with column (4) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the city level are reported in parentheses. The p-values of the Anderson–Rubin Wald test and Hansen J statistic are shown in the bracket.

Table 4
Digital finance, payment convenience, and shared-bike trips.

	(1)	(2)	(3)	(4)
	Mobile payment	Shared bike	Shared bike	Shared bike
Digital finance	0.1798**			
	(0.0900)			
Mobile payment		0.0929***		
		(0.0120)		
Payment			0.1015***	
			(0.0323)	
Coverage breadth				0.0018***
				(0.0006)
Baseline controls	Yes	Yes	Yes	Yes
District/county fixed effect	Yes	Yes	Yes	Yes
N	9421	9421	9421	9421

Notes: All columns report the marginal effects using the Probit model. Baseline controls remain consistent with column (4) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the city level are reported in parentheses.

further exacerbates its role as an alternative to polluting modes of travel, such as private cars. Moreover, we estimate the effect of digital finance on cycling to work, finding a positive coefficient for digital finance in column (4) of Table 7; therefore, using an employment subsample, we further reveal that digital finance contributes to the urban environment by replacing polluting modes of travel and stimulating greener trips.

5.6. Heterogeneity analysis

Previous literature confirmed that digital finance is an inclusive financial model that can benefit more socially disadvantaged groups (e.g., Chen and Zhao, 2021; Lai et al., 2020; Li et al., 2020; Zhao et al., 2022). Whether digital finance is inclusive regarding green trips needs to be further tested; therefore, we try to empirically check this effect through heterogeneity analysis from both micro and macro aspects.

5.6.1. Heterogeneity effects at the micro level

Since our sample includes respondents of different ages, there are differences between the young, middle-aged, and elderly groups, both in terms of digital finance acceptance and motivation to ride shared bikes (Efthymiou et al., 2013; Fishman et al., 2014; Liu et al., 2021; Rixey, 2013). Following the standard of the China Bureau of Statistics, the young group comprises residents aged 35 and younger, middle-aged residents are 36–59, and those aged 60 and above comprise the

Table 5
Digital finance, credit services, and shared-bike trips.

	(1)	(2)	(3)	(4)
	Internet lending	Shared bike	Shared bike	Shared bike
Digital finance	0.1223**			
	(0.0584)			
Internet lending		0.0431***		
		(0.0146)		
Credit			0.3773**	
			(0.1563)	
Digitization level				0.1169
				(0.0924)
Baseline controls	Yes	Yes	Yes	Yes
District/county fixed effect	Yes	Yes	Yes	Yes
N	6001	6001	9421	9421

Notes: All columns report the marginal effects busing the Probit model. Baseline controls remain consistent with column (4) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the city level are reported in parentheses.

elderly group. The first three columns of Table 8 show that although the coefficients in *Digital finance* are all positive, they are only significant in column (1). The potential reasons are manifold. First, the young group is more receptive to new things such as digital finance and shared bikes, and the convenience brought by digital finance provides them with a good solution for short-distance trips (Picasso et al., 2020). Second, urban middle-aged people have more travel options in their daily life. They are less sensitive to routine travel costs and do not have a high level of digital finance (Luo and Zeng, 2020). Third, in column (3), we find the potential of digital finance in promoting green travel modes for the elderly group; however, subject to the digital divide, they may need to spend more transaction costs on using digital financial tools and shared bikes (Vicente, 2022).

In addition, we analyze the education heterogeneity of the impact of digital finance on shared-bike trips. We define the vulnerable low-education group as those whose education level is senior high school and below. The empirical results in column (4) of Table 8 show that, in the subsample of low-education urban householders, digital finance has a more significant promotion effect on shared-bike trips; however, column (5) indicates that the estimated coefficient on *Digital finance* is insignificant for residents with relatively high levels of education. These results imply the inclusive nature of digital finance in China, which is more conducive to enhancing the welfare of educationally disadvantaged groups.

Finally, we test the income heterogeneity in columns (6) and (7) of Table 8. According to the conventional definition, those with household income per capita below the median are considered to be in the low-income group (Zhao et al., 2022). Low-income urban households are relatively vulnerable groups. Not surprisingly, the subsample estimates indicate that digital finance has a more positive effect on low-income residents riding shared bikes, in line with the study by Murphy and Usher (2015). Most urban households with relatively low incomes have more access to public transportation in China. Digital financial tools provide convenience for shared bikes to solve their last-mile problem.

5.6.2. *Heterogeneity effects at the macro level*

We further consider macro-level heterogeneity effects. First, based on the city GDP per capita median, we divide the whole sample into two subsamples. Columns (1) and (2) of Table 9 show that the estimated coefficients on *Digital finance* are significant only in column (2), indicating that the positive impact of digital finance on residents' green travel is mainly in less developed regions. Second, considering that digital finance is closely related to the level of financial development and informatization, as before, we conduct subsample heterogeneity tests based on the median of the two variables, *Loans* and *Informatization*, as the boundaries. In the last four columns of Table 9, the estimates show that the positive impact of digital finance on residents using shared bikes is mainly found in cities with low financial development and informatization levels.

These results further illustrate the inclusive nature of digital finance from a macro perspective. For cities with relatively low levels of economic and financial development and informatization, the development of digital finance has brought more options for people, particularly replacing some financial services provided by traditional financial institutions with high thresholds and transaction costs (Lai et al., 2020;

Table 7
Digital finance and polluting modes of travel.

	(1)	(2)	(3)	(4)
	Taking private cars to work	Taking private cars to work (subsample with cars)	Shared bike	Cycling to work
Digital finance	−0.1331*	−0.6269**		0.0859
	(0.0712)	(0.3185)		(0.0531)
Private car travel			−0.0496***	
			(0.0145)	
Baseline controls	Yes	Yes	Yes	Yes
District/county fixed effect	Yes	Yes	Yes	Yes
N	4536	1992	4520	4520

Notes: All columns report the marginal effects using the Probit model. Baseline controls remain consistent with column (4) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the city level are reported in parentheses.

Table 6
Digital finance, social interactions, and shared-bike trips.

	(1)	(2)	(3)	(4)	(5)
	OLS	Probit	OLS	Probit	OLS
	Online interaction	Shared bike	Offline interaction	Shared bike	Shared bike trip ratio
Digital finance	0.1361**		0.6247***		0.0562*
	(0.0673)		(0.1615)		(0.0297)
Offline interaction		0.0436***			
		(0.0068)			
Offline interaction				0.0029***	
				(0.0008)	
Baseline controls	Yes	Yes	Yes	Yes	Yes
District/county fixed effect	Yes	Yes	Yes	Yes	Yes
N	7413	7413	8716	8716	9434

Notes: Columns (2) and (4) report the marginal effects using the Probit model. Baseline controls remain consistent with column (4) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the city level are reported in parentheses.

Table 8
Heterogeneity effects at the individual and household levels.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Age			Education		Household income	
	The young	The middle-aged	The elderly	Relatively low education	Relatively high education	Relatively high income	Relatively low income
Digital finance	0.2456***	0.1740	0.2182	0.2616***	0.1380	0.1998	0.2934**
	(0.0912)	(0.1145)	(0.2871)	(0.0868)	(0.1995)	(0.1370)	(0.1141)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District/county fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2065	4877	2501	6739	2695	4706	4728

Notes: All columns report the marginal effects using the Probit model. Baseline controls remain consistent with column (4) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the city level are reported in parentheses.

Table 9
Heterogeneity effects at the city level.

	(1)	(2)	(3)	(4)	(5)	(6)
	GDP per capita		Financial development (loans)		Digitalization level	
	Relatively high	Relatively low	Better	Worse	Relatively high	Relatively low
Digital finance	0.2335	0.3211*	0.1097	0.5317***	0.0532	0.3995**
	(0.3888)	(0.1824)	(0.1452)	(0.1791)	(0.2417)	(0.1622)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
District/county fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
N	4637	4797	4714	4720	4710	4724

Notes: All columns report the marginal effects using the Probit model. Baseline controls remain consistent with column (4) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the city level are reported in parentheses.

Yin et al., 2019). As a result, the marginal utility generated by digital finance is higher, and residents in these cities are more encouraged to use the emerging green travel method, shared bikes.

5.7. Robustness checks

To confirm the reliability of the baseline results, we further performed some robustness tests. First, as emphasized in Section 4, we utilize the usual time residents spent riding shared bikes as a proxy-explained variable (*Cycling duration*). In column (1) of Table 10, the coefficient on *Digital finance* is significantly positive, suggesting that our main results are reliable after replacing the core explained variable.

Second, considering the effects of age and health on shared-bike use (Buck et al., 2013; Orvin and Fatmi, 2021), we further control for age and self-rated health fixed effects. Column (2) of Table 10 shows that the estimates of digital finance are quite insensitive to these fixed effects. Moreover, we add the high-dimensional fixed effect. In particular, considering that the behavior of riding shared bikes may differ greatly over age and between regions, we try to include the age-by-city fixed effect for a robustness check in column (3) of Table 10. As expected, our main specification is satisfactory for capturing this high-dimensional fixed effect.

Third, considering that the explanatory variable involves micro-individual behavior, which may be influenced by some community factors (Scannell and Gifford, 2010)—we attempt to control for some community characteristic variables in column (4) of Table 10. In particular, we include the average income, average consumption, and car ownership of households in the same community, as well as the average age, average education, and employment rate of householders. The inclusion of these community variables has little effect on our main results.

Fourth, since our explained variable is located at the individual level, the use of clustering standard errors at the city level may miss heteroscedasticity problems at the individual level. Adopting standard clustering errors at lower levels may help to reduce the potential bias (Cameron and Miller, 2015). Column (5) of Table 10 reports clustering standard errors at the individual level, and we find that the coefficient on *Digital finance* remains unchanged.

Finally, we conducted robustness tests by removing some extreme observations; the results are shown in Table 11. In column (1), we exclude the non-employed observations, owing to the greater daily trip needs of most commuters. In column (2), we remove observations of

Table 10
Robustness checks by using alternative specifications.

	(1)	(2)	(3)	(4)	(5)
	Cycling duration	Controlling for age and self-rated health fixed effects	Controlling for high dimensional fixed effects	Controlling for community-level variables	Standard errors clustered at the individual level
Digital finance	OLS	Probit	Probit	Probit	Probit
	3.0144***	0.2728***	0.3091**	0.3222***	0.2237***
	(1.1193)	(0.0844)	(0.1328)	(0.0973)	(0.0863)
Baseline controls	Yes	Yes	Yes	Yes	Yes
District/county fixed effect	Yes	Yes	Yes	Yes	Yes
N	9400	9434	9434	9434	9434

Notes: Columns (2)–(5) report the marginal effects using the Probit model. Baseline controls remain consistent with column (4) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the city level are reported in parentheses.

Table 11
Robustness checks by removing extreme observations.

	(1)	(2)	(3)	(4)	(5)
	Excluding the non-employed observations	Excluding unhealthy or older householders	Excluding households without transportation expenses	Excluding samples with severe air pollution in their cities	Excluding samples with high slopes in their cities
Digital finance	0.2581***	0.2731***	0.3334***	0.2860***	0.3281***
	(0.0966)	(0.0874)	(0.0950)	(0.1010)	(0.0981)
Baseline controls	Yes	Yes	Yes	Yes	Yes
District/county fixed effect	Yes	Yes	Yes	Yes	Yes
N	6204	8329	7774	7027	7678

Notes: All columns report the marginal effects using the Probit model. Baseline controls remain consistent with column (4) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the city level are reported in parentheses.

people over 80 and those with very unhealthy status, considering that these residents may not meet the conditions of using shared bikes in terms of physical function. In column (3), we delete households without transportation expenditures because most shared bikes in China are commercial and not free. In column (4), we remove cities in our sample with average slopes above the 75% quantile, considering that some cities in China are unsuitable for bike trips due to their high slopes (e.g., Chongqing). In column (5), cities with severe air pollution ($PM_{2.5}$ concentrations above the 75% quantile) are excluded because, as an outdoor activity, the more polluted the air is, the less likely people are to ride shared bikes (Hong et al., 2022; Zhao P et al., 2018). These coefficients in *Digital finance* keep significantly positive.

6. Conclusions and discussion

6.1. Main findings

In recent years, the traffic pollution problem in China has received attention from all walks of life. As a green travel method, shared bikes can play an essential role in improving environmental benefits. Previous studies on the influencing factors of shared-bike adoption rarely addressed digital finance (e.g., Gao et al., 2019; Li C et al., 2021; Ma et al., 2018; Orvin and Fatmi, 2021). This study first identifies the positive impact of shared bikes on reducing environmental pollution using the DID method based on a quasi-natural experiment in which shared bikes entered Chinese cities for the first time. Subsequently, by matching digital financial indexes at the city level and urban household microdata from the CHFS in 2017, we further explore the impact of digital finance on shared-bike trips and determine whether digital finance in China is environmentally friendly.

We find that digital finance significantly facilitates shared-bike trips, supported by IV estimates and other robustness tests. In particular, our main specification suggests that a one standard deviation increase in digital finance increases the probability of riding shared bikes by 0.1468 standard deviations. Moreover, we explore several potential mechanisms by which digital finance affects shared-bike trips, finding that payment convenience, credit services, and social interactions can explain the positive impact of digital finance. Interestingly, our investigation confirms that the positive effect of digital finance on shared bikes can replace private car trips for Chinese urban households, suggesting the environmental friendliness of digital finance by replacing polluting modes of travel. Furthermore, our estimates show that this positive effect of digital finance is valid for some socially vulnerable groups, such as residents with low education and low income living and in cities with less developed economic and financial development or lower digitalization levels. These results indicate the inclusiveness of digital finance in China, providing short-distance green travel options to more socially disadvantaged groups.

6.2. Theoretical implications

The rapid development of digital finance in China and its benefits have received academic attention (e.g., Chen and Zhao, 2021; Lai et al., 2020; Li et al., 2020; Liu et al., 2021; Luo and Zeng, 2020; Shen et al., 2021); however, they were rarely linked to the green travel behavior. Based on the particular perspective of shared bikes, we provide evidence that digital finance is environmentally friendly in China. Additionally, as the shared bike is a green mode of travel, factors influencing people's adoption of bike-sharing trips have been the focus of scholars' attention (e.g., Gao et al., 2019; Godavarthy and Taleqani, 2017; Li C et al., 2021; Li et al., 2019; Ricci, 2015; Winters et al., 2016; Yang et al., 2020; Yao et al., 2019). Our findings explain a reason for the rapid growth of shared bikes in China from a new field, digital finance.

Furthermore, our results suggest that thanks to the convenience of payment and credit services provided by digital finance and the social interaction effect, the shared bike, as a green travel way, has been successful in China; therefore, we provide a reference for subsequent studies exploring the mechanisms of the impact of digital finance on green travel patterns. Moreover, this paper also complements the research field on the inclusiveness of digital finance (Chen and Zhao, 2021; Lai et al., 2020; Liu et al., 2021; Zhao et al., 2022).

6.3. Practical implications

This study has important policy implications. First, our study informed the active promotion of digital finance in developing countries to enhance residents' willingness to travel green. The empirical evidence in this paper suggests that the environmental benefits of digital finance can be effective in developing countries like China; therefore, it is essential to further extend these experiences to more countries and regions to exploit their environmentally friendly benefits by improving policies related to the development of digital finance and technology finance. Second, our findings indicate that using digital finance and bike sharing is more helpful for more socially vulnerable groups, thus it is vital to actively adjust the development strategy of digital finance to make it more suitable for these groups. Possible approaches include strengthening promotion, lowering the threshold for use, and applying discounts to particular groups.

6.4. Limitations and future directions

This paper has some limitations. First, subject to the availability of microdata on shared-bike trips, we only used the cross-sectional data from the fourth wave of CHFS in 2017. In a subsequent study, panel data can be constructed to reflect the dynamics of shared-bike trips in digital finance development. Second, as the survey on bicycle sharing in CHFS data is relatively simple, little information exists concerning residents' willingness to use bicycle sharing, subjective attitudes, potential risks, and timing of use, which follow-up studies should further supplement. Third, future research is also needed to develop other mechanisms of digital finance affecting residents' green travel behaviors.

Declaration of competing interest

The authors have no conflicts of interest.

Data availability

Data will be made available on request.

Acknowledgements

We appreciate the Survey and Research Center for China Household Finance at Southwest University of Finance and Economics for providing us with the China Household Finance Survey (CHFS) data. This work was supported by the Humanities and Social Sciences Youth Foundation, Ministry of Education of the People's Republic of China (No. 22YJC790030), the Natural Science Foundation of Hunan Province (No. 2021JJ40546), and the Scientific Research Project of Hunan Provincial Education Department (21B0108).

Appendix

Table A1
Variables definitions and descriptive statistics for shared-bike entry and air pollution reduction (N = 1599)

Variables	Definitions	Mean	SD
AQI	A dimensionless index that contains six air pollutants to quantitatively describe air quality conditions	91.1085	35.6520
$PM_{2.5}$	Average concentrations of $PM_{2.5}$ ($\mu g/m^3$)	58.5741	31.6759
PM_{10}	Average concentration of PM_{10} ($\mu g/m^3$)	94.6735	43.7986
SO_2	Average concentration of SO_2 ($\mu g/m^3$)	26.7543	23.3487
CO	Average concentration of CO (mg/m^3)	1.1536	0.4670
NO_2	Average concentration of NO_2 ($\mu g/m^3$)	35.0221	13.9692
O_3	Average concentration of O_3 ($\mu g/m^3$)	79.7982	21.7688
Shared-bike entry	= 1 if shared bikes have entered a city, 0 otherwise	0.1063	0.3083
GDP per capita	Total GDP divided by population (10,000 CNY/person)	5.6261	3.3074
Fiscal expenditure	Fiscal expenditures divided by GDP	0.1954	0.0829
Fiscal revenue	Fiscal revenue divided by GDP	0.0793	0.0261
Road infrastructure	Annual road freight volume (100 million tons)	1.2751	1.0696
FDI	Amount of foreign direct investment divided by GDP	1.5676	1.5196
Enterprise	Number of enterprises above the scale (logarithmic)	6.7367	1.0025
Financial development	Total loans and total deposits of financial institutions divided by GDP	2.5111	1.1025
Average temperature	Annual average temperature (°C)	15.5986	4.8386
Relative humidity	Relative humidity	0.6942	0.0952
Average rainfall	Average annual rainfall (100 mm)	0.2499	1.6440

Table A2
Variables definition and descriptive statistics for digital finance and shared-bike trips (N = 9434)

Variables	Definitions	Mean	SD
Shared bike	= 1 if the householder had shared-bike trips in the last week, 0 otherwise	0.1481	0.3552
Digital finance	Digital finance aggregation index at the city level (divided by 100)	2.4085	0.2331
Age	Age of the householder	50.3774	13.5929
Age-squared	Square of the age of the householder (divided by 100)	27.2262	14.0815
Hukou	The *hukou* status of the householder: Rural = 1; urban = 0	0.2974	0.4572
Gender	Gender of the householder: Female = 1; male = 0	0.2423	0.4285
Married	= 1 if the householder is married and has a spouse, 0 otherwise	0.8583	0.3488
Education	Unschooled = 1; primary school = 2; junior middle school = 3; senior or technical high school = 4; college, vocational school, or bachelor's degree = 5; master's degree or doctorate degree = 6	3.7276	1.0709
Self-rated health	Very bad = 1; bad = 2; indifferent = 3; good = 4; very good = 5	3.6485	0.9102
Employment	= 1 if the householder has a job: 0 otherwise	0.6587	0.4742
Household income	Household income per capita (10,000 CNY)	4.3166	2.2278
Household consumption	Household consumption per capita (10,000 CNY)	2.8244	2.6395
Transportation expenses	Household total transportation expenses, logarithm, the CNY	6.4260	3.2480
Household savings	Sum of household fixed and current deposits (10,000 CNY)	7.0764	15.0649
Car ownership	Yes = 1; no = 0	0.3759	0.4844
House ownership	Yes = 1; no = 0	0.8018	0.3987
GDP per capita	Total GDP divided by population (10,000 CNY/person)	10.1389	8.0106
GDP growth rate	Year-on-year growth rate of city GDP (%)	6.9515	4.5349
Buses per capita	Number of buses per 10,000 people	0.8273	1.1235
Taxis per capita	Number of taxis per 10,000 people	1.5902	1.3283
Scope	Average slope of the city	2.1696	1.7316
Salary per capita	Annual salary per capita in the city (1000 CNY)	7.2232	2.1306
Loans	The ratio of total loans from financial institutions to GDP	2.1674	1.0364
R&D expenditure	The ratio of government spending on R&D to GDP (‰)	2.0028	1.1442
Digitalization	The percentage of household Internet access in the city	0.3333	0.2286

References

Acemoglu, D., Moscona, J., Robinson, J.A., 2016. State capacity and American technology: evidence from the nineteenth century. Am. Econ. Rev. 106 (5), 61–67.

Angrist, J.D., Pischke, J., 2008. Mostly Harmless Econometrics: an Empiricist's Companion. Princeton University Press.

Apparicio, P., Carrier, M., Gelb, J., Séguin, A.M., Kingham, S., 2016. Cyclists' exposure to air pollution and road traffic noise in central city neighbourhoods of Montreal. J. Transport Geogr. 57, 63–69.

Barbour, N., Zhang, Y., Mannering, F., 2019. A statistical analysis of bike sharing usage and its potential as an auto-trip substitute. J. Transport Health 12, 253–262.

Bonilla-Alicea, R.J., Watson, B.C., Shen, Z., Tamayo, L., Telenko, C., 2020. Life cycle assessment to quantify the impact of technology improvements in bike-sharing systems. J. Ind. Ecol. 24 (1), 138–148.

Bramoullé, Y., Djebbari, H., Fortin, B., 2009. Identification of peer effects through social networks. J. Econom. 150 (1), 41–55.

Brunekreef, B., Holgate, S.T., 2002. Air pollution and health. Lancet 360 (9341), 1233–1242.

Buck, D., Buehler, R., Happ, P., Rawls, B., Chung, P., Borecki, N., 2013. Are bikeshare users different from regular cyclists? A first look at short-term users, annual members, and area cyclists in the Washington, DC, region. Transport. Res. Rec. 2387 (1), 112–119.

Buekers, J., Dons, E., Elen, B., Panis, L.I., 2015. Health impact model for modal shift from car use to cycling or walking in Flanders: application to two bicycle highways. J. Transport Health 2 (4), 549–562.

Calabrese, S., Epple, D., Romer, T., Sieg, H., 2006. Local public good provision: voting, peer effects, and mobility. J. Publ. Econ. 90 (6–7), 959–981.

Cameron, A.C., Miller, D.L., 2015. A practitioner's guide to cluster-robust inference. J. Hum. Resour. 50 (2), 317–372.

Campbell, A.A., Cherry, C.R., Ryerson, M.S., Yang, X., 2016. Factors influencing the choice of shared bicycles and shared electric bikes in Beijing. Transport. Res. C Emerg. Technol. 67, 399–414.

Campbell, K.B., Brakewood, C., 2017. Sharing riders: how bikesharing impacts bus ridership in New York City. Transport. Res. Pol. Pract. 100, 264–282.

Cao, Y., Shen, D., 2019. Contribution of shared bikes to carbon dioxide emission reduction and the economy in Beijing. Sustain. Cities Soc. 51, 101749.

Chen, B., Zhao, C., 2021. Poverty reduction in rural China: does the digital finance matter? PLoS One 16 (12), e0261214.

Chen, C.F., Huang, C.Y., 2021. Investigating the effects of a shared bike for tourism use on the tourist experience and its consequences. Curr. Issues Tourism 24 (1), 134–148.

Chen, S., Zhang, H., 2021. Does digital finance promote manufacturing servitization: micro evidence from China. Int. Rev. Econ. Finance 76, 856–869.

Chen, S.Y., Lu, C.C., 2016. A model of green acceptance and intentions to use bike-sharing: YouBike users in Taiwan. Network. Spatial Econ. 16 (4), 1103–1124.

Chen, W., Liu, Q., Zhang, C., Mi, Z., Zhu, D., Liu, G., 2020. Characterizing the stocks, flows, and carbon impact of dockless sharing bikes in China. Resour. Conserv. Recycl. 162, 105038.

Cheng, B., Li, J., Su, H., Lu, K., Chen, H., Huang, J., 2022. Life cycle assessment of greenhouse gas emission reduction through bike-sharing for sustainable cities. Sustain. Energy Technol. Assessments 53, 102789.

Chu, J., Duan, Y., Yang, X., Wang, L., 2021. The last mile matters: impact of dockless bike sharing on subway housing price premium. Manag. Sci. 67 (1), 297–316.

Chung, J.E., 2013. Social interaction in online support groups: preference for online social interaction over offline social interaction. Comput. Hum. Behav. 29 (4), 1408–1414.

Creutzig, F., Jochem, P., Edelenbosch, O.Y., Mattauch, L., Vuuren, D.P.V., McCollum, D., Minx, J., 2015. Transport: a roadblock to climate change mitigation? Science 350 (6263), 911–912.

Ebrahimi, Z.D., Momenitabar, M., Nasri, A.A., Mattson, J., 2022. Using a GIS-based spatial approach to determine the optimal locations of bikeshare stations: the case of Washington DC. Transport Pol. 127, 48–60.

Eren, E., Uz, V.E., 2020. A review on bike-sharing: the factors affecting bike-sharing demand. Sustain. Cities Soc. 54, 101882.

Efthymiou, D., Antoniou, C., Waddell, P., 2013. Factors affecting the adoption of vehicle sharing systems by young drivers. Transport Pol. 29, 64–73.

Faghih-Imani, A., Eluru, N., El-Geneidy, A.M., Rabbat, M., Haq, U., 2014. How land-use and urban form impact bicycle flows: evidence from the bicycle-sharing system (BIXI) in Montreal. J. Transport Geogr. 41, 306–314.

Fan, Y., Zheng, S., 2020. Dockless bike sharing alleviates road congestion by complementing subway travel: evidence from Beijing. Cities 107, 102895.

Fishman, E., Washington, S., Haworth, N., 2013. Bike share: a synthesis of the literature. Transport Rev. 33 (2), 148–165.

Fishman, E., Washington, S., Haworth, N., 2014. Bike share's impact on car use: evidence from the United States, Great Britain, and Australia. Transport. Res. Transport Environ. 31, 13–20.

Fricker, C., Gast, N., 2016. Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. EURO Journal on Transportation and Logistics 5 (3), 261–291.

Fuller, D., Gauvin, L., Morency, P., Kestens, Y., Drouin, L., 2013. The impact of implementing a public bicycle share program on the likelihood of collisions and near misses in Montreal, Canada. Prev. Med. 57 (6), 920–924.

Gao, S., Li, Y., Guo, H., 2019. Understanding the adoption of bike sharing systems: by combining technology diffusion theories and perceived risk. Journal of Hospitality and Tourism Technology 10 (3), 464–478.

Giallouros, G., Kouis, P., Papatheodorou, S.I., Woodcock, J., Tainio, M., 2020. The long-term impact of restricting cycling and walking during high air pollution days on all-cause mortality: health impact Assessment study. Environ. Int. 140, 105679.

Godavarthy, R.P., Taleqani, A.R., 2017. Winter bikesharing in US: user willingness, and operator's challenges and best practices. Sustain. Cities Soc. 30, 254–262.

Goodman, A., Cheshire, J., 2014. Inequalities in the London bicycle sharing system revisited: impacts of extending the scheme to poorer areas but then doubling prices. J. Transport Geogr. 41, 272–279.

Guo, Y., Zhou, J., Wu, Y., Li, Z., 2017. Identifying the factors affecting bike-sharing usage and degree of satisfaction in Ningbo, China. PLoS One 12 (9), e0185100.

Han, S.S., 2020. The spatial spread of dockless bike-sharing programs among Chinese cities. J. Transport Geogr. 86, 102782.

He, J., Li, Q., 2020. Can online social interaction improve the digital finance participation of rural households? China Agric. Econ. Rev. 12 (2), 295–312.

He, X., Jiang, S., 2021. Effects of vehicle purchase restrictions on urban air quality: empirical study on cities in China. Energy Pol. 148, 112001.

Hernández-Pérez, H., Salazar-González, J.J., 2004. A branch-and-cut algorithm for a traveling salesman problem with pickup and delivery. Discrete Appl. Math. 145 (1), 126–139.

Hong, J., McArthur, D.P., Sim, J., Kim, C.H., 2022. Did air pollution continue to affect bike share usage in Seoul during the COVID-19 pandemic? J. Transport Health 24, 101342.

Huang, G., Zhang, W., Xu, D., 2022. How do technology-enabled bike-sharing services improve urban air pollution? Empirical evidence from China. J. Clean. Prod. 379, 134771.

Huang, Y., Wang, X., Wang, X., 2020. Mobile payment in China: practice and its effects. Asian Econ. Pap. 19 (3), 1–18.

Jie, F., Standing, C., Biermann, S., Standing, S., Le, T., 2021. Factors affecting the adoption of shared mobility systems: evidence from Australia. Research in Transportation Business & Management 41, 100651.

Kim, J., Kim, M., Choi, J., Trivedi, M., 2019. Offline social interactions and online shopping demand: does the degree of social interactions matter? J. Bus. Res. 99, 373–381.

Lai, J.T., Yan, I.K., Yi, X., Zhang, H., 2020. Digital financial inclusion and consumption smoothing in China. China World Econ. 28 (1), 64–93.

Li, C., Xiao, W., Zhang, D., Ji, Q., 2021. Low-carbon transformation of cities: understanding the demand for dockless bike sharing in China. Energy Pol. 159, 112631.

Li, J., Song, Q., Wu, Y., Huang, B., 2021. The effects of online consumer credit on household consumption level and structure: evidence from China. J. Consum. Aff. 55 (4), 1614–1632.

Li, J., Wu, Y., Xiao, J.J., 2020. The impact of digital finance on household consumption: evidence from China. Econ. Modell. 86, 317–326.

Li, W., Tian, L., Gao, X., Batool, H., 2019. Effects of dockless bike-sharing system on public bike system: case study in Nanjing, China. Energy Proc. 158, 3754–3759.

Li, X., Zhao, C., 2022. Can National Forest City Construction Mitigate Air Pollution in China? Evidence from a Quasi-Natural Experiment. Environmental Geochemistry and Health. https://doi.org/10.1007/s10653-022-01386-7.

Lipp, J., 2007. Lessons for effective renewable electricity policy from Denmark, Germany and the United Kingdom. Energy Pol. 35 (11), 5481–5495.

Liu, Y., Liu, C., Zhou, M., 2021. Does digital inclusive finance promote agricultural production for rural households in China? Research based on the Chinese family database (CFD). China Agric. Econ. Rev. 13 (2), 475–494.

Lu, M., Hsu, S.C., Chen, P.C., Lee, W.Y., 2018. Improving the sustainability of integrated transportation system with bike-sharing: a spatial agent-based approach. Sustain. Cities Soc. 41, 44–51.

Luo, H., Kou, Z., Zhao, F., Cai, H., 2019. Comparative life cycle assessment of station-based and dock-less bike sharing systems. Resour. Conserv. Recycl. 146, 180–189.

Luo, H., Zhao, F., Chen, W.Q., Cai, H., 2020. Optimizing bike sharing systems from the life cycle greenhouse gas emissions perspective. Transport. Res. C Emerg. Technol. 117, 102705.

Luo, Y., Zeng, L., 2020. Digital financial capabilities and household entrepreneurship. Economic and Political Studies 8 (2), 165–202.

Ma, L., Zhang, X., Ding, X., Wang, G., 2018. Bike sharing and users' subjective well-being: an empirical study in China. Transport. Res. Pol. Pract. 118, 14–24.

Mao, G., Hou, T., Liu, X., Zuo, J., Kiyawa, A.H.I., Shi, P., Sandhu, S., 2021. How can bicycle-sharing have a sustainable future? A research based on life cycle assessment. J. Clean. Prod. 282, 125081.

Molina-García, J., Castillo, I., Queralt, A., Sallis, J.F., 2015. Bicycling to university: evaluation of a bicycle-sharing program in Spain. Health Promot. Int. 30 (2), 350–358.

Monazzam, M.R., Karimi, E., Shahbazi, H., Shahidzadeh, H., 2021. Effect of cycling development as a non-motorized transport on reducing air and noise pollution-case study: central districts of Tehran. Urban Clim. 38, 100887.

Murphy, E., Usher, J., 2015. The role of bicycle-sharing in the city: analysis of the Irish experience. International Journal of Sustainable Transportation 9 (2), 116–125.

Orvin, M., Fatmi, M.R., 2021. Why individuals choose dockless bike sharing services. Travel Behaviour and Society 22, 199–206.

Ozili, P.K., 2018. Impact of digital finance on financial inclusion and stability. Borsa Istanbul Review 18 (4), 329–340.

Papanikolaou, D., 2021. The cost of sharing bikes and docks: quantifying rent and work requirements between four bike sharing systems in North America. Travel Behaviour and Society 25, 193–208.

Picasso, E., Postorino, M.N., Bonoli-Escobar, M., Stewart-Harris, M., 2020. Car-sharing vs bike-sharing: a choice experiment to understand young people behaviour. Transport Pol. 97 (C), 121–128.

Ricci, M., 2015. Bike sharing: a review of evidence on impacts and processes of implementation and operation. Research in Transportation Business & Management 15, 28–38.
Rixey, R.A., 2013. Station-level forecasting of bikesharing ridership: station network effects in three US systems. Transport. Res. Rec. 2387 (1), 46–55.
Scannell, L., Gifford, R., 2010. The relations between natural and civic place attachment and pro-environmental behavior. J. Environ. Psychol. 30 (3), 289–297.
Shaheen, S.A., Zhang, H., Martin, E., Guzman, S., 2011. China's Hangzhou public bicycle: understanding early adoption and behavioral response to bikesharing. Transport. Res. Rec. 2247 (1), 33–41.
Shen, Y., Hueng, C.J., Hu, W., 2021. Measurement and spillover effect of digital financial inclusion: a cross-country analysis. Appl. Econ. Lett. 28 (20), 1738–1743.
Stock, J.H., Yogo, M., 2005. Testing for weak instruments in linear IV regression. In: Identification and Inference for Econometric Models: Essay in Honor of Thomas Rothenberg. Cambridge University Press.
Sun, S., Ertz, M., 2021. Contribution of bike-sharing to urban resource conservation: the case of free-floating bike-sharing. J. Clean. Prod. 280, 124416.
Teixeira, J.F., Silva, C., Moura e Sá, F., 2021. Empirical evidence on the impacts of bikesharing: a literature review. Transport Rev. 41 (3), 329–351.
Teixeira, J.F., Silva, C., e Sá, F.M., 2022. The strengths and weaknesses of bike sharing as an alternative mode during disruptive public health crisis: a qualitative analysis on the users' motivations during COVID-19. Transport Pol. 129, 24–37.
Tongwane, M., Piketh, S., Stevens, L., Ramotubei, T., 2015. Greenhouse gas emissions from road transport in South Africa and Lesotho between 2000 and 2009. Transport. Res. Transport Environ. 37, 1–13.
Vicente, M.R., 2022. ICT for healthy and active aging: the elderly as first and last movers. Telecommun. Pol. 46 (3), 102262.
Vogel, M., Hamon, R., Lozenguez, G., Merchez, L., Abry, P., Barnier, J., et al., 2014. From bicycle sharing system movements to users: a typology of Vélo'v cyclists in Lyon based on large-scale behavioural dataset. J. Transport Geogr. 41, 280–291.
Wang, Q., Yang, J., Chiu, Y.H., Lin, T.Y., 2020. The impact of digital finance on financial efficiency. Manag. Decis. Econ. 41 (7), 1225–1236.
Waserhole, A., Jost, V., 2016. Pricing in vehicle sharing systems: optimization in queuing networks with product forms. EURO Journal on Transportation and Logistics 5 (3), 293–320.
Wei, S., 2015. Internet lending in China: status quo, potential risks and regulatory options. Comput. Law Secur. Rep. 31 (6), 793–809.
Winters, M., Teschke, K., Brauer, M., Fuller, D., 2016. Bike Score®: associations between urban bikeability and cycling behavior in 24 cities. Int. J. Behav. Nutr. Phys. Activ. 13 (1), 1–10.
Xiao, K., Li, F., Dong, C., Cai, Y., Li, Y., Ye, P., Zhang, J., 2020. Unraveling effects of coal output cut policy on air pollution abatement in China using a CGE model. J. Clean. Prod. 269, 122369.
Xu, P., Chen, Y., Ye, X., 2013. Haze, air pollution, and health in China. Lancet 382 (9910), 2067.
Yang, J., Guo, B., Wang, Z., Ma, Y., 2020. Hierarchical prediction based on network-representation-learning-enhanced clustering for bike-sharing system in smart city. IEEE Internet Things J. 8 (8), 6416–6424.
Yang, Y., Jiang, L., Zhang, Z., 2021. Tourists on shared bikes: can bike-sharing boost attraction demand? Tourism Manag. 86, 104328.
Yao, Y., Jiang, X., Li, Z., 2019. Spatiotemporal characteristics of green travel: a classification study on a public bicycle system. J. Clean. Prod. 238, 117892.
Yin, Z., Gong, X., Guo, P., Wu, T., 2019. What drives entrepreneurship in digital economy? Evidence from China. Econ. Modell. 82, 66–73.
Zhang, D., Yu, C., Desai, J., Lau, H.Y.K., Srivathsan, S., 2017. A time-space network flow approach to dynamic repositioning in bicycle sharing systems. Transp. Res. Part B Methodol. 103, 188–207.
Zhang, X., Li, W., 2023. Effects of a bike sharing system and COVID-19 on low-carbon traffic modal shift and emission reduction. Transport Pol. 132, 42–64.
Zhang, Y., Mi, Z., 2018. Environmental benefits of bike sharing: a big data-based analysis. Appl. Energy 220, 296–301.
Zhao, C., Deng, M., Cao, X., 2021. Does haze pollution damage urban innovation? Empirical evidence from China. Environ. Sci. Pollut. Control Ser. 28 (13), 16334–16349.
Zhao, C., Guo, J., 2022. Are veterans happy? Long-term military service and the life satisfaction of elderly individuals in China. J. Happiness Stud. 23 (2), 477–508.
Zhao, C., Wu, Y., Guo, J., 2022. Mobile payment and Chinese rural household consumption. China Econ. Rev. 71, 101719.
Zhao, J., Deng, W., Song, Y., 2014. Ridership and effectiveness of bikesharing: the effects of urban features and system characteristics on daily use and turnover rate of public bikes in China. Transport Pol. 35, 253–264.
Zhao, N., Zhang, X., Banks, M.S., Xiong, M., 2018. Bicycle sharing in China: past, present, and future. In: Proceedings of the Southern Association for Information Systems Conference, vol. 732. Atlanta, GA, USA.
Zhao, P., Li, S., Li, P., Liu, J., Long, K., 2018. How does air pollution influence cycling behaviour? Evidence from Beijing. Transport. Res. Transport Environ. 63, 826–838.

Available online at www.sciencedirect.com

ScienceDirect

Journal of Policy Modeling 45 (2023) 320–344

www.elsevier.com/locate/jpm

ICT in education can improve students' achievements in rural China: The role of parents, educators and authorities

Chunkai Zhao [a,1], Boou Chen [b,*,2]

[a] *College of Economics and Management, South China Agricultural University, No. 483, Wushan Street, Tianhe District, Guangzhou, Guangdong Province, China*

[b] *School of Economics, Nanjing University of Finance and Economics, No. 3 Wenyuan Road, Qixia District, Nanjing, Jiangsu Province, China*

Received 20 December 2022; Received in revised form 5 February 2023; Accepted 17 February 2023
Available online 7 March 2023

Abstract

Based on a quasi-natural experiment of the Three Links Project (TLP) implemented in rural China, we explore the impact of ICT in education on students' achievements. By using the difference-in-differences (DID) method, we find that TLP implementations result in a 3.4340-point increase in students' test scores and cognitive and noncognitive ability scores increased by 0.1631 and 0.0405 points, respectively. Results of mechanism tests show that this positive effect is explained by the improvement of teachers' efforts and enthusiasm, the increase of parental involvement in education, and positive student responses. Moreover, heterogeneity analysis results suggest that TLP implementations have a greater positive effect on achievements of some disadvantaged students, such as girls, non-only children, and those from relatively low-income families. This study highlights the benefits of ICT in education and reveals the positive role of parents, educators, and authorities in promoting student achievement and human capital accumulation. Furthermore, our findings complement the research field on educational inequity and have some insights for China and other developing countries.
© 2023 The Society for Policy Modeling. Published by Elsevier Inc. All rights reserved.

Keywords: ICT in education; Students' achievements; Three Links Project; Human capital accumulation; Rural China

* Corresponding author.
E-mail address: drchen2015@163.com (B. Chen).
[1] ORCID: 0000-0003-1918-9598.
[2] ORICD: 0000-0003-1735-7535.

https://doi.org/10.1016/j.jpolmod.2023.02.007
0161-8938/© 2023 The Society for Policy Modeling. Published by Elsevier Inc. All rights reserved.

1. Introduction

As a new generation, students are familiar with Information and Communication Technology (ICT) from an early age. ICT applications have gradually become an important component of high-quality education (Li & Ranieri, 2013; Luo et al., 2020). However, due to differences in economic and technological levels, there is a large gap between developing and developed countries in ICT in education. The latest report from the Program for International Student Assessment (PISA) showed that the computer-student ratio in developed countries such as the UK and the US has reached 1.25, however, in some developing countries, such as Brazil, Morocco, Turkey, and Vietnam, this ratio is about 0.25, which is only one-fifth of developed countries (OECD, 2020).

As a developing country, China is not much better off. The PISA investigated Beijing, Shanghai, Jiangsu, and Zhejiang, which are the most developed regions in China, and the report suggests that the average computer-student ratio is above 0.5 slightly (OECD, 2020). The situation in rural areas is even worse. As represented in Table 1, in 2006, the computer-student ratios in rural primary schools, junior middle schools, and senior high schools were only 0.03, 0.05, and 0.08, respectively. In recent years, with the adoption of several initiatives including ICT in education in rural and distant locations, the situation has improved recently (Li & Ranieri, 2013), but the gap between urban and rural areas remains obvious.

Considering the disadvantages of ICT in education in rural China, the resulting digital divide may have a negative impact on student achievements (e.g., Jonassen & Kwon, 2001; Li & Ranieri, 2013; Machin et al., 2007; Subrahmanyam et al., 2001). Since educational and labor market outcomes in adulthood are largely pre-determined by the academic performances in childhood (Heckman et al., 2006), the potential influencing factors that affect student achievements have drawn a lot of interest from scholars. In addition to the digital divide mentioned above, other factors that were considered to be significant for student achievement, including family background (Avvisati et al., 2014; Zhao & Chen, 2022), school characteristics (Fan et al., 2020; Pop-Eleches & Urquiola, 2013), teacher characteristics (Chen & Zhao, 2022; Duflo et al., 2011; Gong et al., 2018), and socioeconomic conditions (Hannum & Meiyan, 2006; Malamud et al., 2019). In this paper, by regarding the Three Links Project (TLP) implemented in China as a quasi-natural experiment, we are aimed to identify the causal effect of ICT in education on student achievements in rural China.

Our study falls into the broad research field involving the impact of ICT in education on students. Some scholars found the advantages of ICT in education. Higher Internet self-efficacy, which is associated with better academic achievements and more cognitive capacities, is present among students who use the Internet at school (Bai et al., 2016; Li & Ranieri, 2013; Zhao & Xu, 2010). Moreover, ICT in education may foster student enthusiasm, boost classroom interaction,

Table 1
Comparison of computer-student ratios between urban and rural schools in China.

	Primary schools			Junior middle schools			Senior high schools		
	Total	Urban	Rural	Total	Urban	Rural	Total	Urban	Rural
2006	0.0377	0.0752	0.0311	0.0537	0.0769	0.0493	0.1058	0.1503	0.0846
2010	0.0414	0.0719	0.0345	0.0635	0.0779	0.0599	0.1093	0.1479	0.0882
2015	0.0846	0.1011	0.0769	0.1265	0.1407	0.1194	0.1587	0.1961	0.1265
2019	0.1342	0.1333	0.1348	0.1857	0.1996	0.1774	0.2400	0.2928	0.1897

and interest, as well as promote digital skills (Straker & Pollock, 2005; Pagani et al., 2016). More importantly, students' horizons can be expanded by the Internet's wealth of knowledge and information, which can help increase learning effectiveness (Jonassen & Kwon, 2001). The PISA recently showed that high-speed Internet access in schools is linked to greater results, even after completely accounting for socioeconomic background (OECD, 2020).

However, some studies considered the potential negative effects of the ICT and Internet use on students' academic preforms, such as Internet addiction, reducing social interaction, lack of active participation in extracurricular activities, and decreasing classroom attention (Livingstone, 2012; McDool et al., 2020; Subrahmanyam et al., 2001). Besides, some scholars found that even an increase in information technology or applications for teaching did not help much in improving students' performance (Angrist & Lavy, 2002; Rouse & Krueger, 2004). While there are a number of reasons for the heterogeneity of results, the findings are consistent for disadvantaged students. For example, Jackson et al. (2006) found that computer use significantly improved the reading achievements of students from low-income families in the US. Coincidentally, based on a sample of students from most remote school districts in Michigan, Hampton et al. (2021) drew a similar conclusion. In Chinese cases, Yang et al. (2013) indicated that using computers in school was good for improving academic performances of students from poor communities. Another evidence from rural China concluded that using ICT in the classroom helped pupils perform better in English (Bai et al., 2016).

Although the impact of ICT in education on student achievements has received some scholarly attention, there are few studies evaluating the effectiveness of China's policies for ICT in education (Bai et al., 2016), especially causal effects and potential mechanisms. In fact, China has started putting the ICT in education construction initiative into practice since 2011.[3] This project required that each class in rural China must have access to ICT resources and technology by the year 2015. More importantly, it was carried out in batches, which provides us with an available quasi-natural experiment to identify the causal effect of ICT in education on students' achievement. In addition to the role of the authorities, we analyze potential mechanisms at the level of parents, teachers, and students. Thus, we seek to reveal how parents, educators, and authorities play a role in ICT in education to enhance student achievement and promote human capital accumulation.

In the following respects, we have some meaningful findings, which add to the body of literature already in existence. First, by using TLP implementation as a proxy for ICT in education in rural China, we adopt a difference-in-differences (DID) method to establish a causal link between ICT in education and students' achievement. Our empirical results provide evidence that TLP implementations significantly improve students' test scores, cognitive abilities, and non-cognitive abilities. This means that our findings support the positive outcomes of ICT in education (e.g., Li & Ranieri, 2013; Zhao & Xu, 2010; Straker & Pollock, 2005).

Second, we are also interested in potential mechanisms of ICT in education for student achievement. The results of mechanism analysis suggest that TLP implementations may work in three channels to enhance students' achievement: parents, teachers, and students. Specifically, we find that TLP implementations can improve teachers' efforts and enthusiasm, increase parental involvement in education, and bring positive student responses. Thus, based on the

[3] Website for information about the TLP: http://www.moe.gov.cn/srcsite/A16/s3342/201206/t20120626_139233.html.

distinctive Chinese context, our research framework is well represented in how parents, educators, and authorities promote human capital accumulation through ICT in education.

Third, heterogeneity analysis results show that the TLP have greater positive effects on the achievement of some disadvantaged students. With the advancement of ICT in education, China's education inequality and urban-rural digital divide may be alleviated. Taken together, our study complements the role of ICT in education on the human capital accumulation of disadvantaged students (Bai et al., 2016; Hampton et al., 2021), and expands the body of knowledge on reducing educational inequality (Agasisti et al., 2021; Asadullah et al., 2020; Hannum & Meiyan, 2006; Li & Ranieri, 2013; Machin et al., 2007). Also, this study may have implications for other developing countries to formulate relevant ICT in education public policies.

2. Research framework and mechanism analysis

In this section, we construct a theoretical framework about the mechanisms by which ICT in education affects student achievements at three levels: parents, teachers, and students.

2.1. Parental involvement in education

We consider that ICT in education may promote student achievements by increasing parental involvement in education. Not surprisingly, the ICT provides more opportunities and convenience for parental involvement in their children's education. On the one hand, ICT in education has a demonstration effect. Through ICT tools, parents can better understand more scientific models of family education and improve the effectiveness of family education. On the other hand, ICT in education creates better conditions for home-school contact, allowing parents to be more aware of their children's school performance and learning dynamics (del Carmen Ramírez-Rueda et al., 2021). Students' achievements are also improved through a more efficient model of home-school co-education (Chen & Zhao, 2022; Sheridan et al., 2012).

More specifically, the relatively low penetration of private computers and the Internet in rural China has led to difficulties in accessing ICT devices in home education. As a result, students are limited to traditional and inefficient learning tools at home (Zhao & Xu, 2010). However, the implementations of a range of programs for ICT in education could well compensate for this loss (Cleveland & Krashinsky, 1998), which may further have a spillover effect on parental involvement in education (Fan et al., 2020).

Many benefits of parental involvement in education have been widely acknowledged by prior research, such as social skills (Sheridan et al., 2012), attendances (Avvisati et al., 2014), and mental health (Wang & Sheikh-Khalil, 2014). It is clear that these factors are also closely related to student achievements (Chen & Zhao, 2022; De Paola & Gioia, 2017). Taken together, ICT in education may influence parental involvement in education, which in turn contributes to student achievement.

2.2. Teachers' efforts and enthusiasm

ICT in education provides teachers with a variety of pedagogical conveniences that can further promote their efforts and stimulate enthusiasm. First, ICT in education raises expectations for teachers' information technology skills and comprehensive quality (Bai et al., 2016; Livingstone, 2012; Peeraer & Van Petegem, 2011; Wang et al., 2019); therefore, it further

motivates teachers to put more efforts to improve their teaching skills and maintain good learning status (Qaddumi et al., 2021). Moreover, the high efficiency of ICT in education is likely to endogenously motivate teachers to take the initiative to use these electronic devices to enhance teaching quality and improve student academic performance (Bai et al., 2016; Siddiq et al., 2016). More importantly, teachers in rural schools may conclude their teaching methods through action and reflection to optimize teaching and create a more efficient classroom with ICT (Luo et al., 2020; Wang et al., 2019).

Second, ICT in education helps stimulate teachers' subjectivity and wisdom, and enhances sensual enthusiasm and creativity in rural schools. Teachers' enthusiasm has traditionally been defined as a lively and motivating teaching style that includes a range of behaviors that reflect a strong interest in the subject (Collins, 1978). The new teaching resources and equipment brought by ICT can reduce teachers' workload, decrease the time spent on simple and repetitive tasks (Peeraer & Van Petegem, 2011; Zhao & Xu, 2010), which is helpful for reducing burnout and promoting enthusiasm. Also, the new teaching concepts created by ICT in education can lead to a greater sense of achievements and self-efficacy for teachers (Wang et al., 2019), especially in rural schools where the ICT devices were lacking in the past.

Some studies have confirmed that teachers' efforts and enthusiasm for teaching helps correct students' learning attitudes (Frenzel et al., 2019; Meroni et al., 2015; Keller et al., 2016) and improve students' motivation, positivism, and vitality in learning (Chen & Zhao, 2022). Thus, we believe that by stimulating the efforts and enthusiasm of teachers, ICT in education may improve student achievements in rural schools.

2.3. *Students' responses*

There is no doubt that ICT in education can affect rural school students in several ways. ICT in education motivates students to study harder (Peeraer & Van Petegem, 2011). ICT in education provides new learning resources, opens up horizons and enriches students' academic knowledge (Bai et al., 2016; Wang et al., 2019). Therefore, students need to put more effort to acquire and master more knowledge (Straker & Pollock, 2005; Siddiq et al., 2016). In addition, ICT in education brings new learning experiences and more scientific learning models for students, which helps to stimulate interest in learning (Hampton et al., 2021; Voogt et al., 2013). Moreover, ICT in education may change students' learning abilities and attitudes in rural schools. One of the benefits of ICT in education is that it can better supervise students' learning (Wang et al., 2022), both at school and at home. Specifically, teachers can better help students shape good learning attitudes and correct bad habits through the use of these ICT teaching devices (Qaddumi et al., 2021). As ICT in education is promoted in rural schools, students' learning abilities and attitudes are improved (Bai et al., 2016; Wu et al., 2019).

In addition, ICT in education gives more children in rural schools hope and raises educational expectations (Wang et al., 2019). ICT in education further reduces the digital divide between urban and rural areas, allowing students in rural schools to enjoy better teaching resources, bridging the gap with urban schools through ICT educational devices (Li & Ranieri, 2013). Last, ICT in education may enhance students' sense of belonging at school. Some studies confirmed that students reporting a greater sense of belonging at school tend to display higher academic motivation, self-esteem, and achievements (Sirin & Rogers-Sirin, 2004). ICT in education may improve students' sense of belonging at school by better learning equipment and environment, higher quality campus management, a more outstanding disciplined climate, and more enthusiastic teachers (Dotterer et al., 2007).

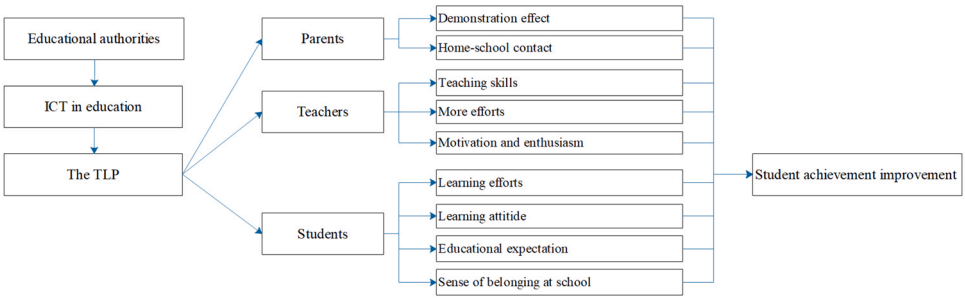


Fig. 1. Research framework.

Taken together, using the government-led TLP as a proxy for ICT in education, we construct an integrated research framework to observe how ICT in education influences the behaviors and responses of parents, teachers, and students. Moreover, we seek to analyze how parents, educators, and authorities come to play a role in student achievements and human capital accumulation through ICT in education. The research framework is shown in Fig. 1.

3. Institutional background

In recent years, the Chinese government has tried to promote some ICT reforms in education in rural schools. In 2003, the Ministry of Education, the National Development and Reform Commission, and the Ministry of Finance jointly implemented the "Project of Distance Education in Rural Schools" (PDERS), which aimed to provide some digital devices for primary and secondary schools in rural areas, such as digital satellite receiver, computer rooms, teaching CD-ROM players. By the end of 2007, this project had been basically completed. The total investment is about 11.1 billion, of which the central special funds and local funds were 5 billion and 6.1 billion, respectively, benefiting more than 100 million rural primary and secondary school students.

The TLP can be regarded as the continuation of PDERS. In 2011, the Chinese government issued the "Decade Plan for Development of ICT in education (2011–2020)," which stated the need to construct an educational information network covering all types of schools and educational institutions, further achieving broadband access to every school by 2020. Then, the Ministry of Education put forward the TLP, requiring three goals to be achieved by 2015. First, every classroom should be equipped with computers, the Internet, projectors, and other teaching digital equipment. Second, develop a wealth of quality teaching resources and deliver them to each classroom for general use in the teaching process. Third, ensure that all teachers and students in schools have their own online learning space.

During the implementation of TLP, all local governments were required to submit timelines for 2013, 2014, and 2015 to the Ministry of Education, which also regularly checked and monitored implementation processes. This top-down model ensures the achievements of the ICT goals of each pilot school (Meng & Su, 2021). More importantly, the TLP also emphasized that resources should be tilted to backward rural areas. Obviously, the TLP has become an important measure for China to improve ICT conditions in rural schools, alleviate the urban-rural digital divide, and promote educational equity (Becker, 2012; Zhao & Xu, 2010).

4. Methodology

4.1. Data

The data originates from two waves of the China Education Panel Survey (CEPS) conducted in the academic years of 2013–2014 and 2014–2015. The CEPS is a large-scale and nationally representative longitudinal survey undertaken by Renmin University of China, concentrating on junior middle school students. More crucially, the CEPS is China's first database that offers a plethora of information on schools, teachers, parents, and students, disclosing the impact of families, schools, and other socioeconomic status on student achievements (Chen & Zhao, 2022; Fan et al., 2020). We process the samples as follows. First, we keep only the tracking samples to accommodate the DID method.[4] Second, since the TLP targets rural schools, we eliminate all non-rural schools.[5] Third, considering that only two phases of data are available and 2013 is not the first year of TLP implementations. Therefore, in the DID framework, we remove samples that were already TLP pilots at the baseline wave.[6] That is, in fact, we evaluate the short-term effects of ICT in education, observing changes in student achievements one year after TLP implementations in rural schools.

4.2. Empirical model

As noted above, the TLP was issued by the Ministry of Education of China in 2011 and launched pilot construction in schools. Although the implementation time of the TLP varied in different schools, it was mainly concentrated between 2012 and 2015, which provided good conditions for the policy evaluation by adopting the DID approach. Specifically, we regard that the treatment group consists of rural schools that implemented the TLP. On the contrary, those rural schools that have never become TLP pilots as a control group, where students would not be affected. The empirical model is set up as follows:

$$ACH_{ist} = \alpha_0 + \alpha_1 TLP_{st} + \alpha_2 X_{ist} + \gamma_s + \rho_t + \delta_{ct} + \varepsilon_{ist} \tag{1}$$

Where ACH_{ist} indicates achievements of the student i at school s in the year t and TLP_{st} represents whether the school s has implemented the TLP in the year t. X_{ist} is a series of control variables which covers student, family, and school characteristics, and details are given below. γ_s and ρ_t are the school fixed effect (FE) and year FE, respectively. δ_{ct} is the year-by-county FE, which is used to capture time trends in different counties. High-dimensional FEs have been commonly seen in DID estimations to further mitigate potential endogeneity issues (e.g., Duflo et al., 2011; Zhao & Guo, 2022). ε_{ist} is a random error term. To control for serial correlation as well as heterogeneity of variables, we cluster the standard errors at the school level.

[4] They were 7th grade in the baseline wave and 8th grade in the second wave.

[5] The CEPS provides information on the location of schools with five options: (1) center of the city/town; (2) outskirts of the city/town; (3) rural-urban fringe zone of the city/town; (4) towns outside of the city/town; (5) rural areas. According to the focus area of Ministry of Education of China for TLP implementation, in the main specification, we remove the first two categories of non-rural schools by adopting a broad definition for rural schools.

[6] If we do not remove these samples, there would be a large number of observations from the treatment group in the base period, which may lead to unsatisfied parallel trends and invalid DID estimates (Sun & Abraham, 2021).

4.3. Variables

4.3.1. Students' achievements

Our explained variables are student achievements, measured by average scores, cognitive abilities, and noncognitive abilities, which are widely used in previous studies (e.g., Chen & Zhao, 2022; Duflo et al., 2011; Feng & Li, 2016; Gong et al., 2018). Specifically, *Average scores* is expressed by the average scores of three main subjects (including Chinese, Math, and English) in the most recent midterm exam (Chen & Zhao, 2022; Fan et al., 2020). *Cognitive abilities* is measured the standardized scores of the cognitive ability test[7] provided by the CEPS database (Cunha & Heckman, 2008; Fiorini, 2010). Moreover, according to existing literature (Cunha & Heckman, 2008; Gong et al., 2018; Romero et al., 2017; Zhao & Chen, 2022), *Noncognitive abilities* is a composite indicator consisting of confidence, openness, emotional stability, and agreeableness.

Adopting multiple variables to measure student achievements is motivated by two main considerations. First, three variables help to more accurately reflect the positive effect of ICT in education on students' achievements in rural schools. Second, the CEPS only provides results for one test and only includes three main subjects, and test papers were inconsistent from school to school, with differences in difficulty levels. Therefore, the explained variable *Average scores* is incomplete and contingent (Chen & Zhao, 2022). Conversely, using the cognitive ability scores would more objectively capture student achievement, as it was tested by the CEPS group by using the same questions.

4.3.2. The TLP

Although the TLP has been implemented in some schools since 2011, given that the CEPS has not announced the specific names of counties and schools, it is impossible to determine which rural schools in the sample are really affected by the policy. Fortunately, in the school leadership questionnaire, the respondents were asked whether the school has implemented the TLP since a certain year. According to the CEPS, we find that 24 rural schools in 14 counties are TLP pilots in the first wave,[8] and 6 schools in 4 counties are added in the second wave. Based on the DID method, the setting rule of our key explanatory variable is as follows. If a rural school becomes the TLP pilot in the second wave, it is set to 1, and 0 otherwise.

4.3.3. Control variables

We include a series of control variables. First, in terms of student characteristics, *Gender*, *Birth year*, *Only child*, *Hukou*, and *Self-rated health* are controlled for, because these determinants of student achievements that is routinely examined in the literature (Cunha & Heckman, 2008; Gong et al., 2018). From the perspective of family characteristics, *Self-rated wealthy*, *Basic living allowance*, *Schooling of father*, *Schooling of mother*, and *Occupation of father* are added in the control set, as household economic status and parental educational attainment are important factors that influence student achievements (Asadullah et al., 2020; Feng & Li, 2016). From the school characteristics, there are five variables, namely *Ranking of school*, *Appropriation per student*, *Teacher quality*, *Education of teachers*, and *Infrastructure of*

[7] Test questions included eleven types of three dimensions: language, graphics and space, as well as calculation and logic, and students were required to answer these questions intensively in class.

[8] As highlighted in Section 4, these observations were deleted.

Table 2
Descriptive statistics.

Variables	Full sample			Non-TLP schools	TLP schools	Diff. (T-test)
	N	Mean	SD	Mean	Mean	
Average scores	2006	70.7082	22.6670	70.2135	71.9273	-1.7138
Cognitive abilities	2006	-0.0521	0.8238	-0.1913	0.2908	-0.4821 ***
Noncognitive abilities	2006	3.0321	0.4754	3.0192	3.0638	-0.0446 *
TLP	2006	0.2886	0.4532			
Gender	2006	0.5189	0.4998	0.5123	0.5354	-0.0231
Birth year	2006	2000.32	0.8010	2000.3174	2000.3264	-0.0090
Only child	2006	0.6825	0.4656	0.7162	0.5993	0.1169 ***
Hukou	2006	0.2483	0.4321	0.2249	0.3057	-0.0808 ***
Self-rated health	2006	4.0543	0.8925	4.0224	4.1330	-0.1106 **
Self-rated wealthy	2006	2.6102	0.6629	2.5809	2.6822	-0.1013 ***
Basic living allowance	2006	0.2039	0.4030	0.2172	0.1710	0.0463 **
Schooling of father	2006	9.2423	2.4528	9.2032	9.3385	-0.1353
Schooling of mother	2006	8.2732	3.0009	8.3139	8.1727	0.1412
Occupation of father	2006	0.0763	0.2655	0.0757	0.0777	-0.0020
Ranking of school	2006	3.4227	0.8671	3.4043	3.4680	-0.0637
Appropriation per student	2006	6.5380	1.1592	6.5701	6.4588	0.1113 *
Teacher quality	2006	0.5635	0.2942	0.6052	0.4607	0.1445 ***
Education of teachers	2006	0.7024	0.2640	0.7121	0.6786	0.0335 **
Infrastructure of school	2006	5.9676	1.6591	5.7456	6.5147	-0.7691 ***

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively.

school. Since our core explanatory variable is at the school level, controlling for some school characteristics helps to further mitigate potential endogeneity (Fan et al., 2020; Zhao & Chen, 2022). The specific definitions can be seen in Appendix Table A1.

4.4. Descriptive statistics

The descriptive statistics of main variables mentioned above are shown in Table 2. Our final sample included 2006 students from rural schools, of which 575 students in the TLP pilot schools in the second wave, accounting for 28.86%. Additionally, Table 2 shows that the mean values of students' average scores, cognitive abilities, and noncognitive abilities were 70.7082, – 0.0521, and 3.0321, respectively. According to the T-test, we find that students in the TLP pilot rural schools had better achievement, particularly in cognitive abilities. More specifically, on average, students in TLP schools had higher cognitive ability scores than these in non-TLP schools by 0.4821. Thus, these statistics provide preliminary validation that TLP implementations contributed to improved students' achievements and illustrated the positive role of ICT in education.

5. Results

5.1. The effect of TLP implementations on students' achievement

Based on the quasi-natural experiment of TLP implementations, we employ the DID approach to examine the causal effect of ICT in education on students' achievement. The baseline results are shown in Table 3. It is clear that the coefficients on *TLP* keep significantly positive at

Table 3
The impact of TLP implementations on student achievement: DID estimations.

	(1)	(2)	(3)
	Average scores	Cognitive abilities	Noncognitive abilities
TLP	3.4340 ***	0.1631 ***	0.0405 ***
	(0.1545)	(0.0075)	(0.0042)
Gender	-7.5943 ***	-0.0060	-0.0070
	(1.2418)	(0.0386)	(0.0205)
Birth year	2.8441 ***	0.1572 ***	-0.0008
	(0.6477)	(0.0368)	(0.0113)
Only child	1.0352	0.0382	-0.0251
	(1.2358)	(0.0442)	(0.0225)
Hukou	-0.4342	-0.0679	-0.0467 *
	(1.0720)	(0.0665)	(0.0263)
Self-rated health	-0.3257	0.0021	0.1048 ***
	(0.5509)	(0.0143)	(0.0127)
Self-rated wealthy	-0.0981	0.0214	0.0190
	(0.9301)	(0.0362)	(0.0166)
Basic living allowance	-2.3699 *	-0.0743	-0.0181
	(1.1565)	(0.0599)	(0.0380)
Schooling of father	0.4343 **	0.0079	0.0185 **
	(0.2026)	(0.0073)	(0.0064)
Schooling of mother	-0.1259	-0.0078	0.0096 *
	(0.2020)	(0.0091)	(0.0052)
Occupation of father	-0.8071	0.0011	-0.0025
	(1.0105)	(0.0757)	(0.0452)
Ranking of school	-1.1262 **	0.5855 ***	0.0475 ***
	(0.4656)	(0.0153)	(0.0086)
Appropriation per student	-4.5725 ***	0.0651 ***	0.0433 ***
	(0.0573)	(0.0033)	(0.0019)
Teacher quality	-18.3068 ***	0.1684 ***	0.0792 ***
	(0.6772)	(0.0327)	(0.0190)
Education of teachers	-12.8934 ***	-1.1662 ***	-0.3545 ***
	(0.8419)	(0.0230)	(0.0146)
Infrastructure of school	-0.0714	0.0545 ***	0.0251 ***
	(0.1155)	(0.0033)	(0.0025)
Year FE	Yes	Yes	Yes
School FE	Yes	Yes	Yes
Year-by-county FE	Yes	Yes	Yes
Adj. R-squared	0.4614	0.2924	0.1474
N	2006	2006	2006

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. Standard errors clustered at the school level are reported in parentheses.

the 1% level in all columns, implying that TLP implementations did help to improve student achievements in Chinese rural schools. These estimates are not only statistically significant, but are equally significant in an economic sense. In particular, in column (1), TLP implementations result in a 3.4340-point increase in the average scores, corresponding to an improvement of 4.86%, relative to the mean, 70.7082. By converting to standard deviations, the estimates indicate that the TLP increases average scores by 0.069 standard deviations. Moreover, in columns (2) and (3), the results suggest that the TLP results in an increase in the students'

Table 4
Parallel trend tests.

	(1)	(2)	(3)
	Average scores	Cognitive abilities	Noncognitive abilities
TLP	5.9851 **	0.2765 ***	0.0767
	(2.4836)	(0.0660)	(0.0586)
TLP_1	2.9264	0.1301	0.0416
	(2.9045)	(0.0765)	(0.0693)
Control variables	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
School FE	Yes	Yes	Yes
Year-by-county FE	Yes	Yes	Yes
N	2006	2006	2006

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. Standard errors clustered at the school level are reported in parentheses.

cognitive and noncognitive ability scores by 0.1631 and 0.0405 points, respectively, which is equivalent to 0.0897 and 0.0386 standard deviations.

These results are consistent with some previous studies that confirmed the positive effects of ICT in education (Bai et al., 2016; Li & Ranieri, 2013; Straker & Pollock, 2005; Pagani et al., 2016), especially vulnerable children in rural and remote areas (Bai et al., 2016; Hampton et al., 2021; Yang et al., 2013). Taken together, by using the case of TLP implemented in Chinese rural schools and multiple educational outcome variables, we further prove the advantages of ICT in education.

5.2. Parallel trend tests

The main prerequisites of employing DID models is to satisfy parallel trends. That is, prior to becoming a TLP school, where both the treatment and control groups' trends for students' achievements are staying consistent. In Table 4, we discover that the coefficients on *TLP_1* are insignificant in three columns, indicating that prior to TLP implementations, there were no differences in three explained variables between the treatment and control schools. Additionally, students in the treatment group, however, demonstrate a significant improvement in their achievements after their schools being TLP pilots. These results suggest that the parallel trend is satisfied and that it is perfectly appropriate to use DID methods in this paper.

5.3. Mechanism tests

5.3.1. Parental involvement in education

We construct five variables to measure parental involvement in education (Zhao & Chen, 2022), namely *Family education expenses*, *Home-school communication*, *Parental requirements in children's learning*, *Parent tutoring*, and *Parental educational expectations*. These variables include not only specific behaviors of parents in children's education, such as parental tutoring and home-school communications, but also subjective parental demands and expectations for their children's education. These variables are defined as shown in Panel A of Table A2 in the Appendix.

The results for TLP implementations and parental involvement in education are represented in Table 5. It is clear that the coefficients on *TLP* are significantly positive in columns (1), (2), and

Table 5
Mechanism tests: parental involvement in education.

	(1)	(2)	(3)	(4)	(5)
	Family education expenses	Home-school communication	Parental requirements in children's learning	Parent tutoring	Parental educational expectations
TLP	1.3794 ***	0.0407 ***	0.0035	0.0008	0.3362 ***
	(0.0533)	(0.0074)	(0.0038)	(0.0043)	(0.0096)
Control variables	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes
Year-by-county FE	Yes	Yes	Yes	Yes	Yes
N	1560	1994	1970	1975	1995

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. Standard errors clustered at the school level are reported in parentheses.

Table 6
Mechanism tests: teachers' efforts and enthusiasm.

	(1)	(2)	(3)	(4)
	Hours of correcting homework or test papers	Class hours	Responsibility	Patience
TLP	0.0085 **	0.3354 ***	0.1201 ***	-0.0068
	(0.0034)	(0.0025)	(0.0093)	(0.0100)
Control variables	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes
Year-by-county FE	Yes	Yes	Yes	Yes
N	2006	1648	1994	1994

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. Standard errors clustered at the school level are reported in parentheses.

(5), indicating that the government-led ICT in education program has affected the thinking consciousness of families, resulting in more education expenses as well as stricter academic requirements for their children (Wang et al., 2022). In addition, the results imply that the TLP increase the frequency of home-school contact. Table A3 reports the results between these mechanism variables that are shown to be effective in Table 5 and student achievements. We find that home-school communications and parental educational expectations are significantly and positively related to student achievement. However, there is no association between family education expenditures and three explained variables, suggesting that the TLP may not enhance children's education outcomes by increasing family education expenditures. Our findings provided evidence that ICT in education made up for the disadvantages of rural households in human capital accumulation and had a positive spillover effect on parental involvement in education.

5.3.2. Teachers' efforts and enthusiasm

We measure teachers' efforts and enthusiasm in two ways. The first is teachers' effort in teaching, including *Hours of correcting homework or test papers* and *Class hours*. The second category is the patience and responsibility of teachers as rated by parents (Chen & Zhao, 2022; Cordero & Gil-Izquierdo, 2018). In Panel B of Table A2 in the Appendix, for these four variables, the precise definitions are provided. The results are reported in Table 6. In the first two columns, we find that TLP implementations allow teachers to spend more time correcting homework or test papers, and the impact on their class hours is also significant. These results indicate that ICT in education makes teachers tend to make more efforts in teaching, which is consistent with some previous studies (Luo et al., 2020; Siddiq et al., 2016). We further investigate the impact of TLP implementations on teachers' responsibility and patience. The results suggest that the coefficient on *TLP* is significantly positive in column (3), while insignificant in column (4). Although ICT in education makes teachers more responsible, it may also have a potential negative impact on their patience; after all, the new pedagogical concepts that ICT brings may require teachers to learn more and adapt gradually (Livingstone, 2012; Zhao & Xu, 2010).

In addition, to test the mechanisms completely, we further show the results for the effect of mechanism variables on student achievements in Table A4. We can see that all coefficients are positive, although some are not significant at the 10% level; however, this does not affect our conclusions that teachers' efforts and enthusiasm is an important way for ICT in education to enhance student achievements in rural China.

Table 7
Mechanism tests: students' efforts, learning attitude, expectations, and sense of belonging.

	(1)	(2)	(3)	(4)	(5)	(6)
	Weekend learning	Cram school on weekends	Learning attitude	Responsiveness	Self education expectation	Liking the school
TLP	0.1141 ***	0.0056 **	-0.0035	0.0207 ***	0.2266 ***	0.1089 ***
	(0.0032)	(0.0025)	(0.0077)	(0.0069)	(0.0078)	(0.0068)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-by-county FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2006	2006	1995	1996	1990	1989

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. Standard errors clustered at the school level are reported in parentheses.

5.3.3. Students' responses

According to CEPS questionnaires, we select several variables to measure students' responses. First, two variables are used to reflect children's efforts, namely *Weekend learning* and *Cram school on weekends*. Second, students' attitude toward learning also include *Learning attitude* and *Responsiveness*. Third, students' expectation was expressed in terms of future educational expectations. Last, sense of belonging at school is expressed in terms of how much students like the school. These variables were defined as shown in Panel C of Appendix Table A2 and regression results are represented in Table 7.

Specifically, we find that the coefficient of TLP is significantly positive in all columns except column (3), suggesting that with the advancement of ICT in education, students are working harder and spending more time on their studies. Also, the ICT improves student responsiveness for learning and does raise students' educational expectations in rural schools. These estimates are similar to the findings in previous literature (Bai et al., 2016; Wang et al., 2019). In addition, we find that ICT in education led to a significant increase of 0.1089 points in students' overall liking of the school. As we analyze earlier, ICT in education makes up for the lack of digital resources at home for rural students, giving them a stronger sense of belonging to the school (Bai et al., 2016). In Appendix Table A5, all mechanism variables were significantly and positively correlated with the three measures of student achievements (De Paola & Gioia, 2017), except for the two variables for student efforts, which were not significantly correlated with average scores. Basically, we confirm that ICT in education could contribute to student achievements in rural China by bringing positive student responses such as efforts, expectations, and sense of belonging at school.

5.4. Heterogeneity effects

In this subsection, we explore heterogeneity effects of TLP implementations on student achievements and further examine whether ICT in education could be of greater benefits to disadvantaged students.

First, we analyze the heterogeneity effect by gender. In rural China, gender discrimination and son preference within families persists (Murphy et al., 2011; Zhao & Chen, 2022), which leads to boys having access to more resources for family human capital investments. In Panel A of Table 8, although the TLP estimates are all positive and significant in columns (1)-(5), we

Table 8
Heterogeneity effects.

	(1)	(2)	(3)	(4)	(5)	
	Average scores		Cognitive abilities		Noncognitive abilities	
Panel A. Gender	Girls	Boys	Girls	Boys	Girls	Boys
TLP	4.3287 ***	1.5254 ***	0.2329 ***	0.1026 ***	0.1404 ***	-0.0336 ***
	(0.6480)	(0.4088)	(0.0261)	(0.0149)	(0.0153)	(0.0101)
N	1041	965	1041	965	1041	965
Panel B. Only child	Only child	Non-only child	Only child	Non-only child	Only child	Non-only child
TLP	-1.5610	3.9983 ***	0.1462 ***	0.3541 ***	0.0252 ***	0.2318 ***
	(2.3148)	(0.1516)	(0.0051)	(0.0930)	(0.0029)	(0.0551)
N	498	1508	498	1508	498	1508
Panel C. Household income	Relatively low	Relatively high	Relatively low	Relatively high	Relatively low	Relatively high
TLP	4.2034 ***	-4.8122 ***	0.1646 ***	0.0077	0.0414 ***	0.0538 ***
	(0.7367)	(0.5693)	(0.0268)	(0.0176)	(0.0117)	(0.0088)
N	695	1311	695	1311	695	1311

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. All regressions include baseline control variables, year FE, school FE, and year-by-county FE. Standard errors clustered at the school level are reported in parentheses.

find that positive effects of the TLP on girls' average scores and cognitive abilities is much greater. The ICT increases students' access to the Internet in education, which may increase the risk of Internet addiction in boys with relatively poor self-control (Livingstone, 2012; McDool et al., 2020). Therefore, these results suggest that ICT in education is more conducive to enhancing achievements of girls, a vulnerable group in rural China.

Second, we explore the only-child heterogeneity effect. According to the quantitative-qualitative trade-off theory in household human capital investment, as the number of children in the family increases, the dilution effect of family education resources becomes more pronounced, with fewer resources allocated to each child (Becker & Lewis, 1973). As shown in Panel B of Table 8, the effect of TLP implementations on non-only children is more pronounced, both in terms of average scores and cognitive and noncognitive abilities. These findings indicate that although non-only children may get fewer educational resources from their families, ICT in education may effectively compensate students with siblings in rural China.

Last, we consider the heterogeneity effect of household economic status. We divide the sample into two groups based on self-rated wealth.[9] As presented in Panel C of Table 8, we find that in terms of noncognitive abilities, the impact of TLP implementations is similar for two groups; however, for average scores and cognitive abilities, the TLP has a significantly positive effect on students from relatively low-income families. Also, these results imply a compensatory effect of TLP implementations on the human capital of relatively disadvantaged students, reflecting the need for universal access to ICT in education in rural schools.

5.5. Robustness checks

We further confirm the reliability of the baseline results by some robustness tests in Appendix Fig. A1 and Tables A6–A8. First, in placebo tests, we firstly randomly select TLP pilots from all rural schools as the treatment group. Then, 2013 and 2014, one year is randomly selected as the period being the TLP pilots. By repeating 500 times in Fig. A1. Clearly, the distribution of estimates from random assignments is centered around zero, suggesting that randomly constructed TLP school pilots have no effect on student achievements.

Second, we redefine rural schools in Table A6. In columns (1)-(3), we remove schools in rural-urban fringe zone of the city/town; in columns (4)-(6), we retain only students who attended schools located in rural areas. Results show that these estimates are quite similar to the baseline results.

Third, we consider two interfering factors: school construction activities and reclassification. If there were other policy changes during the same period that the TLP was implemented, the estimates capture the effects of other confounding factors that bias the empirical results (Hallam et al., 2004; Zhao & Wu, & Guo, 2022; Gong et al., 2018). we retain students in schools without major construction activities[10] in columns (1)-(3) of Appendix Table A7. Moreover, we keep a sample of random assignment in columns (4)-(6). Although about half of samples are removed, coefficients of *TLP* remain robust. These results tell us that school construction activities and reclassification do not have a shock on the effect of TLP implementations.

[9] Specifically, "very poor" and "somewhat poor" were regarded as relatively low income, while "moderate", "somewhat rich" and "very rich" are defined as relatively high income.

[10] These major construction activities refer to relocation, expansions, and new constructions.

Lastly, we conduct robustness checks by adding other control variables and the results are reported in Appendix Table A8. Specifically, we control for class FE in Panel A; add some teacher characteristics variables in Panel B[11]; birth year FE and self-rated health FE are included in Panel C. Clearly, coefficients on *TLP* are significantly positive in all columns and Panels, indicating that the conclusion that ICT in education enhances student achievements remained unchanged.

6. Conclusions and policy implications

Although some studies explored the impact of Internet use or ICT in education on student human capital (Bai et al., 2016; Hampton et al., 2021; Yang et al., 2013), causal effects and potential mechanisms are of little concern. Fortunately, the TLP implemented in batches provided us with a good quasi-natural experiment to help identify the causal relationship between ICT in education and students' achievement in rural China.

The findings can be summarized as follows. First, by using the DID methods, we find that TLP implementations are conducive to improving student achievement, indicating that ICT in education could promote student human capital accumulation in rural China. More specifically, TLP implementations increase average test scores by 3.434 points, and the cognitive and non-cognitive ability scores by 0.1631 and 0.0405 points, respectively. Second, we explore several mechanisms in three aspects: teachers, parents, and students, finding that the positive effect of TLP implementations can be explained by increasing teachers' efforts and enthusiasm, improving parental involvement in education, and positive student responses. Third, results of heterogeneous effects show that TLP implementations had a greater positive impact on the achievements of vulnerable students, including girls, non-only children, and those from relatively low-income families.

This study has made a few contributions to the literature. Undoubtedly, our findings support the benefits of ICT in education, which are similar to previous conclusions that emphasized that more ICT should be applied to schools instead of home (Fairlie & Kalil, 2017; Li & Ranieri, 2013; Malamud et al., 2019). Besides, our results based on rural China show that ICT in education has a greater effect on disadvantaged students, suggesting the potential of ICT to mitigate educational inequality and promote human capital accumulation (Bai et al., 2016; Hampton et al., 2021). More importantly, we construct an integrated framework that includes parents, educators, authorities, and students to provide some insight into the subsequent analysis of some key issues, such as the role of ICT in education and educational inequality in China and other developing countries (Becker, 2012; Hannum & Meiyan, 2006; Machin et al., 2007).

Our study has some policy implications. For educational authorities, investment in ICT in education should continue to be increased to improve the computer-student ratio and to further narrow the digital divide in rural China, especially those in less developed and remote areas (Bai et al., 2016). Our China-based study suggests that strengthening educational infrastructure is effective in improving educational inequality and promoting human capital development in backward areas. Thus, this paper may be supportive for those countries to increase the investment in ICT in education to make the basic education more efficient and improve human capital (Asadullah et al., 2020), which provides a new direction for related education policies.

[11] Teacher characteristics variables include gender, age, education, years of teaching experience, and title.

101

In addition, in terms of educators, based on ICT in education, there is a need to continuously improve teaching methods and enhance teaching enthusiasm to improve student achievement, as demonstrated by our mechanism tests. Thus, it is necessary to improve teachers' technology skills and gradually incorporate more efficient ICT tools in their classes (Meroni et al., 2015). After adapting to ICT teaching and learning, teachers should continue to innovate ICT usages, such as the online interaction and home-school connection, to enhance teaching and learning, especially in the post-COVID-19 era. These measures motivate teachers to actively use these electronic devices to improve students' achievement (Bai et al., 2016; Siddiq et al., 2016).

Last, from a parental perspective, in addition to actively using ICT in education opportunities to strengthen connections with teachers and other parents in the class, it is also important to actively guide and monitor children's use of electronic devices to avoid the potential negative effects of the Internet (Fairlie & Kalil, 2017; Livingstone, 2012). Further, learning digital skills is better suited for teachers and schools to increase the involvement of their children's education through ICT in education.

Acknowledgments

We appreciate the National Survey Research Center at Renmin University of China for providing us with the China Education Panel Survey (CEPS) data. The authors acknowledge the support from the Fundamental Research Funds for the Central Universities (Nos. QCDC-2020-10; QCDC-2020-21), the National Social Science Fund of China (No. 21AJY009), the Social Science Fund of Hunan Province (No. 22ZDA023), the Tsinghua Rural Studies PhD Scholarship (No. 202110), and Jiangsu Education Science Planning Project (No. C/2022/01/45).

Appendix A

See Fig. A1 and Tables A1–A8.

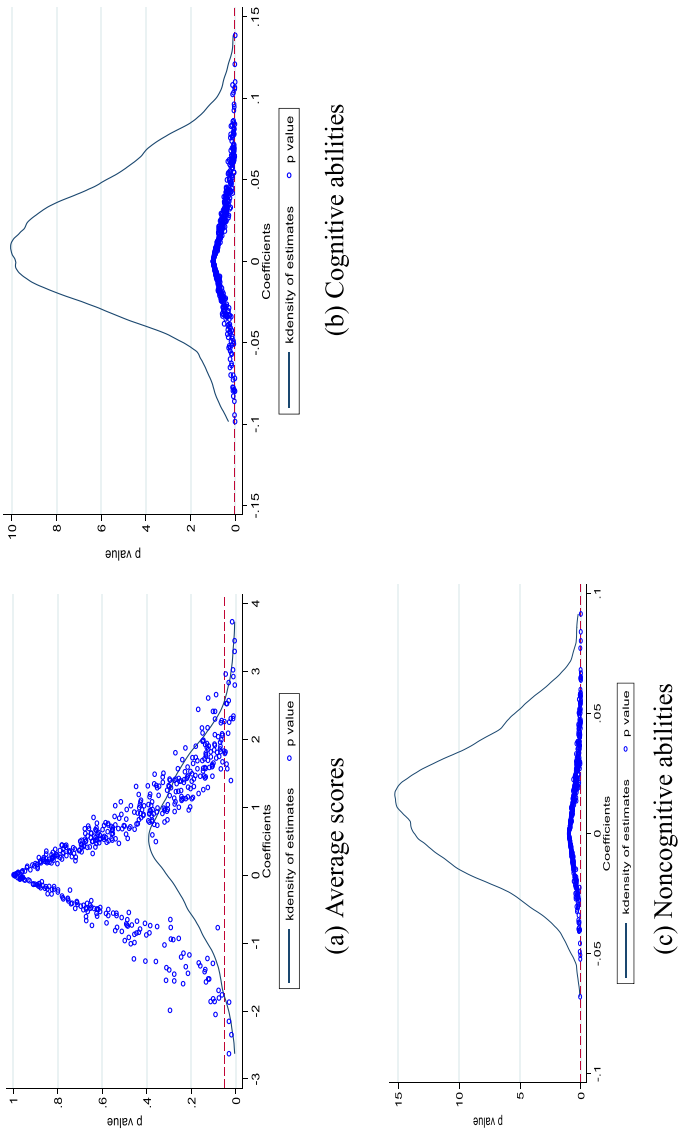


Fig. A1. Placebo tests.

Table A1
Definitions of main variables.

Variables	Definitions
Average scores	Average scores of three subjects, Chinese, Math, and English, on the most recent midterm
Cognitive abilities	Standardized scores of students' cognitive abilities based on the 3PL model
Noncognitive abilities	A composite indicator for average scores across dimensions, including confidence, openness, emotional stability, and agreeableness. Each dimension is assigned an integer from 1 to 4, with larger values indicating higher noncognitive ability scores
TLP	= 1 if the school has implemented TLP in 2014; 0 otherwise
Gender	Girl = 1; boy = 0
Birth year	Year
Only child	Yes = 1; no = 0
Hukou	Urban = 1; rural = 0
Self-rated health	Very poor = 1; not very good = 2; moderate = 3; good = 4; very good = 5
Self-rated wealthy	Very poor = 1; somewhat poor = 2; moderate = 3; somewhat rich = 4; very rich = 5
Basic living allowance	= 1 if the family receives basic living allowance; 0 otherwise
Schooling of father	Unschooled = 0; primary school = 6; junior middle school = 9; (vocationally) high school= 12 junior college = 15; bachelor's degree = 16; master's degree = 19
Schooling of mother	Unschooled = 0; primary school = 6; junior middle school = 9; (vocationally) high school= 12 junior college = 15; bachelor's degree = 16
Occupation of father	= 1 if the occupation is civil servant, manager, technician, and professional (e.g., lawyer and doctors); 0 otherwise (e.g., farmers)
Ranking of school	Near the bottom = 1; below average = 2; average = 3; above average = 4; among the best = 5
Appropriation per student	The fiscal appropriation per student this year
Teacher quality	Percentage of teachers with first-class or higher titles in the school
Education of teachers	Proportion of teachers with a bachelor's degree or higher
Infrastructure of school	Sum scores across dimensions: playground with circular track, laboratory, computer classroom, library, music room, student activity room, student dining room, and gymnasium.

Table A2
Definitions of mechanism variables.

Variables	Definitions
Panel A. Parental education involvement	
Family education expenses	Total household education expenditure every semester (1000 yuan)
Home-school communication	Times of parents contacting the teacher this semester. Never = 1; once = 2; two to four times = 3; five times or more = 4
Parental requirements in children's learning	Average scores across dimension: homework and examination, behavior at school, attendances at school every day.
Parent tutoring	Did the family members help with children's homework last week? Need but no one helped = 1; yes, one or two days = 2; yes, three or four days = 3; yes, almost every day = 4.
Parental educational expectations	1 = junior middle school; 2 = (vocationally) high school; 3 = college/bachelor's degree; 4 = master's degree; 5 = doctoral degree
Panel B. Teachers' efforts and enthusiasm	
Hours of correcting homework or test papers	Hours for homeroom teachers to correct homework or test papers last week
Class hours	Total class hours of homeroom teachers last week
Responsibility	Parents' comments. Not responsible at all = 1; not so responsible = 2; moderate = 3;somewhat responsible = 4; very responsible = 5

(*continued on next page*)

Table A2 (*continued*)

Variables	Definitions
Patience	Parents' comments. Not patient at all = 1; not so patient = 2; moderate = 3;somewhat patient = 4; very patient = 5
Panel C. Students' response	
Weekend learning	= 1 if the child has more than two hours of study time on the weekend; 0 otherwise
Cram school on weekends	= 1 if the child attends cram school on weekends; 0 otherwise
Learning attitude	Parents' comments. Not serious at all = 1; somewhat not serious = 2; moderate = 3;somewhat serious = 4; very serious = 5
Responsiveness	Parents' comments for quick responses of their children. Not fit at all = 1; somewhat not fit = 2; somewhat fit = 3; exactly fit = 4
Self education expectation	1 = junior middle school; 2 = (vocationally) high school; 3 = college/bachelor's degree; 4 = master's degree; 5 = doctoral degree
Liking the school	Strongly dislike = 1; somewhat dislike = 2; somewhat like = 3; like very much = 4

Table A3
Parental education involvement and student achievement.

	(1)	(2)	(3)
	Average scores	Cognitive abilities	Noncognitive abilities
Family education expenses	-0.0862	0.0027	-0.0020
	(0.0668)	(0.0028)	(0.0022)
Home-school communication	0.6067	0.0667 **	0.0422 ***
	(0.4498)	(0.0268)	(0.0108)
Parental educational expectations	4.3396 ***	0.1002 ***	0.0476 ***
	(0.7382)	(0.0121)	(0.0075)

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. All regressions include baseline control variables, year FE, school FE, and year-by-county FE. Standard errors clustered at the school level are reported in parentheses.

Table A4
Teachers' efforts and enthusiasm and student achievement.

	(1)	(2)	(3)
	Average scores	Cognitive abilities	Noncognitive abilities
Hours of correcting homework or test papers	3.4885	0.2999 **	0.0527
	(4.9152)	(0.1091)	(0.0328)
Class hours	17.3497 *	0.2922	0.2144 ***
	(8.8567)	(0.2457)	(0.0476)
Responsibility	3.4141 ***	0.1204 ***	0.0973 ***
	(0.8003)	(0.0253)	(0.0170)

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. All regressions include baseline control variables, year FE, school FE, and year-by-county FE. Standard errors clustered at the school level are reported in parentheses.

Table A5
Students' responses and achievement.

	(1)	(2)	(3)
	Average scores	Cognitive abilities	Noncognitive abilities
Weekend learning	0.5290	0.3379 ***	0.0406 **
	(1.3312)	(0.0773)	(0.0171)
Cram school on weekends	-2.8488	0.2468 ***	0.0805 **
	(2.4857)	(0.0749)	(0.0321)
Responsiveness	2.4515 ***	0.0937 ***	0.2694 ***
	(0.4769)	(0.0228)	(0.0124)
Self education expectation	7.7249 ***	0.1786 ***	0.0935 ***
	(1.0024)	(0.0242)	(0.0143)
Liking the school	2.3154 ***	0.0406 *	0.1930 ***
	(0.3782)	(0.0213)	(0.0157)

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. All regressions include baseline control variables, year FE, school FE, and year-by-county FE. Standard errors clustered at the school level are reported in parentheses.

Table A6
Robustness checks: redefining rural schools.

	(1)	(2)	(3)	(4)	(5)	(6)
	Average scores	Cognitive abilities	Noncognitive abilities	Average scores	Cognitive abilities	Noncognitive abilities
TLP	3.2810 ***	0.1569 ***	0.0399 ***	1.6851 *	0.6054 ***	0.1571 ***
	(0.1533)	(0.0079)	(0.0045)	(0.9623)	(0.0227)	(0.0224)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-by-county FE	Yes	Yes	Yes	Yes	Yes	Yes
N	1719	1719	1719	1082	1082	1082

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. Standard errors clustered at the school level are reported in parentheses.

Table A7
Robustness checks: interfering factors.

	(1)	(2)	(3)	(4)	(5)	(6)
	Construction activities			Reclassification		
	Average scores	Cognitive abilities	Noncognitive abilities	Average scores	Cognitive abilities	Noncognitive abilities
TLP	4.7782 ***	0.6487 ***	0.2334 ***	5.9003 ***	0.3118 ***	0.0786 ***
	(0.7591)	(0.0370)	(0.0175)	(0.3518)	(0.0153)	(0.0100)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-by-county FE	Yes	Yes	Yes	Yes	Yes	Yes
N	1116	1116	1116	1369	1369	1369

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. Standard errors clustered at the school level are reported in parentheses.

Table A8
Robustness checks: adding other variable variables.

	(1)	(2)	(3)
	Average scores	Cognitive abilities	Noncognitive abilities
Panel A. Class FE			
TLP	3.5661 ***	0.1604 ***	0.0432 ***
	(0.1559)	(0.0079)	(0.0048)
Panel B. Homeroom teacher characteristics			
TLP	5.4907 ***	0.1646 ***	0.1415 ***
	(1.1957)	(0.0460)	(0.0115)
Panel C. Birth year FE and self-rated health FE			
TLP	3.7161 ***	0.1681 ***	0.0454 ***
	(0.2042)	(0.0087)	(0.0049)

Note: Significant levels of 1%, 5%, and 10% are indicated by ***, **, and *, respectively. All regressions include baseline control variables, year FE, school FE, and year-by-county FE. Standard errors clustered at the school level are reported in parentheses.

References

Agasisti, T., Longobardi, S., Prete, V., & Russo, F. (2021). The relevance of educational poverty in Europe: Determinants and remedies. *Journal of Policy Modeling, 43*(3), 692–709.

Angrist, J., & Lavy, V. (2002). New evidence on classroom computers and pupil learning. *Economic Journal, 112*(482), 735–765.

Asadullah, M. N., Perera, L. D. H., & Xiao, S. (2020). Vietnam's extraordinary performance in the PISA assessment: A cultural explanation of an education paradox. *Journal of Policy Modeling, 42*(5), 913–932.

Avvisati, F., Gurgand, M., Guyon, N., & Maurin, E. (2014). Getting parents involved: A field experiment in deprived schools. *Review of Economic Studies, 81*(1), 57–83.

Bai, Y., Mo, D., Zhang, L., Boswell, M., & Rozelle, S. (2016). The impact of integrating ICT with teaching: Evidence from a randomized controlled trial in rural schools in China. *Computers & Education, 96*, 1–14.

Becker, G. S. (2012). Growing human capital investment in China compared to falling investment in the United States. *Journal of Policy Modeling, 34*(4), 517–524.

Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy, 81*(2), S279–S288.

Chen, B., & Zhao, C. (2022). More is less: Homeroom teachers' administrative duties and students' achievements in China. *Teaching and Teaching Education, 119*, Article 103857.

Collins, M. L. (1978). Effects of enthusiasm training on preservice elementary teachers. *Journal of Teacher Education, 29*(1), 53–57.

Cordero, J. M., & Gil-Izquierdo, M. (2018). The effect of teaching strategies on student achievement: An analysis using TALIS-PISA-link. *Journal of Policy Modeling, 40*(6), 1313–1331.

Cunha, F., & Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and non-cognitive skill formation. *Journal of Human Resources, 43*(4), 738–782.

De Paola, M., & Gioia, F. (2017). Impatience and academic performance. Less effort and less ambitious goals. *Journal of Policy Modeling, 39*(3), 443–460.

del Carmen Ramírez-Rueda, M., Cózar-Gutiérrez, R., Colmenero, M. J. R., & González-Calero, J. A. (2021). Towards a coordinated vision of ICT in education: A comparative analysis of Preschool and Primary Education teachers' and parents' perceptions. *Teaching and Teacher Education, 100*, Article 103300.

Dotterer, A. M., McHale, S. M., & Crouter, A. C. (2007). Implications of out-of-school activities for school engagement in African American adolescents. *Journal of Youth and Adolescence, 36*, 391–401.

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review, 101*(5), 1739–1774.

Fairlie, R. W., & Kalil, A. (2017). The effects of computers on children's social development and school participation: Evidence from a randomized control experiment. *Economics of Education Review, 57*, 10–19.
Fan, Z., Gao, Y., & Liu, C. (2020). Nutritional intervention, health and education: A study based on the National Nutrition Improvement Program. *Finance and Trade Economics, 41*(6), 21–35.
Feng, H., & Li, J. (2016). Head teachers, peer effects, and student achievement. *China Economic Review, 41*, 268–283.
Fiorini, M. (2010). The effect of home computer use on children's cognitive and non-cognitive skills. *Economics of Education Review, 29*(1), 55–72.
Frenzel, A. C., Taxer, J. L., Schwab, C., & Kuhbandner, C. (2019). Independent and joint effects of teacher enthusiasm and motivation on student motivation and experiences: A field experiment. *Motivation and Emotion, 43*(2), 255–265.
Gong, J., Lu, Y., & Song, H. (2018). The effect of teacher gender on students' academic and noncognitive outcomes. *Journal of Labor Economics, 36*(3), 743–778.
Hallam, S., Ireson, J., & Davies, J. (2004). Primary pupils' experiences of different types of grouping in school. *British Educational Research Journal, 30*(4), 515–533.
Hampton, K. N., Robertson, C. T., Fernandez, L., Shin, I., & Bauer, J. M. (2021). How variation in internet access, digital skills, and media use are related to rural student outcomes: GPA, SAT, and educational aspirations. *Telematics and Informatics, 63*, Article 101666.
Hannum, E., & Meiyan, W. (2006). Geography and educational inequality in China. *China Economic Review, 17*(3), 253–265.
Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*(3), 411–482.
Jackson, L. A., Von Eye, A., Biocca, F. A., Barbatsis, G., Zhao, Y., & Fitzgerald, H. E. (2006). Does home internet use influence the academic performance of low-income children? *Developmental Psychology, 42*(3), 429.
Jonassen, D. H., & Kwon, H. (2001). Communication patterns in computer mediated versus face-to-face group problem solving. *Educational Technology Research and Development, 49*(1), 35–51.
Keller, M. M., Hoy, A. W., Goetz, T., & Frenzel, A. C. (2016). Teacher enthusiasm: Reviewing and redefining a complex construct. *Educational Psychology Review, 28*(4), 743–769.
Li, Y., & Ranieri, M. (2013). Educational and social correlates of the digital divide for rural and urban children: A study on primary school students in a provincial city of China. *Computers & Education, 60*(1), 197–209.
Livingstone, S. (2012). Critical reflections on the benefits of ICT in education. *Oxford Review of Education, 38*(1), 9–24.
Luo, W., Berson, I. R., Berson, M. J., & Li, H. (2020). Are early childhood teachers ready for digital transformation of instruction in Mainland China? A systematic literature review. *Children and Youth Services Review, 120*, Article 105718.
Machin, S., McNally, S., & Silva, O. (2007). New technology in schools: Is there a payoff? *Economic Journal, 117*(522), 1145–1167.
Malamud, O., Cueto, S., Cristia, J., & Beuermann, D. W. (2019). Do children benefit from internet access? Experimental evidence from Peru. *Journal of Development Economics, 138*, 41–56.
McDool, E., Powell, P., Roberts, J., & Taylor, K. (2020). The internet and children's psychological wellbeing. *Journal of Health Economics, 69*, Article 102274.
Meng, T., & Su, Z. (2021). When top-down meets bottom-up: Local officials and selective responsiveness within fiscal policymaking in China. *World Development, 142*, Article 105443.
Meroni, E. C., Vera-Toscano, E., & Costa, P. (2015). Can low skill teachers make good students? Empirical evidence from PIAAC and PISA. *Journal of Policy Modeling, 37*(2), 308–323.
Murphy, R., Tao, R., & Lu, X. (2011). Son preference in rural China: Patrilineal families and socioeconomic change. *Population and Development Review, 37*(4), 665–690.
OECD (2020). *PISA 2018 results (volume V): Effective policies, successful schools.* Organisation for Economic Co-operation and Development OECD.
Pagani, L., Argentin, G., Gui, M., & Stanca, L. (2016). The impact of digital skills on educational outcomes: Evidence from performance tests. *Educational Studies, 42*(2), 137–162.
Peeraer, J., & Van Petegem, P. (2011). ICT in teacher education in an emerging developing country: Vietnam's baseline situation at the start of 'The Year of ICT'. *Computers & Education, 56*(4), 974–982.
Pop-Eleches, C., & Urquiola, M. (2013). Going to a better school: Effects and behavioral responses. *American Economic Review, 103*(4), 1289–1324.
Qaddumi, H., Bartram, B., & Qashmar, A. L. (2021). Evaluating the impact of ICT on teaching and learning: A study of Palestinian students' and teachers' perceptions. *Education and Information Technologies, 26*(2), 1865–1876.

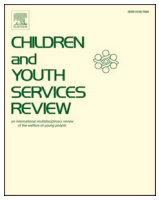
Romero, L. M., Huertas, I. P. M., & Jiménez, M. D. M. S. (2017). Wage effects of cognitive skills and educational mismatch in Europe. *Journal of Policy Modeling, 39*(5), 909–927.
Rouse, C. E., & Krueger, A. B. (2004). Putting computerized instruction to the test: A randomized evaluation of a "scientifically based" reading program. *Economics of Education Review, 23*(4), 323–338.
Sheridan, S. M., Bovaird, J. A., Glover, T. A., Andrew Garbacz, S., Witte, A., & Kwon, K. (2012). A randomized trial examining the effects of conjoint behavioral consultation and the mediating role of the parent-teacher relationship. *School Psychology Review, 41*(1), 23–46.
Siddiq, F., Scherer, R., & Tondeur, J. (2016). Teachers' emphasis on developing students' digital information and communication skills (TEDDICS): A new construct in 21st century education. *Computers & Education, 92*, 1–14.
Sirin, S. R., & Rogers-Sirin, L. (2004). Exploring school engagement of middle-class African American students. *Youth & Society, 35*(3), 323–340.
Straker, L., & Pollock, C. (2005). Optimizing the interaction of children with information and communication technologies. *Ergonomics, 48*(5), 506–521.
Subrahmanyam, K., Greenfield, P., Kraut, R., & Gross, E. (2001). The impact of computer use on children's and students' development. *Journal of Applied Developmental Psychology, 22*(1), 7–30.
Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics, 225*(2), 175–199.
Voogt, J., Knezek, G., Cox, M., Knezek, D., & ten Brummelhuis, A. (2013). Under which conditions does ICT have a positive effect on teaching and learning? A call to action. *Journal of Computer Assisted Learning, 29*(1), 4–14.
Wang, J., Tigelaar, D. E., & Admiraal, W. (2019). Connecting rural schools to quality education: Rural teachers' use of digital educational resources. *Computers in Human Behavior, 101*, 68–76.
Wang, J., Tigelaar, D. E., & Admiraal, W. (2022). From policy to practice: Integrating ICT in Chinese rural schools. *Technology, Pedagogy and Education, 31*(4), 509–524.
Wang, M. T., & Sheikh-Khalil, S. (2014). Does parental involvement matter for student achievement and mental health in high school? *Child Development, 85*(2), 610–625.
Wu, D., Li, C. C., Zhou, W. T., Tsai, C. C., & Lu, C. (2019). Relationship between ICT supporting conditions and ICT application in Chinese urban and rural basic education. *Asia Pacific Education Review, 20*(1), 147–157.
Yang, Y., Zhang, L., Zeng, J., Pang, X., Lai, F., & Rozelle, S. (2013). Computers and the academic performance of elementary school-aged girls in China's poor communities. *Computers & Education, 60*(1), 335–346.
Zhao, C., & Chen, B. (2022). Parental migration and non-cognitive abilities of left-behind children in rural China: Causal effects by an instrumental variable approach. *Child Abuse & Neglect, 123*, Article 105389.
Zhao, C., & Guo, J. (2022). Are veterans happy? Long-term military service and the life satisfaction of elderly individuals in China. *Journal of Happiness Studies, 23*(2), 477–508.
Zhao, C., Wu, Y., & Guo, J. (2022). Mobile payment and Chinese rural household consumption. *China Economic Review, 71*, Article 101719.
Zhao, J., & Xu, F. (2010). The state of ICT education in China: A literature review. *Frontiers of Education in China, 5*(1), 50–73.

Children and Youth Services Review 159 (2024) 107519

Contents lists available at ScienceDirect

Children and Youth Services Review

journal homepage: www.elsevier.com/locate/childyouth



School nutritious feeding and cognitive abilities of students in poverty: Evidence from the nutrition improvement program in China

Chunkai Zhao [a,b], Boou Chen [c,*], Zhiyong Song [d]

[a] College of Economics and Management, South China Agricultural University, No.483, Wushan Street, Tianhe District, Guangzhou, Guangdong Province, China
[b] Rural Revitalization Laboratory, South China Agricultural University, No.483, Wushan Street, Tianhe District, Guangzhou, Guangdong Province, China
[c] School of Economics, Nanjing University of Finance and Economics, No.3 Wenyuan Road, Qixia District, Nanjing, Jiangsu Province, China
[d] School of Urban and Regional Science, Shanghai University of Finance and Economics, No.777 Guoding Road, Yangpu District, Shanghai, China

ARTICLE INFO

Keywords:
School nutritious feeding
Cognitive abilities (CAs)
Nutrition Improvement Program (NIP)
Human capital
Students in poverty

ABSTRACT

While considerable attention has been paid to the benefits of school nutritious feeding, little literature focused on the effect on human capital accumulation for students in poverty. By drawing on the quasi-natural experiment of the Nutrition Improvement Program (NIP) and adopting the Difference-in-Differences (DID) method, we sought to determine the causal relationship between school nutritious feeding and the cognitive ability (CA) development among poor students in China. Specifically, our estimates confirmed that NIP implementations significantly improved the CAs of students in poverty, which is supported by several robustness checks. Moreover, the potential mechanisms of this positive effect included improved health status, enhanced sense of belonging at school, and increased positive expectations for the future. Our study further enriches the literature on the positive effects of school nutritious feeding and provides insights for developing countries to customize nutritional poverty alleviation programs to promote human capital development.

1. Introduction

School nutritious feeding programs have been commonly implemented in many countries around the world, with the main goal of promoting students' health and human capital accumulation by improving their nutritional intake (Anderson et al., 2018; Gundersen et al., 2012; Jomaa et al., 2011; Lundborg et al., 2022; Mostert, 2021; Zheng et al., 2023). A global survey showed that at least 388 million students in pre-primary, primary, and secondary schools received school nutritious feeding in 161 countries, such as India, Brazil, Sweden, and the United States (World Food Programme, 2020). China was no exception. To improve the nutritional status of students in the compulsory education stage,[1] especially those in poor areas and from poor families, the Chinese government issued a guidance document "Plan for the Implementation of the Nutrition Improvement Program for Rural Compulsory Education Students" in 2011, which marks the official implementation of the Nutrition Improvement Program (NIP), providing free nutritious meals to all students in pilot schools. By May 2020, 1762 counties in 29 provincial administrative regions had implemented the NIP, benefiting over 40.6 million students. From 2011 to 2019, the central financial authorities arranged for a cumulative total of 147.2 billion yuan[2] in subsidies for nutritious meals.[3] Moreover, according to the evaluation by the Chinese Center for Disease Control and Prevention in 2021, in NIP pilots, the growth retardation and anemia rates for students aged 6–15 years decreased by 5.7 and 4.7 percentage points, respectively.[4]

Given the widespread implementation of school nutritious feeding programs, scholars were inevitably interested in evaluating the effectiveness of these related programs or policies. Some scholars confirmed the positive effects of school nutritious feeding programs on students, such as physical and mental health (Afridi, 2010; Akin et al., 1983; Millimet et al., 2010), academic performance (Anderson et al., 2018; Imberman & Kugler, 2014; Luo et al., 2012), cognitive and non-cognitive abilities (Aurino et al., 2023; Chakraborty & Jayaraman,

* Corresponding author.
E-mail address: drchen2015@163.com (B. Chen).

[1] Compulsory education includes primary and junior high schools.
[2] This equates to approximately 2.26 billion dollars.
[3] Data source: https://www.moe.gov.cn/jyb_xwfb/s5147/202005/t20200525_458578.html.
[4] Data source: https://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/202210/t 20221014_66 9469.html.

https://doi.org/10.1016/j.childyouth.2024.107519
Received 18 July 2023; Received in revised form 21 January 2024; Accepted 2 March 2024
Available online 4 March 2024
0190-7409/© 2024 Elsevier Ltd. All rights reserved.

2019; Zheng et al., 2023), school integration and social relationships (Wang et al., 2015). Furthermore, a small body of literature addressed the long-term impacts of school nutritious feeding (Bütikofer et al., 2018; Fang & Zhu, 2022; Hinrichs, 2010; Lundborg et al., 2022). Although most of studies emphasized potential positive effects, conclusions are not uniform. Some research found that the benefits of these programs were limited (Hinrichs, 2010; McEwan, 2013; Millimet et al., 2010), due to the intensity, goals, and approaches of program implementations (Imberman & Kugler, 2014).

Based on the Chinese setting, scholars also paid close attention to the impact of school nutritious feeding on student performance. A few studies used randomized control trials to assess the impact of NIP implementations on student health and educational achievements based on regional survey data (Luo et al., 2012; Wang et al., 2015). In recent years, using the nationally representative data, the benefits of NIP implementations on student health, non-cognitive abilities, and academic achievements were also recognized by researchers (Fang & Zhu, 2022; Ren et al., 2023; Wang et al., 2022; Wang & Cheng, 2022; Zheng et al., 2023). However, these empirical studies were more concerned about some non-intellectual elements, but few highlighted the cognitive abilities (CAs). More importantly, previous research overlooked the special group of students in poverty, a key target for the NIP as an important component of China's nutrition poverty alleviation plans.

Unlike non-cognitive abilities which are considered as a collection of individual personality traits such as self-confidence, emotional stability responsibility, and social skills (Chen & Zhao, 2023; Zhao & Chen, 2022), the CAs refer to the human brain's ability to process, store, and extract information, such as intelligence, observation, memory, imagination (Dohmen et al., 2018; Heckman et al., 2006). As a core element of human capital, the CA is an important foundation for personal success and wealth acquisition (Furnham et al., 2003). For students, the CAs directly affect their current academic performances and future achievements (Frisvold, 2015; Heckman et al., 2006). Moreover, prior literature confirmed that household economic status is closely related to children's cognitive development (Li et al., 2021; Liu et al., 2020), implying that the CAs of students in poverty may be relatively backward (Brooks-Gunn & Duncan, 1997; Joe et al., 2016; Li et al., 2021). Therefore, actively enhancing children's CAs plays a crucial role in cutting the intergenerational transmission of poverty.

In view of the importance attached by all walks of life to school nutritious feeding programs and CAs of students in poverty, in this paper, we were concerned with the following two core research questions. First, could school nutritious feeding enhance the CAs among students in poverty based on the quasi-natural experiment of NIP implementations in China? Second, at the theoretical and empirical levels, what are the mechanisms by which school nutritious feeding affects the CAs of poor students? To answer these questions, we used nationally representative data from the China Education Panel Survey (CEPS) and constructed the Difference-in-Differences (DID) model to identify the causal relationship between NIP implementations and CA promotion of poor students. Then, based on the mediating effects model (Baron & Kenny, 1986), we explored the potential mechanisms in terms of students' health status, sense of belonging at school, and expectations for the future.

The contributions of this study are as follows. First, this study is complementary to related research fields of the evaluation of NIP in China. Previous literature on the benefits of NIP implementations for students focused on health, educational achievements, and non-cognitive skills, but little attention was paid to the CAs (Wang et al., 2020; Wang & Cheng, 2022; Wang et al., 2022; Zhang et al., 2015). Our study helped to provide a more comprehensive picture of the effect of NIP implementations on human capital development of students in poverty in China. Second, since the NIP was used as part of the nutritional poverty alleviation program, our findings may help promote China's experience in poverty alleviation, as China has now fully eliminated absolute poverty. Thus, our study further enriched the research area on poverty alleviation and household human capital accumulation (Brady & Bostic, 2015; Li et al., 2021; Nong et al., 2022; Thompson et al., 2016), providing insights for other developing countries facing poverty problems. Third, in addition to health, our study identified other mechanisms by which the benefits of NIP implementation arise, including educational and psychological correlate factors such as student sense of belonging at school and positive expectations for the future. Thus, these findings may offer new directions for subsequent research.

2. Theoretical mechanisms and hypotheses

In this section, we explored the theoretical mechanisms of the school nutritious feeding on the CAs of students in poverty, including three possible mechanisms: student health, sense of belonging at school, and expectations for the future.

2.1. Student health status

Evidence from different countries confirmed the positive effect of school nutritious feeding programs on the physical and mental health of students (e.g., Afridi, 2010; Akin et al., 1983; Gundersen et al., 2012; Liu et al., 2019; Ren et al., 2023; Wang et al., 2022). On one hand, the school nutritious feeding is effective in reducing the incidence of hunger among students at school (Afridi, 2010), and increases their intake of calories and various nutrients, such as vitamins, minerals, and dietary fiber (Bhattacharya et al., 2006; Gleason & Suitor, 2003; Jomaa et al., 2011). On the other hand, the school nutritious feeding can promote dietary diversity among students and help them and their parents to enhance nutrition and health literacy (Bhattacharya et al., 2006). Nutrient supplementation and improved diet quality may reduce the incidence of malnutrition and stunting (Neumann et al., 2007), improve students' resistance to disease, and contribute to their physical health (Afridi, 2010; Akin et al., 1983; Gundersen et al., 2012).

In terms of mental health, dietary diversity is effective in securing students' intake of multiple nutrients such as zinc, folic acid, and magnesium, reducing the risk of depression (Jacka et al., 2012). Moreover, a high-quality diet can positively influence human biological functioning and improve students' mental health (Hatløy et al., 1998; Li et al., 2021). Obviously, for primary and secondary school students, physical and mental health is an important guarantee for the development of their intelligence, memory, observation, and imagination, which are important components of CAs (Dohmen et al., 2018; Heckman et al., 2006; Vittadini et al., 2021). In particular, students from families in poverty often suffer from a number of economic constraints that make them more vulnerable to malnutrition or poorer health (Nong et al., 2022), thus curtailing their cognitive development. Free school nutritious feeding may become an effective way to compensate for their nutritional and health vulnerabilities. Therefore, based on the above analysis, we proposed Hypothesis 1:

Hypothesis 1: The school nutritious feeding may enhance the health status of students in poverty, which helped promote their CAs.

2.2. Students' sense of belonging at school

The school nutritious feeding is likely to improve students' CAs by affecting their sense of belonging at school, particularly for students in poverty. Students' sense of belonging at school reflects how accepted, respected and supported they feel in their social context at school (Goodenow & Grady, 1993). Typically, students with a higher sense of belonging at school are likely to have better peer and teacher-student relationships and tend to exhibit higher levels of academic motivation, self-esteem, and achievements (Goodenow & Grady, 1993; Zhao & Chen, 2023). These outcomes are also favorable for student cognitive development. Moreover, better sense of belonging at school means that students are more actively involved in their schooling, extracurricular

activities, and interpersonal interactions, which can provide a favorable environment for their cognitive, emotional, and behavioral development (Raufelder et al., 2021). Additionally, from a social-psychological perspective, the need-to-belong theory indicates that the sense of belonging at school is beneficial to students' CA development (Baumeister, 2012), and this benefit may be more pronounced among students in poverty.

In many cases, students in poverty may be at a disadvantage in peer relationships due to their poorer family economic and social status, with the potential for a poor sense of belonging at school. However, school nutritious feeding program implementations are likely to motivate students to attend school and reduce absenteeism or tardiness at school (Frisvold, 2015; Zheng et al., 2023), which indicates that the school nutritious feeding is strongly associated with better sense of belonging at school for students. Furthermore, attending school more actively helps to promote better student–teacher and peer relationships, a greater enjoyment of school life, and stronger sense of belonging (Ananga, 2011). Therefore, Hypothesis 2 is proposed as below:

Hypothesis 2: The school nutritious feeding may enhance poor students' sense of belonging at school, which helps to promote their cognitive development.

2.3. Expectations for the future

The school nutritious feeding may also improve poor families' expectations for the future, which further elevates students' CAs. Prior literature confirmed that access to school nutritious feeding programs has significant improvements in students' non-cognitive abilities and academic achievements (e.g., Anderson et al., 2018; Chakraborty & Jayaraman, 2019; Frisvold, 2015; Liu et al., 2019; McEwan, 2013; Zheng et al., 2023), which are important for generating higher returns on household human capital investment in the future (Bütikofer et al., 2018). Considering that parents' investment decisions in their children's education are determined by future net returns, families in poverty need to make more difficult trade-offs about current consumption and investment in education (Wang et al., 2020). Fortunately, the school nutritious feeding can supplement families' food expenditures and improve the economic situation of families in poverty (Long, 1991; Wang & Cheng, 2022), alleviating the economic constraints on human capital investment (Wang et al., 2022), which allows parents to invest more physical capital and energy in their children's learning.

Moreover, NIP implementations are likely to raise underprivileged students' positive self-expectations for their future. On the one hand, the school nutritious feeding saves household expenses and slows down the likelihood that poor students have to run for their livelihood (Chen et al., 2021; Nong et al., 2022), which in turn generates higher educational expectations. On the other hand, the school nutritious feeding gives these students hope for a future out of poverty, which further increases their expectations for their future educational attainment and occupational acquisition. Clearly, according to the Rosenthal effect (e.g., Li et al., 1992; Rosenthal & Rubin, 1982; Turney et al., 2023), as the educational expectations increase, students may become more diligent and more actively involved in activities related to intellectual development. Also, poor students and their parents tend to have lower positive expectations for the future (Thompson et al., 2016), which may be more significantly compensated by school nutritious feeding programs. According to these theoretical analyses, we derived Hypothesis 3:

Hypothesis 3: The school nutritious feeding promotes poor students' and their parents' expectations for the future, which is likely to enhance children's CAs

3. Methods

To explore the impact of the school nutritious feeding program on the CAs of poor students in China, we resorted to the CEPS data and the DID method for analysis. In this section, we introduced the data and samples, variables, and empirical strategy used in the empirical estimation below.

3.1. Data resource

The data used in this study originated from two waves of the CEPS in the 2013–2014 and 2014–2015 school years, which is managed by the National Survey and Research Center of Renmin University of China. The CEPS is a nationally representative survey of junior high school students in China, particularly providing abundant information at four levels: students, households, classes, and schools (Tani et al., 2021; Zhao & Chen, 2022). This database is widely used in the research field on the effects of families, schools, classes, and macro-social structures on individual human capital outcomes (e.g., Chen & Zhao, 2023; Qiu & Liang, 2024; Zhou et al., 2020). In these two waves of surveys, the first wave surveyed 7th and 9th graders, and the second wave surveyed only 8th graders (the 7th graders in the first wave). To explore the effect of NIP implementations on students' CAs in the DID framework, we retained 7th graders in the first wave and 8th graders in the second wave.

In the CEPS questionnaire, in addition to basic student and parent demographic characteristics, family economic status and student performances, such as examination results, CAs, and non-cognitive abilities, are also included. More importantly, this database provides some important information about school characteristics. For example, the school leader was asked if the school is an NIP pilot and when it was implemented (Wang et al., 2022; Zheng et al., 2023). As a result, the CEPS provides an available and reliable data source for investigating the influence of school nutritious feeding on the CAs of students in poverty.

3.2. Definition of poverty and sample selection

As mentioned earlier, the NIP was piloted at the school level and did not target only students in poverty, but based on the theoretical analysis, we believed that NIP implementations may have a more positive impact on CAs for students in poverty. Therefore, in this subsection, we needed to define poverty and select the sample from the CEPS, as we were quite concerned with these disadvantaged students in poverty. Academics define poverty mainly in terms of absolute poverty and relative poverty. The former mainly refers to per capita household income below a government-mandated poverty line, while there is no uniform standard for measuring the latter (Foster, 1998; Ravallion & Chen, 2011). Considering nutrition and dieting as one of the necessary conditions for the growth and development and the promotion of CAs for children, the positive impact of NIP is likely to be greater for those students in absolute poverty, because they may not get enough nutrition from their families (Nelson, 2000; Valdivia, 2004). Instead, relative poverty is dependent on the average level of local socioeconomic development (Brady & Bostic, 2015; Joo, 2011). Students from relatively poor families do not necessarily suffer from undernutrition and nutritional gaps. Thus, in our main specification, we selected students from absolutely poor families as our sample from the CEPS.

Unfortunately, the CEPS did not provide information on specific household income, so we were unable to select the poverty line as a criterion as in the conventional practice. However, the CEPS questionnaire includes information on whether the family received the basic living allowance (BLA). The BLA is an important social security measure to protect low-income households in China, and only residents whose annual per capita household income is below the minimum local subsistence level can apply for the BLA (Chan, 2010). In this paper, we considered students from households receiving the BLA to be in absolute poverty (Wang et al., 2022). Thus, in the main results, our full sample consists of all students from BLA families in the two waves of CEPS data.

3.3. Variables

The dependent variable is *Cognitive ability*, which is measured by the

standardized scores on the students' CA test performed by the CEPS group (Zhao & Chen, 2023; Tani et al., 2021). This test covers 11 different types in three dimensions, including language, graphics and space, and calculation and logic. Also, students were asked to gather in classrooms to answer these questions on this test. Further, based on Item Response Theory (IRT), the CEPS group standardized students' CA scores. The parameters of the IRT model include difficulty index, discriminative powder index, and guessing index. The standardized CA scores calculated based on these three parameters were also provided by the CEPS database.[5] Compared to the original value of the CA scores, standardized scores help to reduce heteroskedasticity and exclude the interference of outliers. Moreover, the CA is more objective and comparable in CEPS than other variables measuring students' human capital, such as examination results and non-cognitive abilities (Chen & Zhao, 2023; Gong et al., 2018), as all students in the same grade are given the same test questions in each wave of the survey.

The independent variable is *Nutrition Improvement Program*, a dummy variable. Although the NIP has been implemented in some schools in some counties since 2011, it is impossible to determine which schools are truly affected by the policy because the CEPS has not announced the specific names of schools and counties. Fortunately, as previously stated, school leaders were asked if the school had become an NIP pilot since a specific year. Based on this information, we found that in the full sample (consisting of students in poverty) from the CEPS, the first wave included 11 schools in 7 counties, and the second wave included 9 pilot schools in 6 counties. However, in this paper, the DID empirical design requires differences from two dimensions, school and school year, which requires at least one school year without the policy pilots. This is because if we included pilot schools from both waves in the DID method, samples from the pilot schools in the first wave did not have a control group of the survey prior to the policy, resulting in the parallel trend assumption not being met and the DID method not being used (De Chaisemartin & d'Haultfoeuille, 2020; Zhao & Chen, 2023). As a result, we could only assess the policy effects of the NIP in the second wave of pilot schools in this study.[6] The setting rule for the variable *NIP* is as follows. If a school becomes an NIP pilot in 2014–2015 school year, it can be set to 1, and 0 otherwise. In the analysis below, we included the NIP pilots in the first survey wave in the regressions to conduct a robustness test, but this is not a DID estimate.

To more accurately identify the causal relationship between the NIP and the CAs of students in poverty, we chose a series of control variables. As highlighted in the previous studies, these variables are regarded as important factors influencing students' CAs (e.g., Frisvold, 2015; Li et al., 2021; Liu et al., 2020; Zhao & Chen, 2022). For school characteristics, *Canteen*, *School ranking*, *Type*, and *Locations* were added in the set of control variables. From the perspective of student characteristics, we controlled for *Gender*, *Year of birth*, *Hukou*, *Kindergarten*, *Left-behind children*, and *Migrant children*. In terms of family characteristics, education of parents and father's occupation were included. Last, we further controlled for students' cognitive endowment with a range of variables including *Ranking at 6th grade*, *Difficulty in learning math*, *Difficulty in learning Chinese*, *Grade repetition experience*, *Learning ability*, *Learning attitude*, *Study effort*, and *Curiosity*. The definitions of above variables are shown in Table A1 of the Appendix.

3.4. Estimation strategy

As mentioned in Section 1, the NIP was introduced gradually, with schools and students as passive recipients of this program. Thus, the NIP can be regarded as an exogenous policy, providing a quasi-natural experiment for us to use the DID approach to confirm the effect of school nutritious feeding on the CAs of students in poverty (Fang & Zhu, 2022; Lundborg et al., 2022; Ren et al., 2023; Wang et al., 2022; Zheng et al., 2023). Specifically, the first difference comes from the school level, that is, the difference between NIP pilot and non-pilot schools, where students in the pilot schools are considered as the treatment group and those in non-pilot schools as the control group; the second difference is from the time level, i.e., the difference in school years before and after the implementation of the NIP. Comparing these two differences captures the short-term changes in CAs of students in poverty after NIP implementations (Wang et al., 2022; Zheng et al., 2023). The standard DID model for baseline results was set up as follows:

$$CA_{ist} = \beta_0 + \beta_1 NIP_{ts} + \beta_2 X_{ist} + \rho_s + \theta_t + \lambda_{ct} + \varepsilon_{ist} \quad (1)$$

where i indicates the student in poverty, t refers to the school year (2013–2014 or 2014–2015), s represents the school, and c is the county.CA_{ist} means the standardized scores for CAs of students in poverty. NIP_{ts} indicates whether the school s has become an NIP pilot in school year t.β_1 is the core DID estimate we focused on, which involved the net effect of NIP implementations on CAs of students in poverty.X_{ist} is a series of control variables as mentioned above. Moreover,ρ_s and θ_t are the school fixed effect (FE) and school year FE, respectively.λ_{ct} is the year-by-county FE, a high-dimensional FE.ε_{ist} is a random error term. To control for serial correlation as well as heterogeneity of variables, we clustered the standard errors at the school level.

It needs to be emphasized that although we used the DID model to identify the causal effect between NIP implementations and CAs of students in poverty, there are still potential endogeneity issues that could cause our estimates to be biased. The first concern is the NIP pilot selection bias. The selection of NIP pilots cannot be randomized (Zhang et al., 2015; Wang & Cheng, 2022), resulting in our DID estimates being challenged by district and school heterogeneity. The second came from student motivation to transfer. Since the NIP pilots were implemented in different schools, parents may transfer to pilot schools for free nutritional meals, which caused some of the samples in the control group to become unsuitable (Lu et al., 2019). The third threat was from the parallel trend assumption for DID methods (Wang & Cheng, 2022), as the CEPS only had two waves of survey data as mentioned earlier.

To alleviate these potential issues, we made the following efforts and explanations under the framework of DID methods. First, to mitigate the selectivity of NIP pilots, we strictly controlled for a number of school characteristic variables as discussed earlier. Moreover, we added the school FE and year-by-county FE, effectively capturing fixed school characteristics and county characteristics that change over time. These measures help to mitigate potential bias from NIP selectivity (Wang & Cheng, 2022; Zhao & Guo, 2022). Second, to ease the challenges of student transfers, we used a sample of students in poverty who may have received focused attention from schools and teachers for a long time in China (Nong et al., 2022), which reduces the likelihood of their transferring to other schools. Furthermore, in the robustness tests, we excluded estimated bias from student transfers as much as possible by excluding migrant children and keeping one school per county. Third, to satisfy the parallel trend assumption, as highlighted above, we only analyzed the second wave of NIP pilot schools so that we could test for the vital parallel trend in the DID empirical design.

Furthermore, to further test the hypotheses presented in Section 2, we constructed the following mediating effects model (Baron & Kenny, 1986):

$$M_{ist} = \alpha_0 + \alpha_1 NIP_{ts} + \alpha_2 X_{ist} + \rho_s + \theta_t + \lambda_{ct} + \mu_{ist} \quad (2)$$

$$CA_{ist} = \delta_0 + \delta_1 NIP_{ts} + \delta_2 M_{ist} + \delta_3 X_{ist} + \rho_s + \theta_t + \lambda_{ct} + \zeta_{ist} \quad (3)$$

Among them,M_{ist} indicates a collection of mediating variables, containing students' health, sense of belonging at school, and expectations

[5] More details on standardized CA calculations can be found at https://ceps.ruc.edu.cn/_local/6/6A/D5/2E0FAE92BBE95B3AF701570D56A_84A8EA20_8EE1D.pdf?e=.pdf.

[6] Specifically, the 433 samples from the first wave of the survey came from 11 schools that were removed.

for the future. In general, the following two conditions need to be met for the mediating effect to hold. On the one hand, the coefficient α_1 of NIP_{ts} in equation (2) and the coefficient δ_2 of M_{ist} in equation (3) are both statistically significant and their signs are in line with the theoretical expectations. On the other hand, the absolute value of the coefficient δ_1 in equation (3) decreases compared to β_1 in equation (1). μ_{ist} and ζ_{ist} are random error terms. The other variables are consistent with equation (1) and will not be repeated.

4. Findings

4.1. Descriptive statistics

Descriptive statistics of the main variables were shown in Table 1. Our full sample size consisting of students in poverty (receiving the BLA) was 1100, of which 99 students were in NIP pilot schools, accounting for 9 %. Moreover, we found that the mean of the standardized CA score was −0.19, suggesting that there may be a deficit in the CAs for students in poverty.[7] The standard deviation was 0.83, which is much higher than the mean, indicating that the CAs of students in poverty may vary significantly. Comparing the treatment and control groups, the T-test results showed that the standardized CA scores of underprivileged students in the pilot schools were slightly higher than those in the non-pilots, but this difference is not statistically significant without considering other control variables. Furthermore, for school characteristics, there were significant differences between pilot and non-pilot schools in terms of geographic location and school rankings. T-test statistics were also significant for most of the variables on student and family characteristics, suggesting that controlling for these variables may help to reduce the differences between these two groups to better capture the causal effects of the NIP on CAs of students in poverty in China.

Table 1
Variable definitions and descriptive statistics.

Variables	Full sample (N = 1100)		Non-NIP schools (N = 1001)	NIP schools (N = 99)	T-tests
	Mean	SD	Mean	Mean	Diff.
CA	−0.19	0.83	−0.190	−0.185	−0.005
NIP	0.09	0.29			
Canteen	0.76	0.43	0.76	0.75	0.01
School ranking	3.76	0.83	3.73	4.07	−0.34***
Type	0.93	0.25	0.93	1.00	−0.07***
Locations	3.13	1.57	3.03	4.11	−1.08***
Gender	0.53	0.50	0.53	0.57	−0.03
Year of birth	2000.34	0.77	2000.35	2000.19	0.16**
Hukou	0.61	0.49	0.60	0.76	−0.16***
Kindergarten	0.76	0.43	0.76	0.68	0.09*
Left-behind children	0.32	0.47	0.34	0.15	0.19***
Migrant children	0.13	0.34	0.14	0.03	0.11***
Only child	0.27	0.45	0.29	0.13	0.16***
Education of father	2.11	0.81	2.13	1.92	0.21**
Education of mother	1.89	0.89	1.90	1.78	0.12
Job of father	0.60	0.49	0.57	0.91	−0.34***
Ranking at 6th grade	2.40	0.49	2.41	2.29	0.12**
Difficulty in learning math	2.32	1.02	2.32	2.25	0.07
Difficulty in learning Chinese	2.02	0.86	2.01	2.09	−0.08
Grade repetition experience	0.27	0.44	0.26	0.39	−0.14***
Learning ability	2.91	0.81	2.92	2.88	0.04
Learning attitude	3.50	0.78	3.50	3.60	−0.10
Study effort	3.32	0.89	3.30	3.47	−0.17*
Curiosity	3.44	0.77	3.43	3.52	−0.09

Note: ***, **, and * denote significance levels of 1%, 5%, and 10%, respectively.

[7] The average CA score for non-poor students in the CEPS is 0.19.

4.2. The effect of school nutritious feeding on the CAs of students in poverty

Table 2 reported the baseline empirical results on the effect of NIP on CAs of students in poverty by using the DID model. In column (1), we controlled for all FEs and school characteristic variables, which are likely to be closely related to NIP pilot selections, finding that the estimated coefficient on *NIP* was significant being at the 1 % level. Moreover, in columns (2)-(4), we further added variables on student characteristics, parent characteristics, and student cognitive

Table 2
The effect of school nutritious feeding on CAs of students in poverty: DID estimations.

	(1)	(2)	(3)	(4)
NIP	0.24***	0.33***	0.36***	0.33***
	(0.00)	(0.08)	(0.08)	(0.10)
Canteen	−0.51***	−0.45***	−0.45***	−0.45***
	(0.00)	(0.01)	(0.02)	(0.03)
Locations	0.74***	0.72***	0.71***	0.66***
	(0.00)	(0.02)	(0.02)	(0.02)
Type	0.24***	0.33***	0.40***	0.21*
	(0.00)	(0.07)	(0.08)	(0.11)
School ranking	−0.52***	−0.50***	−0.53***	−0.47***
	(0.00)	(0.02)	(0.03)	(0.03)
Gender		−0.05	−0.05	−0.10*
		(0.06)	(0.06)	(0.05)
Year of birth		0.22***	0.23***	0.19***
		(0.04)	(0.04)	(0.03)
Hukou		−0.02	−0.03	−0.05
		(0.06)	(0.06)	(0.06)
Kindergarten		0.04	0.03	0.03
		(0.06)	(0.06)	(0.05)
Left-behind children		0.04	0.04	0.06
		(0.07)	(0.07)	(0.07)
Migrant children		0.02	0.01	0.02
		(0.10)	(0.10)	(0.10)
Only child		−0.04	−0.03	−0.02
		(0.06)	(0.06)	(0.07)
Education of father			−0.04	−0.05
			(0.04)	(0.04)
Education of mother			−0.03	−0.03
			(0.04)	(0.04)
Job of father			−0.02	−0.04
			(0.06)	(0.06)
Ranking at 6th grade				−0.23***
				(0.07)
Difficulty in learning math				−0.14***
				(0.03)
Difficulty in learning Chinese				0.01
				(0.03)
Grade repetition experience				−0.04
				(0.06)
Learning ability				−0.01
				(0.03)
Learning attitude				−0.00
				(0.04)
Study effort				0.05*
				(0.03)
Curiosity				0.04
				(0.03)
Year FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes
Year-by-county FE	Yes	Yes	Yes	Yes
R-squared	0.21	0.25	0.25	0.30
N	1100	1100	1100	1100

Notes: ***, **, and * denote significance levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the school level are reported in parentheses.

endowment, and the DID estimates remained unchanged. In particular, in column (4), we controlled for a number of variables related to students' cognitive endowment and study habits, fingding that the coefficient of *NIP* kept significantly positive at the 1 % statistical level. These results suggest that NIP implementations were likely to enhance the CAs of students in poverty.

The results of the preferred main specification, column (4), show that NIP implementations resulted in a 0.33-point improvement in the CA scores of students in poverty. Consistent with what was highlighted in the previous literature (Ren et al., 2023; Wang et al., 2020; Wang et al., 2022; Wang & Cheng, 2022; Zheng et al., 2023), our baseline results also showed a positive effect of school nutritious feeding in promoting students' human capital accumulation. Importantly, new evidence from China suggests that the anti-poverty nutrition policy may have the potential benefits of promoting educational equity and improving the welfare of poor households.

4.3. Robustness checks

4.3.1. Parallel trend assumption testing

A central premise of using the DID model is the need to satisfy the parallel trend assumption (De Chaisemartin & d'Haultfoeuille, 2020; Lundborg et al., 2022; Zhao et al., 2022). That is, before becoming pilot NIP schools, the trend in CAs was flat for students in the treatment and control groups. Following conventional practice and drawing on the idea of event analysis (e.g., Fang & Zhu, 2022; Lu et al., 2019; Zhao & Chen, 2023; Zhou et al., 2020), we constructed a dummy variable *NIP_1* equal to 1, assuming that NIP schools in the second wave were piloted in the first wave, and 0 otherwise. In column (1) of Table 3, the coefficient on *NIP* kept significant while the coefficient on *NIP_1* was insignificant. These estimates suggested that prior to the NIP implementations, there were no differences in CAs between all students in poverty in these two groups. In contrast, after being NIP pilots, poor students in the treatment group showed a significant increase in standardized CA scores. In summary, these results implied that the parallel trend assumption was satisfied and it was reasonable to use DID estimation in this study.

4.3.2. Alternative specifications

We further performed robustness tests by using alternative specifications. First, in the main results, to construct a standard DID model and test the parallel trend hypothesis, we excluded the first wave of NIP pilots. In the robustness checks, we included these samples in column (2) of Table 3. Apparently, we found that the estimate for *NIP* is very similar to the baseline results. The OLS results also imply that the school nutritious feeding helped to raise the CAs of Chinese students in poverty.

Second, because the NIP policy was not yet fully rolled out in 2014, we found a gap in sample size between the treatment and control groups (as shown in Table 1). In this case, it is common practice to use the propensity score matching (PSM) method to reduce this difference to enhance the comparability of two groups (e.g., Nong et al., 2022; Pierce et al., 2022; Zhou et al., 2020). Column (3) in Table 3 reports the empirical results using PSM-DID estimation. The coefficient on *NIP* remains significantly positive despite the sample size being reduced by over 65 % compared to the baseline regressions. Thus, these results indicated that after minimizing the difference in sample size between the treatment and control groups, our baseline results continued to be reliable.

Third, in terms of estimation methodology, in addition to the PSM-DID approach, we further used the single-difference method for robustness testing. The single-difference method aims to explore the changes in CAs of students in poverty before and after the NIP implementations and does not take into account the differences between the treatment group and control group (Slaughter, 2001). In column (4) of Table 3, we discovered that the coefficient for *NIP* remained significantly positive at the 1 % statistical level despite a somewhat change in the magnitude. These results suggest that after changing the estimation method, the conclusion that NIP implementations promoted the CAs of students in poverty remained unchanged.

Fourth, although our sample selection from the CEPS relies on absolute poverty and does not consider the relative poverty status, as local socioeconomic conditions are not uniform, households in absolute poverty on BLA are not necessarily in relative poverty. Therefore, we defined poverty more strictly by excluding families that are not in relative poverty. Based on previous studies (Ravallion & Chen, 2011; Wang et al., 2020), relative poverty was identified based on the answer to the question, "Which one of the following best describes the financial conditions of your family at present?" Options included "very poor", "somewhat poor", "moderate", "somewhat rich", and "very rich". In column (5) of Table 3, we excluded all observations answering as "moderate" from the full sample,[8] finding a significant NIP estimate. In addition, the coefficient is larger compared to the baseline results, suggesting that for more disadvantaged students (simultaneous absolute and relative poverty), the positive effect of NIP on their CA improvement may be greater.

Last, to further remove potential endogeneity, we more tightly controlled for some FEs (Zhao & Guo, 2022; Zhao et al., 2023). Specifically, in column (6) of Table 3, we additionally controlled for school rank FE, school location FE, and student birth year FE. The results show that the estimated coefficient on *NIP* was still significantly positive. Therefore, after tighter controls to exclude potential interference from school characteristics and student birth cohorts, our main specification was not shocked. Moreover, as mentioned in Section 3, to exclude interference in baseline estimates due to student transfers, we kept only one school in each county in column (7) and excluded all migrant children in column (8) of Table 3. The former avoided intra-county migration as much as possible, and the latter overcame disruptions from the inter-county migration sample. We observed that these measures have little effects on our main results.

4.3.3. Placebo tests

To cope with the estimation bias caused by unpredictable factors, we conducted a placebo test (e.g., Li et al., 2016; Zhao et al., 2023). On the one hand, we randomly selected NIP pilots from all schools as a treatment group. Sample sizes are consistent with the baseline results. On the other hand, the 2013–2014 or 2014–2015, one school year is randomly selected as the period being the NIP pilots. The pseudo-policy variable is noted as NIP^p, which was regarded as a new DID estimate. By repeating 500 times in placebo tests, the identification power can be further improved. Results were reported in Fig. 1. We could see that the distribution of estimates from random assignments is centered around 0, suggesting that randomly constructed NIP pilots do not affect the CAs of poor students. Additionally, the vast majority of coefficients lie outside the 95th percentile of the distribution. Taken together, the unobserved factors may not have a significant impact on our main results.

4.4. Mechanism tests

In this subsection, we adopted empirical methods to test the mechanisms and hypotheses, including student health status, students' sense of belonging at school, and expectations for the future.

4.4.1. Health status

In Hypothesis 1, we considered that the school nutritious feeding has the potential to enhance the CAs of poor students by improving their health. To test this mechanism, according to the CEPS, we selected five health-related mechanism variables (e.g., Gundersen et al., 2012; Mostert, 2021; Ren et al., 2023; Wang et al., 2023), namely *Self-rated health*, *Frequent illness*, *Height*, *Weight*, and *Underweight*. Among these, the first

[8] In the sample (consisting of students in absolute poverty), there were no respondents whose answer were "somewhat rich", and "very rich".

Table 3
Parallel trend tests and alternative specifications.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Parallel trend tests	Including NIP schools in 2013	PSM-DID	Single-difference method	Stricter definition of poor families	More FEs	Keeping one school in each county	Migrant children
NIP	0.38**	0.29***	0.44**	1.68***	1.03**	0.70***	1.16***	0.40***
	(0.12)	(0.03)	(0.08)	(0.24)	(0.40)	(0.10)	(0.23)	(0.13)
NIP_1	0.09							
	(0.18)							
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-by-county FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1100	1400	355	143	475	1100	475	746

Notes: The control variables are consistent with column (4) of Table 2. In column (6), we additionally controlled for birth year FE, school location FE, and school ranking FE. ***, **, and * denote significance levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the school level are reported in parentheses.

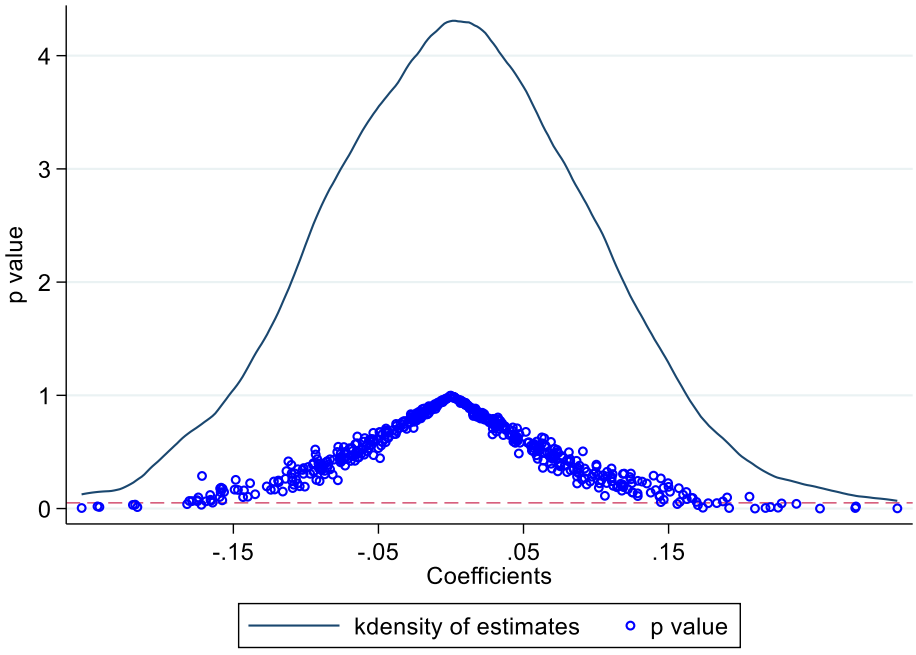


Fig. 1. Placebo tests.

two primarily measure students' overall health status and illness, while the last three relate to students' physical development. Definitions of these variables are shown in Panel A of Table A2 in the Appendix.

In Panel A of Table 4, we found that the DID estimates for *NIP* were significantly positive in column (1) and significantly negative in column (2), indicating that NIP implementations promoted the overall self-rated health of students in poverty and the likelihood of them getting sick. In terms of poor students' physical development, we found that compared to poor students in non-NIP pilot schools, those in NIP pilot schools were 3.53 cm taller, 1.53 kg heavier, and had a 19 % lower probability of being underweight. Not surprisingly, our estimates reveal the benefits of NIP implementation in improving the health of students in poverty, which is similar to the findings in the previous literature (Fang & Zhu, 2022; Ren et al., 2023; Wang et al., 2023; Zhang et al., 2015).

We noted that compared to the baseline results in Table 2, there are some missing values for these mediating variables. Thus, to accurately capture the mediating effects of health promotion, we re-reported the baseline results in Panel B after removing the missing values of the corresponding mediating variables. It is clear that the coefficients on *NIP* remained significantly positive. Moreover, in Panel C of Table 4, we conducted mediation effects tests based on equation (3). We could see that self-rated health is significantly positively associated with CAs of students, while frequent illness is significantly negatively associated with the CAs, suggesting that being in poor health would similarly lower CAs for students in poverty in our sample. As discussed in prior research, health is an important guarantee for the development of children's intelligence, memory, observation and imagination, which are also components of CAs (Dohmen et al., 2018; Heckman et al., 2006).

We further focused on the coefficients on *NIP* in Panel C. Clearly, the absolute values of the estimated coefficients on *NIP* decline in columns (1) and (2) compared to those in Panel B, indicating that self-rated health and frequent illness partially explain the positive effect of NIP implementations in boosting the CAs of students in poverty. In contrast, the mediating effect of self-rated health is more pronounced, with an even more substantial drop in the magnitude. Taken together, we provided supporting evidence for **Hypothesis** 1, confirming that the school nutritious feeding had a positive effect on the physical health of students in poverty, laying the necessary nutritional foundation for their CA improvement.

Table 4
Mechanism tests for health status.

	(1)	(2)	(3)	(4)	(5)
Panel A. The NIP and health status for students in poverty					
	Self-rated health	Frequent illness	Height	Weight	Underweight
NIP	1.52***	−0.08**	3.53***	1.53***	−0.19**
	(0.13)	(0.04)	(1.19)	(0.22)	(0.09)
N	1069	1075	1030	1021	1013
Panel B. The NIP and students' CAs (Baseline results, excluding missing values of mediating variables)					
	Students' CAs	Students' CAs	Students' CAs	Students' CAs	Students' CAs
NIP	0.30***	0.34***	0.37***	0.56***	0.56***
	(0.10)	(0.10)	(0.10)	(0.17)	(0.17)
N	1069	1075	1030	1021	1013
Panel C. Mediating effects of health status					
	Students' CAs	Students' CAs	Students' CAs	Students' CAs	Students' CAs
NIP	0.25**	0.33***	0.39***	0.56***	0.55***
	(0.11)	(0.10)	(0.10)	(0.17)	(0.17)
Self-rated health	0.032*				
	(0.018)				
Frequent illness		−0.16*			
		(0.09)			
Height			−0.01		
			(0.00)		
Weight				0.00	
				(0.03)	
Underweight					−0.03
					(0.08)
N	1069	1075	1030	1021	1013

Notes: Baseline control variables, year FE, school FE, and year-by-county FE are added to all regressions in three Panels. The baseline control variables are consistent with column (4) of Table 2. ***, **, and * denote significance levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the school level are reported in parentheses.

4.4.2. *Sense of belonging at school*

Hypothesis 2 emphasized that school nutritious feeding may enhance underprivileged students' sense of belonging to the school, thus contributing to their cognitive development. Referring to previous literature and the information provided by the CEPS (Huang, 2022; Zhao & Chen, 2023), four variables were selected to capture students' sense of belonging to their schools, including *Like the school*, *Like the homeroom teacher*, *Like the subject teachers*, and *Like classmates* (see Panel B of Appendix Table A2 for specific definitions).

In Panel A of Table 5, we found that the coefficients of *NIP* are all positive in columns (1) and (2) by using the DID method, suggesting that the school nutritious feeding did enhance the sense of belonging at school of students in poverty, as the school was likely to meet their nutritional needs which they may not be able to obtain from their families (Wang et al., 2022; Wang et al., 2023). In addition, the NIP implementations primarily improved students' liking of their homeroom teachers, rather than the subject teachers and classmates. This may be due to the fact that in China, homeroom teachers are not only classroom managers, but also provide material and psychological support to students (Chen & Zhao, 2022; Chen & Zhao, 2023), especially for students in poverty. The implementation of NIP also requires the communication, management and cooperation of homeroom teachers.

As before, Panel B of Table 5 shows the baseline results after excluding the missing values of the mediating variables related to students' sense of belonging at school. It is clear that our main results remain unchanged. The results of mediating effects tests are reported in Panel C of Table 5. As expected, all coefficients on mediating variables are positive. As emphasized in previous literature, sense of belonging at school is strongly associated with higher academic motivation, self-esteem, and sense of achievement (Goodenow & Grady, 1993; Zhao & Chen, 2023), which can improve the CAs of students in poverty. Moreover, comparing the results from Panel B, we found that the absolute values of the coefficients on *NIP* decrease in the first two columns of Panel C, implying that students' sense of belonging at school is an effective channel. In particular, this mediating effect comes from poor students' favoritism of schools and homeroom teachers.

Thus, our results support **Hypothesis** 2 that the school nutritious

Table 5
Mechanism tests for the sense of belonging at school.

	(1)	(2)	(3)	(4)
Panel A. The NIP and the sense of belonging at school of students in poverty				
	Like the school	Like the homeroom teacher	Like subject teachers	Like classmates
NIP	0.34***	0.63***	−0.07	−0.11
	(0.08)	(0.06)	(0.06)	(0.07)
N	1095	1091	1079	1091
Panel B. The NIP and students' CAs (Baseline results, excluding missing values of mediating variables)				
	Students' CAs	Students' CAs	Students' CAs	Students' CAs
NIP	0.29***	0.36***	0.35***	0.34***
	(0.10)	(0.10)	(0.10)	(0.10)
N	1095	1091	1079	1091
Panel C. Mediating effects of the sense of belonging at school				
	Students' CAs	Students' CAs	Students' CAs	Students' CAs
NIP	0.26***	0.30***	0.36***	0.34***
	(0.10)	(0.11)	(0.10)	(0.10)
Like the school	0.10**			
	(0.05)			
Like the homeroom teacher		0.10*		
		(0.06)		
Like subject teachers			0.15**	
			(0.07)	
Like classmates				0.07
				(0.05)
N	1095	1091	1079	1091

Notes: Baseline control variables, year FE, school FE, and year-by-county FE are added to all regressions in three Panels. The baseline control variables are consistent with column (4) of Table 2. ***, **, and * denote significance levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the school level are reported in parentheses.

feeding made students in poverty enhance their sense of belonging to the school, which is quite helpful in improving their CAs and promoting human capital accumulation.

4.4.3. Expectations for the future

As highlighted in Hypothesis 3, the school nutritious feeding was likely to enhance students' CAs by influencing poor students' and parents' expectations for the future. We selected four variables to describe these expectations at two levels. On the one hand, for student self-expectations, we constructed two variables, namely *Self-educational expectations* and *Self-occupational expectations*. On the other hand, parental expectations included educational expectations and occupational expectations for their children. Also, the details of these mechanism variables are presented in Panel C of Table A2 in the Appendix.

In Panel A of Table 6, we first examined the impact of NIP implementations on poor students' and their parents' expectations for the future through the DID method. We found that the estimated coefficients on *NIP* are significantly positive in the first two columns but not significant in the last two columns. These results suggest that the school nutritious feeding increased student self-expectations for the future. The NIP reduces the expenses of poor families and gives children a glimpse of the future possibilities of getting out of poverty, which motivates them even more to learn.

In Panel B of Table 6, we then re-estimated the baseline results after excluding the missing values of the mediating variables related to expectations for the future, finding that the coefficients on *NIP* were very close to those in Table 2. Not surprisingly, results in Panel C of Table 6 indicate that positive expectations for the future are positively correlated with the CAs for students in poverty. As implied by the Rosenthal effect (e.g., Li et al., 1992; Rosenthal & Rubin, 1982; Turney et al., 2023), both educational and career expectations are significantly and positively associated with students' CA performance, which is in line with the results in Panel C.

We further explored the mediating effect in Panel C of Table 6 and observed that compared to Panel B, the absolute values of the coefficients on *NIP* decreased in columns (1) and (2), but this downward trend is not evident in columns (3) and (4). Also, the positive expectations about the future appear to have a greater mediating effect than health improvement and the sense of belonging at school analyzed above. The potential reason is that for these disadvantaged students in poverty, NIP implementations gave them hope and provided the physical and human capital conditions for realizing beliefs about escaping poverty in the future (Fang & Zhu, 2022). All in all, our findings supported **Hypothesis** 3 that by promoting poor students' expectations for the future, the school nutritious feeding program improved their CA performance.

5. Conclusions and discussion

5.1. Conclusions

In recent years, while some scholars have focused on the impact of school nutritious feeding on students in China (e.g., Fang & Zhu, 2022; Ren et al., 2023; Wang et al., 2020; Wang & Cheng, 2022; Wang et al., 2022), they neglected the CA, a key component of human capital accumulation. More importantly, students in poverty may benefit more from the school nutritious feeding, as they are at greater risk of undernutrition. As an important part of the nutritional poverty alleviation program in China, it is necessary to explore the impact of NIP implementations on the CAs of poor students.

By regarding the NIP implementation as a quasi-natural experiment, we adopted the DID approach to explore the causal relationship between the school nutritious feeding and the CAs of students from poor families in China. We draw the following conclusions from our results. First, we found that NIP implementations resulted in an increase of 0.33 points in the CA scores of students in poverty, confirming the effectiveness of the school nutritious feeding policy in China. Second, the baseline results are supported by parallel trend tests and other robustness checks, like placebo tests and alternative specifications. Third, according to the results of mediation effect analysis, the effect of NIP implementations in promoting the CAs of students in poverty can be partially explained by improving health status, enhancing sense of belonging at school, and increasing positive expectations for the future. In particular, we found that the positive expectations among poor students may have a greater mediating effect than students' health status and sense of belonging at school.

Table 6
Mechanism tests for expectations for the future.

	(1)	(2)	(3)	(4)
Panel A. The NIP and expectations for the future for students in poverty				
	Self-educational expectations	Self-occupational expectations	Parents' educational expectations	Parents' occupational expectations
NIP	0.43***	1.09***	0.00	−0.01
	(0.05)	(0.16)	(0.05)	(0.07)
N	1095	1093	1094	1093
Panel B. The NIP and students' CAs (Baseline results, excluding missing values of mediating variables)				
	Students' CAs	Students' CAs	Students' CAs	Students' CAs
NIP	0.31***	0.32***	0.33***	0.34***
	(0.10)	(0.10)	(0.10)	(0.10)
N	1095	1093	1094	1093
Panel C. Mediating effects of expectations for the future				
	Students' CAs	Students' CAs	Students' CAs	Students' CAs
NIP	0.24**	0.26**	0.33***	0.34***
	(0.12)	(0.11)	(0.10)	(0.10)
Self-educational expectations	0.16*			
	(0.09)			
Self-occupational expectations		0.05**		
		(0.02)		
Parents' educational expectations			0.17***	
			(0.05)	
Parents' occupational expectations				0.14**
				(0.06)
N	1095	1093	1094	1093

Notes: Baseline control variables, year FE, school FE, and year-by-county FE are added to all regressions in three Panels. The baseline control variables are consistent with column (4) of Table 2. ***, **, and * denote significance levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the school level are reported in parentheses.

5.2. Discussion

There are two main theoretical insights from our study. First, as an important part of the national nutrition poverty alleviation plans in China, the effectiveness of NIP implementations has been underappreciated by scholars (Immink, 2010; Li et al., 2021). Our study revealed that the school nutritious feeding program significantly improved the CAs of students in absolute poverty in China, confirming that the nutritional poverty alleviation program is reliable in promoting human capital accumulation. Currently, China has fully eradicated absolute poverty and the NIP has made a contribution. This experience may be replicable in other developing countries. Thus, our study not only complements a series of studies on the positive effects of poverty alleviation (e.g., Joe et al., 2016; Liu et al., 2017; Nong et al., 2022; Zhou et al., 2023), but also provides new research topics on how scholars can draw on China's nutritional poverty alleviation program to promote child development.

Second, with the widespread implementation of school nutritious feeding programs around the world, a considerable number of scholars focused on the positive effect of this program on children development (e.g., Afridi, 2010; Gleason & Suitor, 2003; McEwan, 2013; Wang & Cheng, 2022). Although there are studies covering some developing countries (Afridi, 2010; Chakraborty & Jayaraman, 2019; Jomaa et al., 2011; McEwan, 2013), we further added to this body of literature based on the special Chinese context. More importantly, although previous literature recognized that the NIP provided additional benefits to students, such as health status, academic performance, and non-cognitive abilities (Anderson et al., 2018; Gordon & Ruffini, 2018; Gundersen et al., 2012; Imberman & Kugler, 2014; Millimet et al., 2010; Ren et al., 2023), we enriched this research field on school nutritious feeding programs and children's human capital accumulation at the level of CA development.

In terms of practical insights, we considered that the function of the government-led school nutritious feeding needs to be expanded to encourage the human capital investment among low-income and poor households. On one hand, the relevant supporting measures of school nutritious feeding should be further enhanced and financial subsidies for schools in poor, low-income, and remote areas should be increased to guarantee the basic nutritional needs of students in poverty. Possible measures included renovating the canteen, customizing nutritious recipes, and improving the dining environment. On the other hand, since poor families may neglect child nutrition (Morris et al., 2019), it is important to actively promote the positive effects of improved student nutrition by school nutritious feeding. For schools, training of teachers and canteen staffs on relevant nutrition knowledge needs to be strengthened (Wang et al., 2022). For grassroots organizations, more nutrition knowledge should also be actively disseminated through village committees, so that poor families can give priority to the nutritional status of their children, thus giving full play to the poverty alleviation effect of school nutritious feeding programs. In conclusion, we emphasized the importance of strengthening school nutrition feeding as an effective way to achieve equity in education and enhance the human capital development of poor students.

Some shortcomings also deserved to be considered by subsequent studies. First, for the measurement of poverty, we resorted to the specific indicator of BLA, due to the lack of key information on household income in the CEPS. However, the BLA and absolute poverty are not strictly consistent. Therefore, subsequent studies focusing on NIP and the human capital of poor households need to be more carefully defined. Second, in the mechanism analysis, we mainly emphasized some variables related to schools and students and may not have paid enough attention to the responses of parents and families. In fact, NIP implementations affected household welfare and economic behaviors as well, especially for poor families (Wang & Cheng, 2022). Third, since the CEPS focused on middle school students in China, their nutritional needs may differ from those of elementary or senior high school students. Therefore, the conclusions of this paper need further confirmation if extended to these groups.

CRediT authorship contribution statement

Chunkai Zhao: Conceptualization, Funding acquisition, Writing-original draft, Writing-reviewing & editing, Methodology, Project administration, Software. **Boou Chen:** Funding acquisition, Data curation, Writing-Original Draft, Visualization, Formal analysis, Resources. **Zhiyong Song:** Writing-reviewing & editing, Investigation, Validation, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

We appreciate the National Survey Research Center at Renmin University of China for providing us with the China Education Panel Survey (CEPS) data.

Funding: The work was supported by the National Natural Science Foundation of China (72303090; 72304102); the National Social Science Fund of China (21AJY009); Social Science Fund of Hunan Province (22ZDA023); Jiangsu Education Science Planning Project (C/2022/01/45); Philosophy and Social Science Program Discipline Co-construction Project of Guangdong Province (GD23XYJ68).

Appendix

Table A1. Variable definitions.

Variables	Definitions
CA	Standardized scores of students' CAs
NIP	= 1 if the school is a NIP pilot; 0 otherwise
Canteen	= 1 if there is a canteen in schools; 0 otherwise
School ranking	Current ranking of the school in the local county/district. 1 = near the bottom; 2 = below average; 3 = average; 4 = above average; 5 = among the best
Type	Public schools = 1; private schools = 0
Locations	1 = center of the city/town; 2 = outskirts of the city/town; 3 = rural–urban fringe zone of the city/town; 4 = towns outside of the city/town; 5 = rural areas

(continued on next page)

(*continued*)

Variables	Definitions
Gender	Boys = 1; girls = 0
Year of birth	Students' year of birth
Hukou	Rural = 1; urban = 0
Kindergarten	= 1 if the student attended kindergarten before starting primary school; 0 otherwise
Left-behind children	= 1 if at least one parent migrates; 0 otherwise
Migrant children	= 1 if students are from other counties; 0 otherwise
Only child	Only child = 1; non-only child = 0
Education of father	0 = unschooled; 1 = primary school; 2 = junior middle school; 3 = (vocationally) high school; 4 = college/bachelor's degree and above
Education of mother	0 = unschooled; 1 = primary school; 2 = junior middle school; 3 = (vocationally) high school; 4 = college/bachelor's degree and above
Job of father	= 1 if the father is a farmer, unemployed person, or freelancer; 0 otherwise
Ranking at 6th grade	= 1 if ranked in the top third; = 2 if ranked between one-third and two-thirds; = 3 if ranked in the bottom third
Difficulty in learning math	The difficulty of learning math in 6th grade. Not difficult at all = 1; not very difficult = 2; a bit difficult = 3; very difficult = 4
Difficulty in learning Chinese	The difficulty of learning Chinese in 6th grade. Not difficult at all = 1; not very difficult = 2; a bit difficult = 3; very difficult = 4
Grade repetition experience	= 1 if had grade repetition experience in elementary school; 0 otherwise
Learning ability	In 6th grade, how much do students agree with the statement that "I was fast learner"? Strongly disagree = 1; somewhat disagree = 2; somewhat agree = 3; strongly agree = 4
Learning attitude	In 6th grade, how much do students agree with the statement that "I would try my best to go to school even if I was not feeling very well"? Strongly disagree = 1; somewhat disagree = 2; somewhat agree = 3; strongly agree = 4
Study effort	In 6th grade, how much do students agree with the statement that "I would try my best to finish my homework, even if it would take me quite a long time"? Strongly disagree = 1; somewhat disagree = 2; somewhat agree = 3; strongly agree = 4
Curiosity	In 6th grade, how much do students agree with the statement that "I was curious about new stuff"? Strongly disagree = 1; somewhat disagree = 2; somewhat agree = 3; strongly agree = 4

Table A2. Variable definitions for mechanism Variables.

Variables	Definitions
Panel A. Health outcomes	
Self-rated health	Very poor = 1; not very good = 2; moderate = 3; good = 4; very good = 5
Frequent illness	Parents' comments. Yes = 1; no = 0
Height	cm
Weight	kg
Underweight	= 1 if girls and boys in 7th grade with BMI (kg/m^2) less than 14.7 and 15.4, respectively, and girls and boys in 8th grade with BMI less than 15.2 and 15.6, respectively; 0 otherwise.
Panel B. Sense of belonging at school	
Like the school	Somewhat like or like very much = 1; strongly dislike or somewhat dislike = 0
Like the homeroom teacher	Somewhat like or like very much = 1; strongly dislike or somewhat dislike = 0
Like the subject teachers	Somewhat like or like very much = 1; strongly dislike or somewhat dislike = 0
Like classmates	Somewhat like or like very much = 1; strongly dislike or somewhat dislike = 0
Panel C. Expectations for the future	
Self-educational expectations	Bachelor's degree and above = 1; junior middle school or (vocationally) high school = 0
Self-occupational expectations	= 1 if the future occupation is civil servant, manager, technician, and professional (e.g., lawyer and doctors); 0 otherwise (e.g., farmers)
Parents' educational expectations	Parents' responses. Bachelor's degree and above = 1; junior middle school or (vocationally) high school = 0
Parents' occupational expectations	= 1 if the future occupation is civil servant, manager, technician, and professional (e.g., lawyer and doctors); 0 otherwise (e.g., farmers)

References

Afridi, F. (2010). Child welfare programs and child nutrition: Evidence from a mandated school meal program in India. *Journal of Development Economics, 92*(2), 152–165.

Akin, J. S., Guilkey, D. K., & Popkin, B. M. (1983). The school lunch program and nutrient intake: A switching regression analysis. *American Journal of Agricultural Economics, 65*(3), 477–485.

Ananga, E. D. (2011). Typology of school dropout: The dimensions and dynamics of dropout in Ghana. *International Journal of Educational Development, 31*(4), 374–381.

Anderson, M. L., Gallagher, J., & Ramirez R., E. (2018). School meal quality and academic performance. *Journal of Public Economics, 168*, 81-93.

Aurino, E., Gelli, A., Adamba, C., Osei-Akoto, I., & Alderman, H. (2023). Food for thought? *Journal of Human Resources, 58*(1), 74–111.

Baumeister, R. F. (2012). *Need-to-belong theory*. Handbook of Theories of Social Psychology: SAGE Publications Ltd.

Bhattacharya, J., Currie, J., & Haider, S. J. (2006). Breakfast of champions? *Journal of Human Resources, XLI, 3*, 445–466.

Brady, D., & Bostic, A. (2015). Paradoxes of social policy: Welfare transfers, relative poverty, and redistribution preferences. *American Sociological Review, 80*(2), 268–298.

Brooks-Gunn, J., & Duncan, G. J. (1997). The effects of poverty on children. *Future of Children, 7*(2), 55–71.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173.

Bütikofer, A., Mølland, E., & Salvanes, K. G. (2018). Childhood nutrition and labor market outcomes: Evidence from a school breakfast program. *Journal of Public Economics, 168*, 62–80.

Chakraborty, T., & Jayaraman, R. (2019). School feeding and learning achievement: Evidence from India's midday meal program. *Journal of Development Economics, 139*, 249–265.

Chan, C. K. (2010). Re-thinking the incrementalist thesis in China: A reflection on the development of the minimum standard of living scheme in urban and rural areas. *Journal of Social Policy, 39*(4), 627–645.

Chen, B., & Zhao, C. (2022). More is less: Homeroom teachers' administrative duties and students' achievements in China. *Teaching and Teacher Education, 119*, Article 103857.

Chen, B., & Zhao, C. (2023). Teachers' administrative positions and students' non-cognitive abilities in China: Evidence from a quasi-natural experiment. *Applied Economics*. https://doi.org/10.1080/00036846.2023.2267819

Chen, K., Liu, C., Liu, X., Wang, Z., Luo, R., Li, S., Yu, Y., & Alderman, H. (2021). Nutrition, cognition, and social emotion among preschoolers in poor, rural areas of south Central China: Status and correlates. *Nutrients, 13*(4), 1322.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review, 110*(9), 2964–2996.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2018). On the relationship between cognitive ability and risk preference. *Journal of Economic Perspectives, 32*(2), 115–134.
Fang, G., & Zhu, Y. (2022). Long-term impacts of school nutrition: Evidence from China's school meal reform. *World Development, 153*, Article 105854.
Frisvold, D. E. (2015). Nutrition and cognitive achievement: An evaluation of the school breakfast program. *Journal of Public Economics, 124*, 91–104.
Foster, J. E. (1998). Absolute versus relative poverty. *American Economic Review, 88*(2), 335–341.
Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences, 14*(1), 47–64.
Gleason, P. M., & Suitor, C. W. (2003). Eating at school: How the National School Lunch Program affects children's diets. *American Journal of Agricultural Economics, 85*(4), 1047–1061.
Gong, J., Lu, Y., & Song, H. (2018). The effect of teacher gender on students' academic and noncognitive outcomes. *Journal of Labor Economics, 36*(3), 743–778.
Goodenow, C., & Grady, K. E. (1993). The relationship of school belonging and friends' values to academic motivation among urban adolescent students. *Journal of Experimental Education, 62*(1), 60–71.
Gordon, N. E., & Ruffini, K. J. (2018). *School nutrition and student discipline: Effects of schoolwide free meals* (p. w24986). No: National Bureau of Economic Research.
Gundersen, C., Kreider, B., & Pepper, J. (2012). The impact of the National School Lunch Program on child health: A nonparametric bounds analysis. *Journal of Econometrics, 166*(1), 79–91.
Hatløy, A., Torheim, L., & Oshaug, A. (1998). Food variety—a good indicator of nutritional adequacy of the diet? a case study from an urban area in Mali, West Africa. *European Journal of Clinical Nutrition, 52*(12), 891–898.
Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*(3), 411–482.
Hinrichs, P. (2010). The effects of the National School Lunch Program on education and health. *Journal of Policy Analysis and Management, 29*(3), 479–505.
Huang, L. (2022). Exploring the relationship between school bullying and academic performance: The mediating role of students' sense of belonging at school. *Educational Studies, 48*(2), 216–232.
Imberman, S. A., & Kugler, A. D. (2014). The effect of providing breakfast in class on student performance. *Journal of Policy Analysis and Management, 33*(3), 669–699.
Immink, M. D. (2010). Nutrition, poverty alleviation, and development in Central America and Panama. *Food and Nutrition Bulletin, 31*(1), 161–172.
Jacka, F. N., Maes, M., Pasco, J. A., Williams, L. J., & Berk, M. (2012). Nutrient intakes and the common mental disorders in women. *Journal of Affective Disorders, 141*(1), 79–85.
Joe, W., Rajaram, R., & Subramanian, S. V. (2016). Understanding the null-to-small association between increased macroeconomic growth and reducing child undernutrition in India: Role of development expenditures and poverty alleviation. *Maternal & Child Nutrition, 12*, 196–209.
Jomaa, L. H., McDonnell, E., & Probart, C. (2011). School feeding programs in developing countries: Impacts on children's health and educational outcomes. *Nutrition Reviews, 69*(2), 83–98.
Joo, M. (2011). Effects of federal programs on children: Absolute poverty, relative poverty, and income inequality. *Children and Youth Services Review, 33*(7), 1203–1211.
Li, C, M., Feldman, S. S., & Rosenthal, D. A. (1992). The influence of immigration on parental behavior and adolescent distress in Chinese families residing in two Western nations. *Journal of Research on Adolescence, 2*(3), 205-239.
Li, P., Lu, Y., & Wang, J. (2016). Does flattening government improve economic performance? evidence from China. *Journal of Development Economics, 123*, 18–37.
Li, S., Chen, K., Liu, C., Bi, J., He, Z., Luo, R., Yu, Y., & Wang, Z. (2021). Association of dietary diversity and cognition in preschoolers in rural China. *Nutrition, 91–92*, Article 111470.
Li, Y., Yang, H., & Luo, L. (2021). Poverty exposure and cognitive abilities of children in rural China: Causation and the roles of family investments. *Children and Youth Services Review, 121*, Article 105747.
Liu, T., Zhang, X., & Jiang, Y. (2020). Family socioeconomic status and the cognitive competence of very young children from migrant and non-migrant chinese families: The mediating role of parenting self-efficacy and parental involvement. *Early Childhood Research Quarterly, 51*, 229–241.
Liu, X., Zhao, Q., & Chen, Q. (2019). Better nutrition, healthier mind? experimental evidence from primary schools in rural northwestern China. *Journal of Integrative Agriculture, 18*(8), 1768–1779.
Liu, Y., Liu, J., & Zhou, Y. (2017). Spatio-temporal patterns of rural poverty in China and targeted poverty alleviation strategies. *Journal of Rural Studies, 52*, 66–75.
Long, S. K. (1991). Do the school nutrition programs supplement household food expenditures? *Journal of Human Resources, 26*(4), 654.
Lu, Y., Wang, J., & Zhu, L. (2019). Place-based policies, creation, and agglomeration economies: Evidence from China's economic zone program. *American Economic Journal: Economic Policy, 11*(3), 325–360.
Lundborg, P., Rooth, D. O., & Alex-Petersen, J. (2022). Long-term effects of childhood nutrition: Evidence from a school lunch reform. *Review of Economic Studies, 89*(2), 876–908.
Luo, R., Shi, Y., Zhang, L., Liu, C., Rozelle, S., Sharbono, B., Yue, A., Zhao, Q., & Martorell, R. (2012). Nutrition and educational performance in rural China's elementary schools: Results of a randomized control trial in Shaanxi Province. *Economic Development and Cultural Change, 60*(4), 735–772.
McEwan, P. J. (2013). The impact of Chile's school feeding program on education outcomes. *Economics of Education Review, 32*, 122–139.
Millimet, D. L., Tchernis, R., & Husain, M. (2010). School nutrition programs and the incidence of childhood obesity. *Journal of Human Resources, 45*(3), 640–654.
Morris, M. C., Marco, M., Maguire-Jack, K., Kouros, C. D., Im, W., White, C., & Garber, J. (2019). County-level socioeconomic and crime risk factors for substantiated child abuse and neglect. *Child Abuse & Neglect, 90*, 127–138.
Mostert, C. M. (2021). The impact of the school feeding programme on the education and health outcomes of south african children. *Children and Youth Services Review, 126*, Article 106029.
Nelson, M. (2000). Childhood nutrition and poverty. *Proceedings of the Nutrition Society, 59*(2), 307–315.
Neumann, C. G., Murphy, S. P., Gewa, C., Grillenberger, M., & Bwibo, N. O. (2007). Meat supplementation improves growth, cognitive, and behavioral outcomes in kenyan children. *Journal of Nutrition, 137*(4), 1119–1123.
Nong, H., Zhang, Q., Zhu, H., & Zhu, R. (2022). Targeted poverty alleviation and children's academic performance in China. *Review of Income and Wealth, 68*(4), 951–969.
Pierce, B. J., Muzzey, F. K., Bloomquist, K. R., & Imburgia, T. M. (2022). Effectiveness of family centered treatment on reunification and days in care: Propensity score matched sample from Indiana child welfare data. *Children and Youth Services Review, 136*, Article 106395.
Qiu, H., & Liang, X. (2024). Segmented assimilation of migrant children in urban public schools of China: Impact of family resources and school segregation on their educational aspirations. *Children and Youth Services Review, 156*, Article 107300.
Raufelder, D., Neumann, N., Domin, M., Lorenz, R. C., Gleich, T., Golde, S., Romund, L., Beck, A., & Hoferichter, F. (2021). Do belonging and social exclusion at school affect structural brain development during adolescence? *Child Development, 92*(6), 2213–2223.
Ravallion, M., & Chen, S. (2011). Weakly relative poverty. *Review of Economics and Statistics, 93*(4), 1251–1261.
Ren, J., Zheng, X., Smith, R., & Fang, X. (2023). School feeding program and urban-rural inequality of child health: Evidence from China. *Agribusiness*. https://doi.org/10.1002/agr.21862
Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74*(2), 166.
Slaughter, M. J. (2001). Trade liberalization and per capita income convergence: A difference-in-differences analysis. *Journal of International Economics, 55*(1), 203–228.
Tani, M., Xu, L., & Zhu, Y. (2021). The impact of an un (der) funded inclusive education policy: Evidence from the 2013 China education panel survey. *Journal of Economic Behavior & Organization, 190*, 768–784.
Thompson, I., McNicholl, J., & Menter, I. (2016). Student teachers' perceptions of poverty and educational achievement. *Oxford Review of Education, 42*(2), 214–229.
Turney, K., Testa, A., & Jackson, D. B. (2023). Anticipated unmet educational expectations following youth police contact. *Children and Youth Services Review, 155*, Article 107197.
Valdivia, M. (2004). Poverty, health infrastructure and the nutrition of peruvian children. *Economics & Human Biology, 2*(3), 489–510.
Vittadini, G., Sturaro, C., & Folloni, G. (2021). Non-cognitive skills and cognitive skills to measure school efficiency. *Socio-Economic Planning Sciences, 81*, Article 101058.
Wang, D., Stewart, D., Chang, C., & Shi, Y. (2015). Effect of a school-based nutrition education program on adolescents' nutrition-related knowledge, attitudes and behaviour in rural areas of China. *Environmental Health and Preventive Medicine, 20*(4), 271–278.
Wang, H., & Cheng, Z. (2022). Kids eat free: School feeding and family spending on education. *Journal of Economic Behavior & Organization, 193*, 196–212.
Wang, H., Zhao, Q., Bai, Y., Zhang, L., & Yu, X. (2020). Poverty and subjective poverty in rural China. *Social Indicators Research, 150*(1), 219–242.
Wang, J., Hernandez, M. A., & Deng, G. (2023). Large-scale school meal programs and student health: Evidence from rural China. *China Economic Review, 79*, Article 101974.
Wang, J., Zhou, L., & Yao, S. (2022). The impact of the nutrition improvement program on children's health in rural areas: Evidence from China. *Emerging Markets Finance and Trade, 58*(1), 267–289.
World Food Programme. (2020). *State of school feeding worldwide 2020*. Rome: World Food Programme.
Zhang, F., Hu, X., Tian, Z., Zhang, Q., & Ma, G. (2015). Literature research of the nutrition improvement programme for rural compulsory education students in China. *Public Health Nutrition, 18*(5), 936–943.
Zhao, C., & Chen, B. (2022). Parental migration and non-cognitive abilities of left-behind children in rural China: Causal effects by an instrumental variable approach. *Child Abuse & Neglect, 123*, Article 105389.
Zhao, C., & Chen, B. (2023). ICT in education can improve students' achievements in rural China: The role of parents, educators and authorities. *Journal of Policy Modeling, 45*(2), 320–344.
Zhao, C., Chen, B., & Li, X. (2023). Rising housing prices and marriage delays in China: Evidence from the urban land transaction policy. *Cities, 135*, Article 104214.
Zhao, C., & Guo, J. (2022). Are veterans happy? long-term military service and the life satisfaction of elderly individuals in China. *Journal of Happiness Studies, 23*(2), 477–508.

Zhao, C., Wu, Y., & Guo, J. (2022). Mobile payment and chinese rural household consumption. *China Economic Review, 71*, Article 101719.
Zheng, X., Ren, J., Chen, D., & Fang, X. (2023). School feeding and children's noncognitive skills: Evidence from the nutrition improvement program in rural China. *Applied Economics, 55*, 6459–6478.
Zhou, L., Jiang, B., & Wang, J. (2020). Do cash transfers have impacts on student academic, cognitive, and enrollment outcomes? evidence from rural China. *Children and Youth Services Review, 116*, Article 105158.
Zhou, Y., Huang, X., Shen, Y., & Tian, L. (2023). Does targeted poverty alleviation policy lead to happy life? evidence from rural China. *China Economic Review, 81*, Article 102037.

北京大学
中国
PEKING
UNIVERSITY
CHINA

经济科学

Economic Science

2023 | Number 2

ISSN 1002-5839
9 771002 583235 04>

2023年
第2期
(总第254期)

目　录

基于人民幸福感的共同富裕指标体系构建及测度
…… 万广华　蓝　菁　刘　震（5）
中国居民绝对与相对贫困的综合测度：1989—2015
…… 陈铭聪　程振源　杨彦欣（26）
地方政府债务与宏观经济波动
——地方政府激励与央地互动的视角
…… 夏　芊　赵扶扬（49）
国税地税征管体制改革、信息获取与企业避税
——基于国税地税合作的证据
…… 张明昂　陈斌开　岳林峰（71）
城市金融发展与外商直接投资
——基于古寺院金融的研究视角
…… 黄新飞　王升泉　杜灵珊　叶梓南（89）
信息基础设施建设与企业投资
——基于“宽带中国”试点政策的准自然实验
…… 孔东民　陶云清（106）
加工贸易转移如何影响承接地区企业利润率
——基于时期差异视角
…… 胡浩然　施炳展　宋颜群（125）
制度质量如何影响外资企业进入模式
——来自中国设立开发区的经验证据
…… 吴　敏　黄玖立　丁思琪（145）
价格规制降低药品招标价格了吗？
——基于基本药物制度和取消药品加成政策准自然实验的研究
…… 岳　阳　龙小宁　朱恒鹏（163）
法定育儿假补贴、人力资本投资与经济增长
…… 赵美洁　严成樑（183）
抵御风险视角下农户的借贷和赊账行为
…… 马　祥　李承政（198）
宗族组织、信贷约束与农村人力资本积累
——基于“大学扩招”的实证研究
…… 梁若冰　任伟聪（213）
来自军营的印记
——军人职业与老年人健康
…… 赵纯凯　湛　泳　陈华帅（227）

 DOI: 10.12088/PKU.jjkx.2023.02.13

来自军营的印记*
——军人职业与老年人健康

赵纯凯　湛　泳　陈华帅

摘　要：本文基于2002—2014年中国老年健康长寿影响因素调查（CLHLS）数据，采用带有工具变量的条件混合过程（CMP）估计方法，首次考察了军人职业对中国男性老年人健康的影响。研究发现，军人职业显著提升了老年人的躯体健康、认知能力以及自评健康，这一结论得到一系列稳健性检验的支持。军人职业改善老年人健康的机制包括健康行为、教育和收入溢价以及社会保障。同时，本文发现军人职业降低了老年人患癌症的风险，但对心理健康的影响不大。此外，短期从军经历同样有助于改善老年人健康。

关键词：军人职业　健康　退役军人　社会保障　条件混合过程

中图分类号：F240　　**JEL分类号**：I12　J14

一、引　言

长期以来，中国政府高度重视退役军人健康和社会保障问题。十九大报告指出“要维护军人军属合法权益，让军人成为全社会尊崇的职业。”2018年3月，中华人民共和国退役军人事务部正式成立，并出台了一系列政策文件，要求省市县各级均建立退役军人事务管理与服务机构。2021年，《中华人民共和国退役军人保障法》和《中华人民共和国军人地位和权益保障法》正式实施，进一步完善了中国退役军人管理、抚恤和社会保障制度。根据中国民政部的统计，截至2017年，中国60岁以上的老年退役军人达到411.7万，需要优抚和优待的退役军人、伤残军人以及烈属遗属高达858.5万。在从中央到地方都热切关注退役军人健康的背景下，本文探究军人职业对老年人健康的影响具有重要的现实意义。

生命历程理论认为生命历程中特定的事件、经历和环境对个体健康具有长期性的影响（Spiro等，2016）。对老年退役军人而言，军人职业作为生命历程中的一段特殊经历，也可能影响其当前的健康状况（Angrist和Chen，2011；Sheehan和Hayward，2019）。一方面，长期的军事训练和部队生活所留下的“军营印记”，能通过改变退役军人的性格特征、生活习惯和价值观念来影响健康状况（Benmelech和Frydman，2015；Wilmoth等，2010）。另一方面，军人职业经历还可能影响个体教育和收入，带来更好的就业机会

* 赵纯凯（通信作者），华南农业大学经济管理学院，E-mail：ckz_zhao@163.com；湛泳，湘潭大学商学院，E-mail：zhanyong225@163.com；陈华帅，湘潭大学商学院，E-mail：huashuai.chen@gmail.com。本文为国家社会科学基金重点项目“新时代中国退役军人社会保障制度体系优化研究”（项目编号：21AJY009）和湖南省教育厅重点项目“新时代退役军人社会保障制度研究”（项目编号：19A481）的阶段性成果。本文的前期版本曾在第十九届中国经济学年会（劳动经济学分会场）、第9期香樟经济学Seminar（杭州）会议和第三届健康经济发展论坛上汇报过，感谢与会专家和学者给本文的建设性意见。作者感谢审稿专家和编辑部在本文写作过程中提出的宝贵意见，文责自负。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

(Zhang，2015；王兵和杨宝，2019)，有助于在老年阶段获得更好的医疗保障和社会福利(Angrist，1998；Waller 和 McGuire，2011)，这同样可能改善健康状况。

即使如此，关于军人职业对健康的影响，现有文献并没有得到一致的结论。一些研究发现由于“健康士兵效应”(healthy soldier effect)的存在(McLaughlin 等，2008；Waller 和 McGuire，2011)，参军能提升健康水平，并降低老年人的死亡率(Sheehan 和 Hayward，2019)。然而，一些基于参战退役军人的研究却得出了相反的结论，认为军人职业会降低个体健康状况(Bedard 和 Deschênes，2006；Teachman，2011)。一些心理学和精神病学的研究也发现参战退役军人面临更多的心理疾病(Hoge 等，2004)。还有文献特别关注了老年退役军人的身心健康。一方面，参军会带来较大的职业心理压力(Hoge 等，2004)并面临更多的风险暴露(Johnston 等，2016)，所造成的损伤会形成劣势积累效应，导致老年退役军人的患病率更高(Johnston 等，2016；Williamson 等，2019)。另一方面，也有部分研究显示一些社会弱势群体，如少数族裔、女性、贫困家庭和农村个体，能够通过参军获得相对更高的社会经济地位(Hou 等，2020；Padavic 和 Prokos，2017；Sheehan 和 Hayward，2019)，这能够帮助其在老年阶段获得更大的健康溢出效应(Landes 等，2017；Sheehan 和 Hayward，2019)。此外，也有少量文献关注了中国退役军人福利，不过都以农村退役军人为研究对象，研究发现参军对农村青年就业和收入存在积极影响(王兵和杨宝，2019；汪建华，2011；Hou 等，2020；Zhang，2015)。

从现有文献来看，关于军人职业或服役经历对健康影响的研究结论是存在差异的，且大部分研究着眼于欧美发达国家。而基于中国等发展中国家所展开的研究主要关注退役军人的就业和收入状况，较少评估军人职业对老年人健康的影响。显然，无论是在战争历史、兵役制度和文化背景方面，还是在老年退役军人的生活环境、社会保障、经济待遇和生活习性方面，中国等发展中国家与发达国家都存在较大的差异。随着退役军人事务部正式成立和一系列政策和法规的落地，迫切需要对老年退役军人的健康和医疗需求状况进行评估。然而，目前关于中国老年退役军人健康方面的经验研究还很欠缺。本文运用中国老年健康长寿影响因素调查(Chinese Longitudinal Healthy Longevity Survey，CLHLS)数据，试图进一步将军人职业对老年人健康的影响这一研究领域拓展到发展中国家，以便为政策制定者提供微观经验支撑。

本文首次利用微观数据探究军人职业对中国男性老年人健康的影响，并可能在以下三个方面有所创新。首先，从躯体健康、认知能力和自评健康等多个角度，本文全面检验了军人职业经历所留下的“军营印记”对中国男性老年人健康的影响，进一步拓宽现有研究的范围和填补该领域在发展中国家的研究空白。其次，考虑到参军决策的自选择性和部队招募士兵的筛选行为，从军和老年人健康之间可能存在内生性问题。本文通过采取带有工具变量的条件混合过程(CMP)估计法进行处理，这为识别军人职业等特殊职业经历的内生性问题提供了参考。最后，本文从健康行为、教育和收入溢价以及社会保障等多个层面检验了军人职业影响老年人健康的机制，这有助于理解“军营印记”对老年人生命质量和社会福祉的具体影响。

二、文献综述和理论机制

尽管现有文献关于军人职业或军事服役经历对个体健康的影响并未形成统一的观点，但是学者们依旧从多个方面探究了从军经历影响个体健康的潜在机制。根据这些研究，本文从人力资本和收入、健康行为、政治资本和社会网络以及社会保障等方面来进行文

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

献梳理。

参军的人力资本效应是学者们重点关注的一个领域。有证据表明，得益于对退役军人实施的一系列教育培训政策，在越南战争时期的美国退役军人能够获得更多接受高等教育的机会（Angrist 和 Chen，2011；Bound 和 Turner，2002）。同时，军队本身能提供一些低成本的军事教育并创造有利于学习的环境，这能够提高退役军人的认知能力（Lindqvist 和 Vestman，2011）。同时，军队强调权威、职责、团队合作、不畏风险和自律自强，这也可能会提升退役军人的非认知技能（Benmelech 和 Frydman，2015）。鉴于参军对人力资本积累的正面影响，学者们进一步评估了参军所带来的收入溢价。尽管关于第二次世界大战和越南战争期间美国退役军人是否存在收入溢价的研究结论有分歧（Angrist 和 Krueger，1994；Card 和 Cardoso，2012），但是在参军帮助社会弱势群体获得更多收入的观点上达成一致（Hirsch 和 Mehay，2003；Padavic 和 Prokos，2017）。一些基于中国农村样本的研究也同样发现参军能够带来较高收入溢价（Hou 等，2020；Zhang，2015；王兵和杨宝，2019）。显然，人力资本积累和收入溢价优势对提升退役军人的健康状况至关重要（Sheehan 和 Hayward，2019）。

还有一些文献关注退役军人的健康行为。一方面，为对抗军事职业的高风险，军人可能通过吸烟和喝酒等不健康行为来舒缓压力。尤其是面临战争威胁时，军人更容易形成对香烟、酒精和药物的依赖（Bedard 和 Deschênes，2006）。同时，军旅集体生活有很强的外部性，很难让军人保持独立的生活习惯而远离这些不健康行为（Landes 等，2017）。另一方面，军人职业经历可能形成更好的锻炼习惯来提高身体素质（McLaughlin 等，2008），也更加注重营养搭配和传授营养知识。更重要的是，军队中的身体训练更加科学和有效，更能够锻炼个人的意志品质。

此外，退役军人的政治资本和社会网络也得到了一些学者的重视。已有文献强调了对军人进行政治培训和促使军人形成政治意识的重要性。通过政治军事学习，军人能够更清晰了解所承担的政治责任，增强爱国主义精神（Leal 和 Teigen，2018）。因此，参军有助于提升个体政治积极性，促进政治资本的积累（Leal 和 Teigen，2018）。而在社会网络方面，退役军人可能与本地战友、军官和其他退伍军人建立联系并形成社会网络（Hou 等，2020；王兵和杨宝，2019）。部队长期的集体生活以及共同训练和工作的经历，使得战友的情感联结更加紧密。即使非战友，基于退役军人的“身份认同效应”，军人社会网络也可能更加牢固（Spiro 等，2016）。政治资本和社会网络在退役军人就业和社会互动等方面的优势可能在改善老年退役军人的健康中发挥作用（McLaughlin 等，2008；Wilmoth 等，2010）。

最后，由于军人社会保障涉及国家军事人力资本的稳定、军队战斗力甚至是国家安全，所以世界各国都非常重视退役军人的社会保障。一般而言，除常规的社会保障方式外，政府还会提供一些特殊的社会保障方式来对退役军人进行补偿，如就业安置、社会抚恤、军属优待和特殊津贴等（Angrist，1998；Landes 等，2017）。在中国，义务兵退役后可以享受退役金和地方政府经济补助，还可以参与职业教育和技能培训课程；军官退役后可以享受安置、住房、医疗、荣誉和社会保险等各个方面的保障。① 农村老兵还额

① 参见退役军人事务部、中央宣传部、国家发展改革委等 20 部门联合印发的《关于加强军人军属、退役军人和其他优抚对象优待工作的意见》。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

外享有生活补助、养老补贴和养老保险军龄优待。这些特殊保障和优待在提升退役军人健康状况上起到了重要的作用。

三、研究设计

（一）数据来源

本文所使用的数据来自中国老年健康长寿影响因素调查。该调查是由北京大学健康老龄与发展研究中心、美国杜克大学老龄发展研究中心以及国家疾控中心共同主持实施，从1998年开始进行8次跟踪调查。项目调查范围包括全国23个省区市，调查内容涵盖中国老年人的健康状况、健康行为、职业背景以及经济条件等方面，是目前研究中国老龄化和老年人健康问题的重要大型微观数据库。

本文选取2002—2014年5期的混合数据①，同时对数据进行了如下筛选：第一，按照现有文献的做法，考虑到军人主要以男性为主，剔除了女性样本（Hou等，2020；Johnston等，2016）；第二，剔除了在调查时尚未年满60岁的样本；第三，剔除核心变量有缺漏的样本。通过对数据的清理和筛选，最终保留了22 027个观测值。

（二）变量说明

本文的被解释变量为老年人健康状况，包括躯体健康、认知能力和自评健康三个核心变量。其中，躯体健康采用老年人工具性日常活动能力（*IADL*）来衡量（程令国等，2015；Wilmoth等，2010），其依据是观测老年人能否完成做饭、洗衣、外出串门、购物、走远路、提重物、下蹲、乘公交车等基本活动。如果上述活动均能自理，则说明躯体健康未受损，赋值为1；至少有一项需借助他人帮助才能完成则表示躯体健康受损，赋值为0。认知功能是根据国际通行的认知功能简易量表（MMSE）的测量结果，测量内容包括方向定位能力、反应能力、注意力、计算能力以及语言和自我协调能力等，总分值为30分。根据Shen和Zeng（2014）的做法，本文将24分及以上定义为认知能力完好，赋值为1；低于24分则赋值为0。此外，本文还选取了自评健康这一主观指标来综合反映老年人的健康状况，其根据问卷中“您觉得现在您自己的健康状况怎么样”的回答进行赋值。本文将回答为“好”和“很好”的赋值为1，将回答为“一般”、“不好”和“很不好”的赋值为0。

本文核心解释变量为老年人是否有军人职业经历，根据调查问卷中的“您60岁以前主要从事什么工作（职业）”的回答来进行赋值。将回答为“军人”的赋值为1，其他赋值为0。需要说明的是，CLHLS中的职业调查是指早年从事的工作超过20年，所以问卷中所指的是一种将军人作为主要职业的长期服役经历。②

本文还选取了一系列外生控制变量。在老年人的个人特征上，所选取的变量包括年龄、是否有配偶、民族、童年是否常挨饿、童年生病是否能及时治疗和出生次序。家庭特征变量包括父亲是否务农、父亲是否受过教育、母亲是否受过教育、兄弟姐妹数、在

① 由于1998年和2000年周期的数据主要调查80岁以上的高龄老年人，所以本文选取2002年以后调查的数据。

② 根据CLHLS问卷，本文所识别的是以军人职业为主的长期从军经历对老年人健康的影响，并不包括一些短期参军服役的退役军人。不过，在后面的分析中，本文将根据中国家庭追踪调查（CFPS）来分析短期参军对老年人健康的影响。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

世儿子数以及在世女儿数。本文控制的一系列童年特征变量和家庭特征变量很可能与个体的参军决策以及老年后的身体状况相关（Shen 和 Zeng，2014）。因此，加入这些变量有助于降低内生性。

（三）实证策略

为了准确识别军人职业影响老年人健康的因果效应，需要对模型中可能存在的内生性问题加以考虑。一是反向因果问题。本文使用的是老年人样本，而参军行为和职业经历发生在过去，显然老人当期的健康状况无法影响几十年前的从军行为。因此，反向因果问题不需要特别关注。二是自选择问题。考虑到个体从军决策与身体素质、家庭经济状况和性格特征相关，参军有可能是一种自选择行为。例如，早年家庭经济状况较差或偏好风险的个体，从军的意愿可能更高，更倾向于通过从军来获得向上跃迁的机会（Wu 和 Treiman，2004；Hou 等，2020）。而早年生活条件以及性格特征都与个体健康状况息息相关。三是部队选择行为。由于军人存在严格的筛选机制，身体素质越好的个体越容易通过征兵体检，这可能导致老年退役军人的健康状况不能被真实反映。

针对可能存在的内生性问题，本文主要采取工具变量法来进行处理。参照 Bedard 和 Deschênes（2006）、Hou 等（2020）以及 Zhao 和 Guo（2022）的做法，本文所选取的工具变量为基于出生队列的参军比例，从出生年份层面上捕捉军队规模的外生变化。参军比例的上升可能意味着国家对士兵需求量的增加和征兵强度的提升，个体进入部队的可能性会相应提高。相反，参军比例下降可能意味着处于裁减军事武装人员时期，国家对士兵的需求量下降，个体参军的成功率也随之降低。所以，参军比例的变动在一定程度上能够影响个体的参军行为，满足工具变量相关性条件。此外，参军比例的变动取决于兵力需求和国家国防战略，不受个人行为和意志的干扰，且基于出生队列的参军比例是一个历史变量，难以对老人当前的健康水平产生直接影响，所以该工具变量具备理论上的可行性。为了保障工具变量来源的外生性，本文还从另一个老年人微观数据库，中国健康与养老追踪调查（CHARLS）来获取工具变量。CHARLS 在 2014 年所展开的生命历程调查中记录了受访者的参军信息。

同时，由于所选取的工具变量是出生队列层面的，为了缓解未观测到的队列异质性对实证结果的干扰，本文控制了五年期的出生队列固定效应，并采取出生队列层面的聚类标准误进行矫正（Hou 等，2020）。此外，考虑到本文的解释变量和被解释变量均为二元变量，因此采用带有工具变量的条件混合过程（CMP）估计方法。模型设定如下：

$$Health^{*}_{itc} = \alpha_0 + \alpha_1 Military_{ic} + \alpha_2 X_{itc} + \varphi(c) + \lambda_p + \theta_t + \mu_{itc} \quad (1)$$

$$Health_{itc} = \begin{cases} 1, & \text{if} \quad Health^{*}_{itc} > 0; \\ 0, & \text{otherwise} \end{cases}$$

$$Military^{*}_{ic} = \beta_0 + \beta_1 Z_c + \beta_2 X_{itc} + \varphi(c) + \lambda_p + \theta_t + \varepsilon_{itc} \quad (2)$$

$$Military_{ic} = \begin{cases} 1, & \text{if} \quad Military^{*}_{ic} > 0; \\ 0, & \text{otherwise} \end{cases}$$

其中，下标 i、c、t、p 分别表示个体、出生队列、调查年份和省份。$Health^{*}_{itc}$ 是老年人的健康状况的潜变量，包括前文提到的躯体健康、认知能力和自评健康。$Military_{ic}$ 指代个体是否有长期的军人职业经历，$Military^{*}_{ic}$ 是个体军人职业倾向的潜变量。X_{itc} 是一系列外生控制变量。$\varphi(c)$、λ_p 和 θ_t 分别表示五年期出生队列固定效应、省份固定效应和年份

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

固定效应。Z_c 是工具变量。α 和 β 为两组待估参数。μ_{itc} 和 ε_{itc} 是随机误差项。

（四）描述性统计①

从描述性统计可以看出，样本中存在军人职业的老年人占比为 1.39%，略低于之前基于青年样本的文献（Hou 等，2020；王兵和杨宝，2019）。平均而言，44.37% 的老年人躯体健康未受损，70.96% 的认知能力完好，46.85% 的认为自己的健康状况较好。另外，从均值差异检验结果可以看到，无论是健康躯体、认知能力还是自评健康，老年退役军人的健康状况都相对更好。平均而言，老年退役军人躯体健康未受损的概率比非退役军人要高 7.04%，认知能力完好的占比要多 5.92%，自评健康的比例要高 5.51%。描述性分析表明，军人职业很可能有利于老年人健康。此外，工具变量的描述性统计结果显示，平均而言，退役军人所在出生队列的平均参军比例比非退役军人高 1.80%，初步说明工具变量是满足相关性假设的。

四、实证结果及分析

（一）基准结果

在基准结果中，本文检验了军人职业对老年人健康的影响。因为老年人健康均为二元变量，所以本文首先采用了普通的 Probit 模型来进行估计，结果如表 1 的前三列所示。在控制一系列个体和家庭特征变量以及固定效应之后，边际效应结果显示无论是客观的躯体健康和认知能力还是主观的自评健康，核心解释变量的系数都显著为正。不过，尽管 Probit 估计结果显示军人职业可能促进老年人健康，但是依旧可能受到潜在内生性问题的干扰。

表 1 的后四列报告了运用带有工具变量的 CMP 方法的估计结果。第（4）列的第一阶段结果显示工具变量与老年人的军人职业经历显著正相关。从最后三列的结果来看，军人职业经历所留下的“军营印记”确实显著促进了老年人健康。平均而言，军人职业

表 1　军人职业对老年人健康的影响

	Probit 模型			CMP 方法			
	躯体健康	认知能力	自评健康	军人职业	躯体健康	认知能力	自评健康
	（1）	（2）	（3）	（4）	（5）	（6）	（7）
军人职业	0.1167***	0.0829***	0.0622**		0.3309***	0.3140***	0.4215**
	（0.0271）	（0.0294）	（0.0304）		（0.1258）	（0.1029）	（0.1802）
平均参军比例				0.0737***			
				（0.0231）			
控制变量	是	是	是	是	是	是	是
五年期出生队列固定效应	是	是	是	是	是	是	是
省份固定效应	是	是	是	是	是	是	是
时间固定效应	是	是	是	是	是	是	是
观测值	22 027	22 027	22 027	22 027	22 027	22 027	22 027

注：***、**、* 分别表示在 1%、5%、10% 的水平下显著，报告的均为边际效应，括号内为出生队列层面的聚类标准误；后同。

① 因篇幅所限，本文省略了描述性统计，感兴趣的读者可在《经济科学》官网论文页面“附录与扩展”栏目下载。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

使老年人躯体健康、认知能力完好以及自评健康的概率分别增加了 33.09、31.40 和 42.15 个百分点。这些结果相较于 Probit 估计在数值上更大，可能是受到工具变量法的局部平均处理效应的影响（Angrist 和 Imbens，1995；程令国等，2015），但并不影响因果效应的识别。

（二）稳健性检验①

（1）两阶段最小二乘估计。在基准回归结果中，本文使用 CMP 估计方法得到了军人职业与中国老年人健康之间的因果关系，但是 CMP 方法无法得到弱工具变量判断的常用统计量，因此本文进一步使用两阶段最小二乘法（2SLS）来进行估计。从结果中不难看到，第一阶段 F 值为 15.2982，大于通常认为的弱工具变量的门槛值 10，同时 Anderson-Rubin 检验也都在 5%的水平上显著，故可以认为并不存在弱工具变量的情形。从估计结果来看，核心解释变量的系数在各列中均显著为正，说明替换估计方法后基准回归的结果保持稳健。

（2）考虑战争因素。由于样本中可能存在一些经历过战争的老年人，而早年的战争或参战经历很可能直接影响老年人的健康状况（Bedard 和 Deschênes，2006），因此本文需要剔除战争因素对基准实证结果的影响。本文首先考虑中华人民共和国成立后的两次重要战役，抗美援朝和中印（边境）自卫反击战。本文将 18—25 岁视为可能适合参战的年龄，并基于老年人的出生队列剔除对应的样本（Zhao 和 Guo，2022）。② 可以看出，军人职业的系数在各列中都显著为正，说明剔除这些潜在的参战样本对本文结果的影响较小。其次，考虑到中华人民共和国成立前的两次全面战争，抗日战争和解放战争，持续时间较长，且这两场大规模的全面战争波及全国一半以上的疆域，不仅会影响参战士兵，而且会对普通民众产生负面冲击。因此，本文根据出生队列分别计算样本中的老年人经历这两次战争的年数③，并分别控制经历战争年数的固定效应来尽可能降低全面战争对实证结果的影响。可以看出，估计结果没有发生太大的变化。

（3）剔除极端观测值。在前文的分析中，为避免样本中只存在调查期仍健在的老年人而引起的样本选择偏误问题，基准回归中同时包括在世和已经去世的老年人样本。但考虑到死亡这一极端健康表现，本文进一步剔除了在调查期（2014 年以前）已逝世的样本。从前三列的实证结果可以看出，各列的核心解释变量依旧为正，说明本文在剔除健康状况极差（最终死亡）的老年人样本后，实证结果依旧保持稳健。此外，考虑到离休是针对中华人民共和国成立以前参加革命工作干部的一项特殊保障制度，这些老年人在离休后可能享受到更多社会福利和社会保障，因此本文进一步剔除这些特殊样本。本文根据问卷中老年人是否享受离休待遇来进行识别，发现共有 525 个老年人享受离休待遇。在后三列的回归结果中，CMP 方法所报告的边际效应结果显示军人职业依旧能够显著促进老年人健康。

（4）加入其他控制变量。虽然本文控制了省级固定效应，但是一些宏观变量也可能

① 稳健性检验结果请见《经济科学》官网“附录与扩展”。

② 抗美援朝战争发生在 1950—1953 年，相应地本文剔除了出生于 1925—1935 年的样本。而中印边境自卫反击战发生在 1962 年，本文对应剔除了出生于 1937—1944 年的老年人。相应的回归结果请见《经济科学》官网“附录与扩展”中表 A3。

③ 抗日战争的时间为 1931—1945 年，解放战争的时间为 1946—1950 年。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

会影响老年人的健康状况。这些宏观变量包括人均 GDP、公共服务①、每万人医疗机构数、每百人病床数和每百人医疗人员数，数据均来自《中国统计年鉴》。在检验结果前三列中，核心解释变量的系数保持显著为正，说明加入这些宏观控制变量后本文的实证结果依旧稳健。此外，考虑到中国幅员辽阔，不同地区的队列异质性可能较大，因此本文进一步控制了五年期出生队列与省份的固定效应，结果如检验结果后三列所示。显然，控制高维度固定效应后的估计结果与表 1 中的基准结果相似。

（三）拓展分析②

（1）军人职业对老年人慢性疾病的影响。一些文献发现参军可能增加老年人患心脏病、肺结核和高血压等慢性疾病的概率（Bedard 和 Deschênes，2006；Johnston 等，2016），但并未具体分析老年退役军人的疾病状况。根据 CLHLS 数据库，本文选取了几种常见慢性病，包括心脏病、高血压、糖尿病、肺结核和癌症。可以看出，平均而言，老年退役军人患癌症的概率要低 7.69 个百分点，但是军人职业经历对其他慢性病的影响并不显著。

（2）军人职业对老年人心理健康的影响。诸多研究显示，参军不仅会影响身体健康，而且会对心理健康造成冲击（Hoge 等，2004；Williamson 等，2019）。CLHLS 问卷中针对老年人心理健康的问题包括“是否常感到紧张和害怕”、“是否常感到孤独”、“是否觉得越老越没用”。本文将至少有一个问题回答为“是”的定义为有负面情绪。同时，本文还分析了军人职业对老年人精神疾病的影响。③ 可以看出，军人职业对中国老年人的心理健康并未造成负面影响。之所以与国外的部分研究存在较大差异，可能是因为国外文献大多以参战士兵作为研究对象，直接参与战争可能会导致军人容易出现创伤后应激障碍症和心理疾病。相较而言，本文样本中直接参战的可能并不多。即使早年有参与抗日战争、解放战争和抗美援朝等全面战争的可能，也都是怀有强烈的反侵略和反压迫的爱国主义情怀，甚至会因为参与战争而获得国家荣誉和长期保障，这可能会弱化参战对心理创伤的影响。

（3）城乡异质性。在前面的分析中，本文发现军人职业对老年人带来了健康溢价，但是并未区分城乡异质性效应，接下来进一步分析军人职业能否影响城镇老年人健康。④ 结果显示，军人职业无法改善出生于城镇的老年人的健康状况，但能促进农村老年人健康。如前文所提到的，现有文献关于服役或者军人职业评估的结论并未统一，但是在关于帮助弱势群体获得更好就业机会和更高收入方面较为一致（Card 和 Cardoso，2012；Hirsch 和 Mehay，2003；Padavic 和 Prokos，2017）。而之前基于中国背景展开的研究显示

① 公共服务用财政支出（万亿元）来表示。

② 拓展分析的回归结果请见《经济科学》官网“附录与扩展”。

③ 严格来说，心理疾病和精神疾病并不完全一致。心理疾病是由致病因素导致的脑功能障碍，而精神疾病是精神活动紊乱，导致正常的心理和精神生活难以维持。日常生活中所能观察到的精神病往往表现为严重的“障碍症”，心理咨询师一般将心理问题分为一般心理问题、心理障碍症和精神疾病三类，所以精神疾病也可以视为一种较为严重的心理疾病。

④ 之前的文献表明，参军是一些农村青年获得城镇户籍和就业机会的重要途径（Hou 等，2020；Zhang，2015），户籍和居住地转换经历可能发生在参军以后，故本文不能根据老年人当前的户籍状况或居住地来进行城乡异质性分析。同时，CLHLS 中仅提供了老年人出生地类型（城镇或农村），本文据此来分析城乡异质性。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

农村青年能够通过参军获得较高的回报（Hou 等，2020；Zhang，2015；王兵和杨宝，2019）。本文的发现与这些研究类似，说明军人职业仅能显著改善初始资源禀赋较差的中国农村老年人的健康状况，而对出生于城镇的老年人健康影响不显著。

（4）短期从军经历与老年人健康。在前文的实证分析中，本文着眼于以军人职业为主的长期从军经历对老年人健康的影响。那么，短期从军经历能否促进男性老年人的健康？由于 CLHLS 中并未披露老年人参军的具体时间，因此本文借助中国家庭追踪调查（CFPS）2018 年的数据来进行进一步分析。CFPS 数据库中详细记录了退役军人的入伍时间和退役时间，并发现老年退役军人的平均参军年限仅为 4.34 年。① 因此，相对于 CLHLS 中以军人职业为主的长期从军经历，CFPS 中记录的大部分退役军人可视为拥有短期从军经历。此外，根据问卷，本文所选取的衡量老年人健康的被解释变量包括自评健康、慢性病和记忆力。② 估计结果显示，短期从军经历依旧能够显著促进老年人的自评健康，但是对慢性病和记忆力的影响并不显著。

五、机制检验

基于前文对理论机制的阐述，结合 CLHLS 数据库，本部分从健康行为、教育和收入状况以及社会保障三个方面来进行机制检验。

（一）健康行为

本文从以下五个方面选取健康行为变量，包括过去常吸烟、过去常喝酒、过去常锻炼身体、饮食均衡③和休闲活动④，实证结果如表 2 所示。可以看出，军人职业的系数在前两列中均不显著，这说明相对于非退役军人，老年退役军人在抽烟和喝酒两项不利于健康的习惯上并没有统计上的差异，这与 Teachman（2011）的研究结果类似。此外，在第（3）列中，核心解释变量的系数显著为正，意味着老年退役军人确实有锻炼身体的好习惯。而在第（5）列中，同样可以发现军人职业使得老年人所参与的休闲活动数量显著增加。表 2 的估计结果在总体上说明，老年退役军人有锻炼身体和积极参与休闲活动等利于健康的习惯，而在抽烟和喝酒等不利于健康的行为习惯上与其他老年人并没有显著差异。显然，常锻炼身体和积极参与休闲活动有助于改善老年人的健康。⑤ 因此，军人职业能通过改善老年人的健康行为来提升健康状况。

---

① 这里删除了有 20 年以上从军经历的个体。

② 在自评健康上，将“非常健康”、“很健康”和“比较健康”设置为 1，将“一般”和“不健康”设置为 0。在慢性病变量上，将半年内有慢性疾病的设置为 1，其他设置为 0。记忆力变量根据“老年人能记住主要事情吗”这一问题的回答来进行赋值，将“只能记住一点点”、“只能记住少数”和“只能记住一半”设置为 0，将“能记住多数”和“完全能记住”设置为 1。

③ 若老年人常吃蔬菜（水果）和肉类（猪肉和鱼肉），饮食均衡变量赋值为 1，其他赋值为 0。

④ 休闲活动包括个人户外活动、种花养宠物、阅读书报、打牌（麻将）、看电视、旅游以及参与有组织的社会活动 7 项。

⑤ 本文发现常锻炼身体和休闲活动与衡量老年人健康的三个变量正相关。相应的回归结果请见《经济科学》官网“附录与扩展”中表 A12。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

表 2　军人职业与老年人健康行为（CMP 方法）

	常吸烟	常喝酒	常锻炼	饮食均衡	休闲活动
	(1)	(2)	(3)	(4)	(5)
军人职业	0.1245	−0.2698	0.5355*	0.1094	0.7760*
	(0.1487)	(0.2318)	(0.2995)	(0.2250)	(0.4350)
控制变量	是	是	是	是	是
五年期出生队列固定效应	是	是	是	是	是
省份固定效应	是	是	是	是	是
时间固定效应	是	是	是	是	是
观测值	22 027	22 027	22 027	21 954	22 026

注：前四列报告的是边际效应，第（5）列报告的是估计系数；括号内为出生队列层面的聚类标准误，控制变量与表 1 保持一致。

（二）教育和收入状况

本文选取了两个衡量老年人受教育状况的变量，分别为受过教育和受教育年限。衡量老年人收入状况的变量包括老年人生活来源是否充足、家庭绝对收入和家庭相对收入。其中，家庭绝对收入用前一年家庭人均收入的对数值表示，相对收入是老年人与当地其他住户相比的主观贫富程度。① 之所以选取多个变量，主要是考虑到老年人可能存在记忆偏差，导致自报的受教育年限和家庭收入失真，通过多个变量对比能够有效降低度量误差（程令国等，2015）。

在表 3 的第（1）列和第（2）列中，核心解释变量的系数都显著为正，说明中国老年退役军人的受教育程度更高。这与一些文献所得到的结论相似（Angrist 和 Chen，2011；Bound 和 Turner，2002），说明军人职业能通过提升受教育程度而改善老年人健康。② 表 3 的后三列考察的是军人职业对收入状况的影响，估计结果显示军人职业不仅增加了家庭绝对收入，而且促进了相对收入的增长。实证结果同样支持个体能够通过参军获得收入溢价的结论（Hou 等，2020；Zhang，2015；王兵和杨宝，2019），且发现这种收入溢价能够延续到老年阶段。因此，估计结果说明军人能够通过改善家庭绝对收入和相对收入来进一步提升老年人健康。③

表 3　军人职业与老年人教育和收入（CMP 方法）

	受过教育	受教育年限	生活来源充足	家庭绝对收入	家庭相对收入
	(1)	(2)	(3)	(4)	(5)
军人职业	0.2293**	1.5236**	0.1891	0.9286***	0.1899***
	(0.1077)	(0.5997)	(0.2770)	(0.2445)	(0.0611)
控制变量	是	是	是	是	是
五年期出生队列固定效应	是	是	是	是	是
省份固定效应	是	是	是	是	是
时间固定效应	是	是	是	是	是
观测值	22 027	22 027	22 027	22 007	22 027

注：第（1）、（2）列和第（4）列报告的是边际效应，第（3）、（5）列和第（6）列报告的为估计系数；括号内为出生队列层面的聚类标准误，控制变量与表 1 保持一致。

① 相对富裕 = 3，一般 = 2，相对贫困 = 1。

② 受过教育和受教育年限与衡量老年人健康的三个变量之间也都是正相关的。相应的回归结果请见《经济科学》官网“附录与扩展”中表 A13。

③ 家庭绝对收入和家庭相对收入与衡量老年人健康的三个变量之间均是正相关关系。相应的回归结果请见《经济科学》官网“附录与扩展”中表 A14。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

（三）社会保障

本文选取了四个社会保障变量，分别为社会医疗保险、社会养老保险、政府医疗保障和政府生活保障。[①] 其中，政府医疗保障用医疗费用是否主要由公共支付来表示，政府生活保障用主要生活来源是否依靠政府来衡量。实证结果如表 4 所示。可以看出，核心解释变量在第（1）列和第（3）列中显著，说明老年退役军人在社会医疗保险和政府医疗保障两个方面拥有显著优势。不过，在社会养老保险和政府生活保障上，老年退役军人并没有显著的优势，背后潜在的原因可能是近些年新农保等社会养老保险的快速推行和低保制度的逐渐完善，老年退役军人和非退役军人之间的差距逐渐缩小。毋庸置疑，社会保障与老年人健康水平提升息息相关。[②] 因此，本文的实证结果显示社会保障优势是解释中国老年退役军人健康溢出效应的重要机制。

表 4　军人职业与社会保障（CMP 方法）

	社会医疗保险 （1）	社会养老保险 （2）	政府医疗保障 （3）	政府生活保障 （4）
军人职业	0.2355**	−0.0151	0.1951***	0.0558
	（0.0969）	（0.1345）	（0.0626）	（0.0586）
控制变量	是	是	是	是
五年期出生队列固定效应	是	是	是	是
省份固定效应	是	是	是	是
时间固定效应	是	是	是	是
观测值	16 534	16 524	21 186	21 186

注：报告的均为边际效应，括号内为出生队列层面的聚类标准误，控制变量与表 1 保持一致。

六、结论与政策启示

老年退役军人的健康状况受到社会各界的普遍关注。本文运用 2002—2014 年中国老年健康长寿影响因素调查中的男性样本，全面探讨了军人职业经历对老年人健康的影响及其内在机制。实证结果表明，军人职业显著提升了老年人的躯体健康、认知能力和自评健康，这一结论在通过克服内生性和一系列稳健性检验之后依旧可靠。进一步地，本文发现军人职业使得老年人患癌症的风险下降，但是对心理健康的影响不大。同时，实证结果显示军人职业并不能改善城镇老年人健康，而短期从军经历同样有助于提升老年人健康水平。

军人职业影响老年人健康的机制分析主要围绕健康行为、受教育程度、收入状况以及社会保障等方面展开。在健康行为上，本文发现老年退役军人常锻炼身体和积极参与休闲活动的健康行为习惯有助于提高其健康水平。同时，在教育和收入状况上，军人职业提升了老年人的受教育程度，并促使绝对收入和相对收入的增加。此外，在社会养老保险和政府医疗保障上，老年退役军人存在明显优势。因此，通过改善健康行为、提升受教育程度和收入溢价以及获得社会保障优势等作用机制，军人职业显著改善了老年人

① CLHLS 中社会医疗保险和社会养老保险的统计始于 2005 年。

② 社会医疗保险和政府医疗保障与衡量老年人健康的三个变量之间也都保持正相关关系。相应的回归结果请见《经济科学》官网“附录与扩展”中表 A15。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

的健康状况。

当前，鲜有文献涉及中国等发展中国家老年退役军人健康问题，本文首次运用中国老年人的微观数据，探究军人职业对男性老年人健康的影响，从新的视角解释了中国男性从参军中所获得的健康溢出效应，是对现有文献的补充和拓展。本文的研究有重要的政策启示。一方面，可以构建针对职业军人的特殊保障体系，进一步发挥政府特殊保障措施在提升老年退役军人健康和生命质量中的作用。根据不同兵种、训练经历、战斗经历和从军年限制定多样化的保障方式，全方位满足老年退役军人的医疗需求。另一方面，应加快退役军人健康保障相关政策措施的落地，进一步强化和细化退役军人社会保障相关政策和法规体系的建设，实现老有所养、病有所医。

当然，本文还存在一些局限性。首先，鉴于军事职业本身的高危险性，本文只能获得调查期存活的老年人样本，潜在的幸存者偏差可能导致实证结果低估。其次，受制于调查问卷，无法准确捕捉一些特殊的社会保障方式，如退役军人的安置、优待和补贴等，这些保障可能对改善老年退役军人的健康有积极作用。最后，由于数据库中关于军人职业的调查信息较少，如军兵种、职务和退出现役的方式等，本文无法进行针对性的异质性探讨。不过，随着退役军人事务管理与服务机构在全国各地铺开，相信会有越来越多针对中国退役军人的专项调查，预期将有更多评估老年退役军人的健康和医疗需求的高质量实证研究，以进一步夯实国家退役军人健康保障的理论和现实基础。

参考文献：

1. 程令国、张晔、沈可：《教育如何影响了人们的健康？——来自中国老年人的证据》[J]，《经济学》（季刊）2015 年第 1 期，第 305—330 页。
2. 王兵、杨宝：《参军经历会导致农村青年收入溢价吗？——以 CFPS 2010 调查 9628 个样本为例》[J]，《南方经济》2019 年第 1 期，第 137—148 页。
3. 汪建华：《参军：制度变迁下的社会分层与个体选择性流动》[J]，《社会》2011 年第 3 期，第 138—154 页。
4. Angrist, J. D., 1998, "Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants" [J], *Econometrica*, Vol. 66, No. 2: 249–228.
5. Angrist, J. D., Chen, S. H., 2011, "Schooling and the Vietnam-era GI Bill: Evidence from the Draft Lottery" [J], *American Economic Journal: Applied Economics*, Vol. 3, No. 2: 96–118.
6. Angrist, J. D., Krueger, A. B., 1994, "Why Do World War II Veterans Earn More than Nonveterans?" [J], *Journal of Labor Economics*, Vol. 12, No. 1: 74–97.
7. Angrist, J. D., Imbens, G. W., 1995, "Two-stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity" [J], *Journal of the American Statistical Association*, Vol. 90, No. 430: 431–442.
8. Bedard, K., Deschênes, O., 2006, "The Long-term Impact of Military Service on Health: Evidence from World War II and Korean War Veterans" [J], *American Economic Review*, Vol. 96, No. 1: 176–194.
9. Benmelech, E., Frydman, C., 2015, "Military CEOs" [J], *Journal of Financial Economics*, Vol. 17, No. 1: 43–59.
10. Bound, J., Turner, S., 2002, "Going to War and Going to College: Did World War II and the GI Bill Increase Educational Attainment for Returning Veterans?" [J], *Journal of Labor Economics*, Vol. 20, No. 4: 784–815.
11. Card, D., Cardoso, A. R., 2012, "Can Compulsory Military Service Raise Civilian Wages? Evidence from the

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

Peacetime Draft in Portugal" [J], *American Economic Journal: Applied Economics*, Vol. 4, No. 4: 57–93.
12. Hirsch, B. T., Mehay, S. L., 2003, "Evaluating the Labor Market Performance of Veterans Using a Matched Comparison Group Design" [J], *Journal of Human Resources*, Vol. 38, No. 3: 673–700.
13. Hoge, C. W., Castro, C. A., Messer, S. C., Mcgurk, D., Cotting, D. I., Koffman, R. L., 2004, "Combat Duty in Iraq and Afghanistan, Mental Health Problems, and Barriers to Care" [J], *New England Journal of Medicine*, Vol. 351, No. 1: 13–22.
14. Hou, B., Liu, H., Wang, S. X., 2020, "Returns to Military Service in Off-farm Wage Employment: Evidence from Rural China [J], *China Economic Review*, Vol. 59, No. 1: 101361.
15. Johnston, D. W., Shields, M. A., Siminski, P., 2016, "Long-term Health Effects of Vietnam-era Military Service: A Quasi-experiment Using Australian Conscription Lotteries" [J], *Journal of Health Economics*, Vol. 45, No. 1: 12–26.
16. Landes, S. D., Wilder, J. A., Williams D., 2017, "The Effect of Race and Birth Cohort on the Veteran Mortality Differential" [J], *Social Science & Medicine*, Vol. 179, No. 8: 36–44.
17. Leal, D. L., Teigen, J. M., 2018, "Military Service and Political Participation in the United States: Institutional Experience and the Vote" [J], *Electoral Studies*, Vol. 53, No. 3: 99–110.
18. Lindqvist, E., Vestman, R., 2011, "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment" [J], *American Economic Journal: Applied Economics*, Vol. 3, No. 1: 101–128.
19. McLaughlin, R., Nielsen, L., Waller, M., 2008, "An Evaluation of the Effect of Military Service on Mortality: Quantifying the Healthy Soldier Effect" [J], *Annals of Epidemiology*, Vol. 18, No. 12: 928–936.
20. Padavic, I., Prokos, A., 2017, "Aiming High: Explaining the Earnings Advantage for Female Veterans" [J], *Armed Forces & Society*, Vol. 43, No. 2: 368–386.
21. Sheehan, C. M., Hayward, M. D., 2019, "Black/white Differences in Mortality among Veteran and Non-veteran Males" [J], *Social Science Research*, Vol. 79, No. 3: 101–114.
22. Shen, K., Zeng, Y., 2014, "Direct and Indirect Effects of Childhood Conditions on Survival and Health among Male and Female Elderly in China" [J], *Social Science & Medicine*, Vol. 119, No. 20: 207–214.
23. Spiro, A., Settersten, R. A., Aldwin, C. M., 2016, "Long-term Outcomes of Military Service in Aging and the Life Course: A Positive Re-envisioning" [J], *Gerontologist*, Vol. 56, No. 1: 5–13.
24. Teachman, J., 2011, "Are Veterans Healthier? Military Service and Health at Age 40 in the All-Volunteer Era" [J], *Social Science Research*, Vol. 40, No. 1: 326–335.
25. Waller, M., McGuire, A. C. L., 2011, "Changes over Time in the Healthy Soldier Effect" [J], *Population Health Metrics*, Vol. 9, No. 1: 1–9.
26. Williamson, V., Harwood, H., Greenberg, K., Stevelink, S. A. M., Greenberg, N., 2019, "The Impact of Military Service on the Mental Health of Older UK Veterans: A Qualitative Study" [J], *International Journal of Geriatric Psychiatry*, Vol. 34, No. 10: 1412–1420.
27. Wilmoth, J. M., London, A. S., Parker, W. M., 2010, "Military Service and Men's Health Trajectories in Later Life" [J], *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, Vol. 65, No. 6: 744–755.
28. Wu, X., Treiman, T. J., 2004, "The Household Registration System and Social Stratification in China: 1955—1996" [J], *Demography*, Vol. 41, No. 2: 363–384.
29. Zhang, X. N., 2015, "Military Service and Life Chances in Contemporary China" [J], *Chinese Sociological Review*, Vol. 47, No. 3: 230 –254.
30. Zhao, C., Guo, J., 2022, "Are Veterans Happy? Long-term Military Service and the Life Satisfaction of Elderly Individuals in China" [J], *Journal of Happiness Studies*, Vol. 23, No. 2: 477–508.

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

Imprint from Barracks: Military Profession and the Health of the Elderly

Zhao Chunkai[1], Zhan Yong[2], Chen Huashuai[2]
(1. College of Economics and Management, South China Agricultural University)
(2. Business School, Xiangtan University)

Abstract: By using the data from Chinese Longitudinal Healthy Longevity Survey (CLHLS) from 2002 to 2014, this paper adopts the conditional mixed process (CMP) method with an instrumental variable (IV) to explore the impact of military profession on the health of the Chinese male elderly for the first time. The results show that military profession significantly improves the physical health, cognitive abilities, and self-reported health of the elderly, which is supported by a series of robustness checks. The mechanisms by which military profession improves the health of the elderly include healthy behavioral habits, education and income premiums, and social security. Moreover, military profession reduces the risk of cancer in the elderly, but has little effect on mental health. Furthermore, short-term military service experience also contributes to health improvement.

Keywords: military profession; health; veterans; social security; conditional mixed process
JEL Classification: I12; J14

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

经济科学

Economic Science

◆ 全国中文核心期刊（经济学类）
◆ 中国社会科学引文索引（CSSCI）来源期刊
◆ 中国人文社会科学核心期刊
◆ 国家社科基金资助期刊

国内刊号：CN11-1564/F
国际刊号：ISSN 1002-5839
国内邮发代号：2-840
国外代号：BM262
每期定价：50.00元

Article

Military Service and Health Outcomes of the Elderly in China

Armed Forces & Society
2024, Vol. 50(3) 781–809
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0095327X221149423
journals.sagepub.com/home/afs
SAGE

Chunkai Zhao[1], Jianhao Guo[2], and Jinchen Yan[3]

Abstract

This article investigated the effects and mechanisms of military service on health outcomes of Chinese elderly men aged 60 and older. While numerous studies explored the effect of military service on health in developed countries, we still knew little about the relationship between military service and later health outcomes in developing countries such as China. Using the data of China Health and Retirement Longitudinal Study, we found that military service is positively associated with better health outcomes of the elderly, in terms of physical health, cognitive abilities, and self-rated health. In addition, results of mechanism analysis show that, compared with nonveterans, elderly veterans had healthier habits, better education, higher individual and household income, and more favorable social medical security. Moreover, heterogeneity analysis indicates that this effect is more pronounced for older, rural, and spouseless elderly people. This article provided insights into elderly veterans' health security measures in developing countries.

Keywords

military service, health outcomes, the elderly, veterans, China

[1]College of Economics and Management, South China Agricultural University, Guangzhou, China
[2]Department of Military Finance, Army Logistics University of PLA, Suizhou, China
[3]School of Economics, Renmin University of China, Beijing, China

Corresponding Author:
Jinchen Yan, School of Economics, Renmin University of China, No. 59 Zhongguancun Street, Haidian District, Beijing 100872, China.
Email: jinchenyan@ruc.edu.cn

Introduction

As a special experience in the life course, military service had a pervasive and profound impact on men's socioeconomic status and health outcomes (Angrist, 1990; Fredland & Little, 1985; Kleykamp, 2009; Smith et al., 2012; Zhao & Guo, 2022). While extensive literature assessed the health effects of military service in developed countries (e.g., Angrist et al., 2010; Bedard & Deschênes, 2006; Dobkin & Shabani, 2009; Johnston et al., 2016), few investigated the health status of elderly veterans in the context of developing countries. To fill this gap, in this article, we examined the implications of military service on the health outcomes of elderly individuals not younger than 60[1] based on a unique Chinese setting.

There are two different reasons why military veterans might exhibit higher levels of individual health. On one hand, due to strict screening of physical conditions before joining the army, veterans may be healthier than civilians, which is called the "healthy soldier effect" (McLaughlin et al., 2008). On the other hand, the advantages of veterans are also reflected in enjoying better social welfare and medical resources (Angrist, 1998; Davies et al., 2015; Oh & Berry, 2021; Waller & McGuire, 2011), having more competitive military skills (Johansen et al., 2014; Lindqvist & Vestman, 2011; MacLean, 2017), improving personal qualities and abilities (Angrist & Chen, 2011; Brown & Routon, 2016; Eynde, 2015; L. Wang et al., 2012), obtaining higher socioeconomic status (e.g., Angrist & Krueger, 1994; Bachman et al., 2000; Brown & Routon, 2016; MacLean & Kleykamp, 2021; Sheehan & Hayward, 2019; Smith et al., 2012), and accumulating political resources (Diamant, 2008; Hou et al., 2020; Wakenhut, 1979; Wu & Treiman, 2004). Hence superior external conditions seem to make veterans have better health outcomes.

However, on the negative side, the physical injuries and dangerous exposures experienced during military service may have adverse effects on veterans' health. Risk events directly related to soldiers' health include training accidents, overtraining injuries, public security maintenance, and major emergency events such as infectious diseases and natural disasters (Chase et al., 2015; Elder et al., 2009; McKinney et al., 1997; Wilmoth et al., 2010; Zhao & Guo, 2022). Especially in nonpeaceful times, soldiers would face significant risks disability, chronic illness, and mental illness (Bedard & Deschênes, 2006; Chase et al., 2015; Davies et al., 2015; Dobkin & Shabani, 2009; Williamson et al., 2019), and may even give their lives. Also, some studies provided evidence of higher mortality rates among veterans who have served in combat (Johnston et al., 2016; Landes et al., 2017; Liu et al., 2005). Moreover, active soldiers have to be separated from their families; after retiring from the military, they need to readjust to the social environment and economic changes (Diamant & O'Brien, 2015). These psychological shocks may also bring about worse health status (Black, 1993).

Although the relationship between military service and health outcomes has been well discussed in previous studies, little attention has been paid to the effects of the military service experience on the elderly (Zhao & Guo, 2022). In most countries,

elderly veterans are in highly technical military professions with better educational status; thus, they could obtain higher incomes during their service than nonveterans. In addition, these veterans enjoy better financial status and special benefits upon retirement (Davies et al., 2015), such as higher retirement payments and pensions, superior political treatment, and better health care resources, which further improve their health outcomes.

Since the founding of the People's Republic of China (PRC), 11 disarmament programs have been implemented and a large number of retired military personnel have left the armed forces for society. In addition, China has one of the largest numbers of retired military personnel in the world. According to statistics from the Ministry of Civil Affairs of China, as of 2017, the number of elderly veterans over 57 million, and the number of retired soldiers, disabled soldiers, and survivors of martyrs who need special care and treatment are 8.585 million. In recent years, the Chinese government attached great importance to the health and social security issues of veterans. In March 2018, the Ministry of Veterans Affairs of the PRC was formally established and issued a series of policy documents to protect the rights and interests of veterans. Moreover, plenty of local veterans affairs management and service agencies were established. Therefore, from a realistic perspective, it is necessary to effectively assess the health outcome of veterans who have served to protect the country and maintain social stability.

This article examined the effect of military service on the health of the elderly, based on a nationally representative sample from the China Health and Retirement Longitudinal Study (CHARLS). We found that military service had a significant and positive effect on the health outcomes of the elderly from three perspectives: physical health, cognitive abilities, and self-rated health. This positive health effect is more pronounced for older, rural, and spouseless elderly veterans. In addition, we found that the effect of military service on health outcomes can be explained by healthier living habits, better education attainment, higher income, and more favorable social medical security.

This article contributed to the existing literature in the following two ways. To the best of our knowledge, existing studies on military service and individual health focused on developed countries in Europe and the United States (e.g., Bedard & Deschênes, 2006; Dobkin & Shabani, 2009; Elder et al., 2009; Liu et al., 2005; Wilmoth et al., 2010), while the research on veterans' health outcomes in developing countries is scarce. Undoubtedly, there are great differences between developing and developed countries in terms of war history, military service system, cultural background, living environment, social security, economic treatment, and living habits of veterans (Zhao & Guo, 2022). Based on the particular perspective of Chinese conditions, we further expand the research field on the effect of military service on individual health and provided micro-empirical evidence for policy makers in developing countries.

In addition, the wealth of information in the CHARLS data facilitates us to elucidate possible mechanisms of the positive health effect of military service in the elderly. The mechanism analysis enriched previous related studies (e.g., Landes

et al., 2017; McKinney et al., 1997; Williamson et al., 2019). Moreover, the results support the finding that military service promoted human capital (e.g., Eynde, 2015; Hou et al., 2020; Lindqvist & Vestman, 2011; MacLean, 2017), while we found that some social security measures were ineffective in improving the health status of older veterans. These findings helped us understand the urgency of improving social security systems for veterans in developing countries like China.

Institutional Background

The Military Service Law of China first promulgated in 1955 and amended 3 times in 1984, 1998, and 2011, obliges all male citizens between 18 and 22 to perform military service regardless of their ethnicity, occupation, family background, religious beliefs, or educational status. According to the law, individuals who have reached 18 shall register for military service and remain in active service until 22. Clearly, only individuals who pass physical examinations and political background check are eligible for enlistment.[2] During enlistment, individuals must give priority to their military duties.

In 1978, China began a special system combining compulsory military service and voluntary military service. Soldiers subject to the compulsory military service system are called conscripts, while military personnel who voluntarily join the army are called noncommissioned officers. The volunteers may be long-term serving conscripts by rising as cadres or citizens with specialized skills from the civilian sector. Moreover, volunteers who perform well can be promoted to military officers. Prior to 1999, the required term of service in the Army, Air Force, and Navy was 3, 4 and 4–5 years, respectively. After 1999, periods of conscript service were reduced to 2 years for all branches of the military. However, the service length for volunteers with military ranks could exceed 30 years.[3]

Furthermore, the rank salary system applies to volunteers. The basic pay and benefits (such as disability pension, hardship assistance, and housing allowance) on active duty are all implemented in accordance with the relevant local regulations for cadres. If approved for separation, demobilization, or eligible for retirement, they receive not only a retirement or pension payment but also a variety of special privileges, such as job resettlement, military insurance, and honors. Moreover, China mainly adopts a lagging compensation approach in the treatment of military personnel.[4] In other words, some veterans, especially officers and elderly soldiers and military officers, are able to receive more compensatory monetary income and social benefits compared with active-duty personnel.

Theoretical Hypotheses

In this section, we discussed in detail the theoretical mechanisms of military service on the health outcomes of the elderly. We classified possible channels into the following categories: healthy habits, education, income, and social security.

Healthy Habits

Military service with high-intensity training, strict discipline, and a barrack culture emphasizing strictness, obedience, norms, and justice, may change individual values, personality characteristics, and living habits (Benmelech & Frydman, 2015; Zhao & Guo, 2022), thereby indirectly affecting veterans' health. Previous studies found that soldiers usually relieve stress by smoking and drinking to combat the high risks of occupation (Bedard & Deschênes, 2006; Elder et al., 2009; McKinney et al., 1997). Worse, there are strong externalities in the military collective life, and it is difficult for soldiers to maintain independent and stay away from these unhealthy behaviors (Wilmoth et al., 2010).

Furthermore, military service also leads individuals to form better exercise habits to improve their physical fitness through strict discipline (McLaughlin et al., 2008). More importantly, physical exercise in the military may be more scientific and effective. In addition, veterans are likely to form military-related social networks through military service by making important connections with local comrades, officers, and other veterans (De Tray, 1982; Hou et al., 2020). Such social networks may lead older adults to become more involved in social activities and leisure sports, which were beneficial to health outcomes of Chinese elderly (e.g., W. Cao et al., 2015; T. Li & Zhang, 2015; Ye & Zhang, 2021). In particular, for elderly veterans, these health-related habits formed during their military careers could persist into their retirement and even later life (Elder et al., 2009; Spiro et al., 2016). Nevertheless, we proposed the first hypothesis:

Hypothesis 1 (H1): Veterans may have healthier living habits, which contributes to their health statues.

Education

We considered that education may be another important channel to explain better health outcomes of elderly veterans. However, the existing literature has not reached a consistent conclusion regarding the relationship between military service and educational achievement. Some scholars found that military service reduced individuals' access to education because soldiers have to stay away from the conventional education system and receive military education and military skills training (Teachman, 2007; Teachman & Call, 1996; Torun & Tumen, 2016). In addition, some evidence suggested that military service in the United States during the Vietnam era had a positive impact on individual education, since GI Bill schooling benefits helped eligible veterans pay for college, graduate school, and training programs (Angrist & Chen, 2011; Bound & Turner, 2002). Also, the military could provide low-cost educational opportunities, training programs, and an atmosphere conducive to learning (Eynde, 2015; Gade et al., 1991; Routon, 2014), which subsequently improves the cognitive abilities of soldiers (Eynde, 2015; Hou et al., 2020; L. Wang et al., 2012).

In addition, military service creates the conditions for an enlisted or officer to receive a military education. In China, a number of officers are selected from military academies. Military colleges[5] not only teach military knowledge, but generalist nonmilitary knowledge and skills are also imparted to students. Even some military colleges that only enroll active-duty military personnel, and the original benefits and perks are not eliminated by attending school (Heaton & Lovejoy, 1986). As a result, elderly veterans may have a higher level of education. Needless to say, the positive impact of education on health has been widely recognized (e.g., Cutler & Lleras-Muney, 2010; Grossman, 2000). Therefore, we put forward the second hypothesis:

Hypothesis 2 (H2): Veterans have better educational attainment, which is positively associated with better health in their elderly stage.

Income

As mentioned earlier, there is controversy about the effect of military experience on earnings. Scholars found that military service during World War II boosted the income of U.S. veterans (e.g., Angrist & Krueger, 1994; De Tray, 1982; MacLean & Kleykamp, 2021; Smith et al., 2012), while Vietnam-era veterans earned less (e.g., Angrist, 1990; Spiro et al., 2016; Teachman & Call, 1996). Even so, most of prior studies affirmed the positive effects of military service on the earnings of some socially disadvantaged groups, such as minorities, rural people, and women (Angrist, 1998; Hou et al., 2020; Padavic & Prokos, 2017; Sheehan & Hayward, 2019), because military service tends to help them gain access to elite jobs and narrow gaps in human and social capital.

In China, people with military service are likely to receive higher incomes. First, some veterans have higher military ranks and sizable retirement payments to protect their livelihoods after retirement. Second, the Chinese government will provide proper employment resettlement for them mainly in state-owned enterprises and the public sectors (Hou et al., 2020; Wu & Treiman, 2004). Third, some veterans were engaged in related high-tech jobs in the military, so that they had opportunities to find high-income jobs, such as engineers, scientific researchers, and university teachers (Zhang, 2015). Due to positive effects of income on elderly people's health outcomes (e.g., Oh & Berry, 2021; Shen & Zeng, 2014), we raised the third hypothesis:

Hypothesis 3 (H3): Veterans with military service may have higher incomes than nonveterans, which benefits their health.

Social Security

The protection of soldier rights and interests involves the stability of national military human capital, army combat effectiveness, and even national security. Therefore,

countries around the world attach great importance to the social security of active and retired soldiers. In general, in addition to regular social security, the government provides some guarantees and supports as compensation (Y. Cao et al., 2019). Taking the United States as an example, veterans enjoy special medical treatment, employment assistance, and education grants (Angrist et al., 2010). In particular, veterans entering a war have higher benefits, includes the income from the Veterans Affairs' Disability Compensation program and Social Security Administration's Disability Insurance program (Davies et al., 2015).

Furthermore, in China, several veterans are those who have contributed to national security and defense constructions, and some of them even participated in the war for national independence. They are likely to receive more social benefits, such as high subsidies, free medical care, government livelihood guarantees, and even more preferential treatment for their families. Therefore, the elderly veterans receive more social benefits and medical insurance in theory. Confidently, social security has played an important role in improving the health of the Chinese elderly (Cheng et al., 2015; Oh & Berry, 2021). To sum up, we derived the fourth hypothesis:

Hypothesis 4 (H4): Elderly veterans have advantages in social security, which is good for their health.

Data, Variables, and Method

Data Resource

The data come from the CHARLS, which has been followed up 5 times from 2008 to 2018. The survey scope of CHARLS includes 28 provinces (autonomous regions or municipalities), covering over 97% of the population in China. The CHARLS records in detail the health outcome, healthy behaviors, living habits, economic conditions, and family background of Chinese middle-aged and elderly people. It is an important large-scale micro-database for the current research on China's aging and elderly health issues (Tan et al., 2017). More importantly, in the 2014 special wave, Life History Survey, CHARLS provided respondents with information on whether they had military service. We selected three waves of follow-up data from 2011 to 2015.[6] Moreover, the females were excluded considering that soldiers are predominantly male (Angrist, 1990; Dobkin & Shabani, 2009; Johnston et al., 2016). Of course, observations under 60 at the time of each survey were also removed.[7]

Variables and Measurements

Dependent Variables. We measured the health outcomes of the elderly using three indicators: instrumental activities of daily living (IADL), cognitive abilities (CA), and self-rated health (SRH). First, IADL assesses the physical health of the elderly. The basis is to observe whether the elderly can complete the eight activities.[8]

According to Barberger-Gateau et al. (1999) and Shen & Zeng (2014), *IADL* will be set to 1 if the elderly person is able to complete all eight activities without assistance, meaning there is no instrumental dysfunction, and 0 otherwise.

Second, CA has been widely used as a screening instrument for the detection of cognitive impairment (Folstein et al., 1975; Zeng & Vaupel, 2002). There are 25 questions in the CHARLS used to measure the cognitive abilities of the elderly from three areas: calculation ability, logical thinking, and language ability. We defined elderly individuals who received a score of 20 and above (reaching the top 80%) as cognitively unimpaired and assigned the variable *CA* a value of 1, and 0 otherwise.

Third, SRH is an individual's subjective perception of their own health, and can comprehensively reflect older people's illness and the loss of body function (Grimard et al., 2010; Laaksonen et al., 2005). We assigned values based on the answers to the question "Would you say your health is very good, good, fair, poor or very poor?." If the answers were "good" or "very good," we set the value of *SRH* to 1, and 0 for "fair," "poor," or "very poor."

Independent Variable. The key independent variable is *Military*. We assigned values based on the answers to the question "Which of the following category did your job belong to," the answer "Army" to 1 and others to 0. What needs to be emphasized is that the CHALRS here counted all jobs for 6 months or more respondents have held in the past. However, the years in the military is not reported.

Control Variables. According to previous related literature (Bedard & Deschênes, 2006; W. Cao et al., 2015; Grimard et al., 2010; Laaksonen et al., 2005), we selected two types of control variables. First, at the individual level, we controlled for *Age, Married, Hukou*,[9] *Hungry in childhood, Poverty in childhood*, and *Parental death in childhood*. Controlling for childhood experiences helps reduce potential endogeneity due to the fact that these variables may be strongly associated with military service and health outcomes (Hou et al., 2020). Second, household characteristics such as *Father farming, Education of mother, Education of father*, and *Number of siblings* were also taken into consideration. The detailed definition of these variables is shown in Table A1 of Appendix A.

Summary Statistics

The descriptive statistics of the above variables are shown in Table 1. We found that 11.03% of observations had military service. On average, 31.7% of the elderly had no instrumental impairment, 53.60% had intact cognitive abilities, and 76.90% rated themselves as physically healthy. In addition, we reported the results of mean difference tests for veterans and nonveterans. The results indicated that the elderly with military service had 9.08% lower proportions of impaired instrumental function, 8.39% lower rates of cognitive impairment, and 3.06% higher proportions of self-rated health. These descriptive statistics initially validated our expectation that elderly individuals with military service had better health outcomes.

Table 1. Descriptive Statistics.

	Full sample			Nonveterans	Veterans	
Variables	*N*	*M*	*SD*	*M*	*M*	Diff.
IADL	14,653	0.3170	0.4653	0.3058	0.3966	−0.0908***
CA	14,867	0.5360	0.4987	0.5253	0.6092	−0.0839***
SRH	14,746	0.7690	0.4215	0.7656	0.7962	−0.0306***
Military	17,233	0.1103	0.3133			
Age	17,233	67.1423	5.5475	67.1586	67.0110	0.1476
Married	17,233	0.8926	0.3096	0.8904	0.9106	−0.0201***
Hukou	17,233	0.6992	0.4586	0.7127	0.5921	0.1206***
Hungry in childhood	17,233	0.7866	0.4097	0.7792	0.8464	−0.0672***
Poverty in childhood	17,233	0.4116	0.4921	0.4157	0.3787	0.0369***
Parental death in childhood	17,233	0.2824	0.4502	0.2849	0.2620	0.0229**
Father farming	17,233	0.7700	0.4208	0.7685	0.7822	−0.0137
Education of mother	17,233	0.8759	1.3993	0.8900	0.7628	0.1272***
Education of father	17,233	0.5184	1.2974	0.5252	0.4634	0.0617*
Number of siblings	17,233	0.7144	1.5187	0.7173	0.6912	0.0260

Note. IADL = instrumental activities of daily living; CA = cognitive abilities; SRH = self-rated health.

Methods

We considered regression models to estimate the effect of military service on the health outcomes of the elderly. Since all three dependent variables are binary, we employed the nonlinear probit method which can be set as follows:

$$Health_{it}^{*} = \beta_0 + \beta_1 Military_i + \beta_2 X_{it} + \delta_{t-c} + \lambda_p + \theta_t + \varepsilon_{it}, \quad (1)$$

$$Health_{it} = \begin{cases} 1 \;\; if \;\; Health_{it}^{*} > 0; \\ 0 \;\; otherwise, \end{cases} \quad (2)$$

where $Health_{it}^{*}$ is a latent variable representing health outcomes of the old person *i* in year *t*. $Military_i$ is a dummy variable which indicates whether the person *i* had military service. X_{it} refer to a set of control variables mentioned earlier. The subscripts *p* and *c* indicate the province and year of birth, respectively. δ_{t-c} represents the age fixed effect, which takes into account the close relationship between age and health outcomes of the elderly (Cheng et al., 2015; T. Li & Zhang, 2015). λ_p and θ_t mean province and year fixed effects, respectively. ε_{it} denotes the random disturbance term.

Furthermore, we created the following models to further test the mechanism-related hypotheses:

$$MEC_{it} = \alpha_0 + \alpha_1 Military_i + \alpha_2 X_{it} + \delta_{t-c} + \lambda_p + \theta_t + \varepsilon_{it}, \quad (3)$$

$$Health_{it} = \gamma_0 + \gamma_1 MEC_{it} + \gamma_2 X_{it} + \delta_{t-c} + \lambda_p + \theta_t + \varepsilon_{it}, \quad (4)$$

In Equations 3 and 4, the mechanism variable set is denoted by MEC_{it}. α_1 and γ_1 are the main coefficients that we focused on. If they are both statistically significant and the sign is what was anticipated, it means the mechanism is operating.[10]

Results

Baseline Results

We examined the effects of military service on the health of elderly individuals, and the marginal effect results using probit models are presented in Table 2. As mentioned earlier, we controlled for a number of individual- and household-level characteristic variables as well as age, province, and year fixed effects. The estimates indicated that military service is beneficial for the physical health and cognitive abilities, which are objective measures of health outcomes of the elderly. For subjective self-rated health, the coefficient of the core independent variable *Military* is also significantly positive. As expected, the baseline results suggested that the elderly with military service were indeed healthier in China.

Specifically, we found that military service may increase the probability of unimpaired physical health by 3.56 percentage points, intact cognitive ability by 5.09 percentage points, and self-rated health by 5.13 percentage points among elderly individuals. Given that the mean values of IADL, CA, and SRH in the full sample are 0.3058, 0.5253, and 0.7656, respectively (as shown in Table 1), these estimates amount to 11.64%, 9.69%, and 6.70% increase in the probability of unimpaired physical health, intact cognitive ability, and self-rated health for elderly veterans, respectively (e.g., 3.56%/30.58% = 11.64%).

Mechanisms of Health Outcomes

As discussed above, the following mechanisms may be responsible for better health outcomes in elderly veterans: healthy habits, education, income, and social security. To verify hypotheses, we tested these mechanisms in turn by empirical analysis in this section.

Healthy Habits. Four variables were used to measure healthy habits of the elderly according to the CHARLS questionnaire. We constructed three binary variables indicating whether the elderly used to smoke, drink heavily, and exercise regularly. In addition, we set up variable related to social networks in terms of types of leisure

Table 2. The Effect of Military Service on Health of the Elderly: Probit Estimations.

	(1)	(2)	(3)
Variables	IADL	CA	SRH
Military	0.0356***	0.0509***	0.0513***
	(0.0123)	(0.0125)	(0.0115)
Married	0.0190	0.0366***	0.0472***
	(0.0126)	(0.0131)	(0.0110)
Hukou	0.0060	0.2294***	0.0562***
	(0.0110)	(0.0112)	(0.0097)
Hungry in childhood	−0.0119	−0.0507***	0.0339***
	(0.0096)	(0.0098)	(0.0086)
Poverty in childhood	−0.0083	−0.0717***	0.0667***
	(0.0098)	(0.0101)	(0.0085)
Parental death in childhood	0.0412***	−0.0437***	0.0214***
	(0.0079)	(0.0079)	(0.0071)
Father farming	0.0071	0.0283***	0.0143*
	(0.0087)	(0.0087)	(0.0078)
Education of mother	0.0405***	0.0216*	0.0196*
	(0.0127)	(0.0125)	(0.0110)
Education of father	0.0009	−0.0046	0.0071*
	(0.0047)	(0.0047)	(0.0041)
Number of siblings	0.0098**	0.0098**	0.0104***
	(0.0046)	(0.0045)	(0.0040)
Age fixed effect	Yes	Yes	Yes
Province fixed effect	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes
N	14,653	14,867	14,746
Pseudo R^2	0.0224	0.1168	0.0457

Note. Marginal effects are reported for probit method in all columns. IADL = instrumental activities of daily living; CA = cognitive abilities; SRH = self-rated health.
***, **, and * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are shown in parentheses.

activities in which the elderly regularly engaged (T. Li & Zhang, 2015; Ye & Zhang, 2021). Definitions of these variables are shown in Panel A of Appendix Table A2.

In Table 3, the results from probit models show that the coefficient on *Military* is insignificant in Columns (1) and (2), suggesting that there is no association between military service and smoking and drinking, which is similar to the findings of Teachman (2011). In Columns (3) and (4), by using probit and ordinary least square (OLS) methods, we found that elderly veterans were more motivated to exercise and engage in leisure activities. Moreover, based on the estimation results of Equation 4, shown in Table B1 in Appendices, as expected, we found that both exercise and leisure activities are significantly and positively associated with three variables of health

Table 3. Military Service and Health Habits of the Elderly.

	(1)	(2)	(3)	(4)
	Smoking	Alcoholism	Exercises	Leisure activities
Variables	Probit	Probit	Probit	OLS
Military	0.0127	−0.0007	0.0226***	0.0984***
	(0.0116)	(0.0131)	(0.0085)	(0.0122)
Control variables	Yes	Yes	Yes	Yes
Age fixed effect	Yes	Yes	Yes	Yes
Province fixed effect	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes
N	15,088	15,063	14,553	17,233
Pseudo/Adjusted R^2	0.1777	0.0419	0.1458	0.0389

Note. Marginal effects are reported for probit method in first three columns. Baseline control variables and fixed effects are added in all regressions. OLS = ordinary least square.
***, **, and * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are shown in parentheses.

outcomes in the elderly. Therefore, these estimates support Hypothesis 1 and imply that military service helped elderly individuals develop better health habits, primarily in physical exercises and social networks, which is good for their health status.

Education and Income. Next, we considered the education and income mechanisms, as laid out in Hypotheses 2 and 3. We selected two variables to measure the education, namely, *Schooled* and *Years of schooling*, Moreover, two income variables are *Individual income* and *Household income*. The details of these variables are in Panels B and C of Table A2 in Appendix A.

In the first two columns of Table 4, the probit or OLS estimates show a significant positive association between military service and education, suggesting that elderly individuals with military service in China have a higher level of education, which is similar to some previous results (Angrist & Chen, 2011; Bound & Turner, 2002). In addition, we found that the estimates of military service are positive and significant in Columns (3) and (4) of Table 4, suggesting that military service increased individual and household income for older adults. These finding are consistent with previous findings based on Chinese young samples (Hou et al., 2020; Zhang, 2015).

Furthermore, as before, in Table B2 in Appendix B, we could see that both education and income are significantly and positively associated with the health of the elderly, which is consistent with what has been highlighted in previous studies (Cutler & Lleras-Muney, 2010; Shen & Zeng, 2014). All in all, these results provided supportive evidence for Hypotheses 2 and 3 that military service can improve health outcomes of the elderly by increasing their education and income.

Table 4. Military Service and Education and Income of the Elderly.

	(1)	(2)	(3)	(4)
	Schooled	Years of schooling	Individual income	Household income
Variables	Probit	OLS	OLS	OLS
Military	0.0277***	0.3181***	0.1012***	0.5838**
	(0.0078)	(0.0618)	(0.0373)	(0.2880)
Control variables	Yes	Yes	Yes	Yes
Age fixed effect	Yes	Yes	Yes	Yes
Province fixed effect	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes
N	17,181	16,578	8,680	17,176
Pseudo/Adjusted R^2	0.1776	0.1021	0.4574	0.0296

Note. Marginal effects are reported for probit in column (1). Baseline control variables and fixed effects are added in all regressions. OLS = ordinary least square.
***, **, and * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are shown in parentheses.

Social Security. The CHARLS recorded two basic types of social security related to the elderly, medical insurance and pension insurance. We constructed four mechanism variables, *Social medical insurance, Social pension insurance, Government medical support*, and *Government pension support*, of which specific definitions are shown in Panel D of Table A2 in Appendix A. Among them, the first two are conventional for all seniors, while the latter two are that the government grants to certain specific groups that are vulnerable or make contributions to the country and society.

Table 5 presents the empirical results of the effect of military service on social security, and coefficients of *Military* is significantly positive in Columns (1), (3), and (4), indicating that elderly veterans have more medical security and special government guarantees, as expected in the Hypothesis 4. Furthermore, we checked the relationship between social security variables and health variables in older adults, as presented in Table B3 of Appendix B. However, the results suggest that although the estimated coefficients in Panels B and C are positive, they were mostly insignificant, especially for government medical support. These findings imply that although elderly veterans receive more government benefits, which does not translate into more positive health outcomes. Taken together, our estimates partly support the Hypothesis 4 that military service improves the health of the elderly through social medical security.

Heterogeneity Effects

We further analyzed heterogeneity effects of military service on health outcomes of elderly individuals from three perspectives: age, *hukou*, and spouse.

Table 5. Military Service and Social Security: Probit Estimations.

Variables	(1) Social medical insurance	(2) Social pension insurance	(3) Government medical support	(4) Government pension support
Military	0.0059*	0.0077	0.0437***	0.0104***
	(0.0033)	(0.0050)	(0.0059)	(0.0028)
Control variables	Yes	Yes	Yes	Yes
Age fixed effect	Yes	Yes	Yes	Yes
Province fixed effect	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes
N	16,381	11,880	17,019	16,158
Pseudo R^2	0.1517	0.8016	0.0593	0.1982

Note. Marginal effects are reported for probit method in all columns. Baseline control variables and fixed effects are added in all regressions.
***, **, and * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are shown in parentheses.

First, previous studies suggested that the "healthy soldier effect" faded with age (Liu et al., 2005; Wilmoth et al., 2010). In line with conventional practice, we divided the full sample into two groups, using age 65 (the median) as the dividing line. The marginal effects using probit models are shown in Panels A and B of Table 6. We found that the effect of military serving is not significant for the elderly younger than 65 years, but the positive effect is greater for the older seniors. Clearly, older adults have worse health overall, and using these older adults as controls, veterans have better health, suggesting a long-term effect of military service on health outcomes of Chinese elderly. Our findings differ from the existing literature and may be explained by better exercise habits and greater access to social benefits for older veterans.

Second, due to the disparities in health care conditions, public services, and nutritional provision between urban and rural areas in China (W. Cao et al., 2015; Zhang, 2015), it is necessary to explore urban–rural heterogeneity effects. Results in Panels C and D of Table 6 show that the differences in effects of military service on physical health between urban and rural elderly individuals are not obvious, but rural veterans have more significant advantages in cognitive abilities and self-rated health. Using Chinese older adults in rural areas, who have lower educational attainment and poorer health endowment (Ye & Zhang, 2021), as controls, veterans had better cognitive ability and self-rated health.

Third, a spouse can enhance the health of the elderly by forming good lifestyle habits, improving the household division of labor, and providing emotional support (Brockmann & Klein, 2004), therefore, whether elderly individuals have a spouse may face different health risks. In Panels E and F of Table 6, we found that in the spouseless subsample, military service formed a greater positive effect on health outcomes of elderly individuals. As before, veterans were in better health relative to

Table 6. Heterogeneity Effects: Probit Estimations.

	(1)	(2)	(3)
Variables	IADL	CA	SRH
Panel A. Age≦65			
Military	0.0257	−0.0148	0.0437***
	(0.0175)	(0.0170)	(0.0157)
N	6,698	7,905	6,863
Panel B. Age > 66			
Military	0.0390**	0.1098***	0.0507***
	(0.0172)	(0.0175)	(0.0170)
N	7,783	6,941	7,864
Panel C. Urban			
Military	0.0121	0.0413**	0.0334*
	(0.0193)	(0.0166)	(0.0183)
N	9,977	10,163	10,177
Panel D. Rural			
Military	0.0487***	0.0728***	0.0698***
	(0.0160)	(0.0191)	(0.0153)
N	4,500	4,671	4,534
Panel E. With a spouse			
Military	0.0052	0.0381***	0.0735*
	(0.0412)	(0.0130)	(0.0436)
N	12,910	13,403	13,171
Panel F. Without a spouse			
Military	0.0348***	0.1742***	0.0472***
	(0.0128)	(0.0453)	(0.0118)
N	1,563	1,432	1,535

Note. Marginal effects are reported for probit method in all columns. Baseline control variables and fixed effects are added in all regressions. IADL = instrumental activities of daily living; CA = cognitive abilities; SRH = self-rated health.
***, **, and * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are shown in parentheses.

vulnerable older adults without a spouse. These estimates suggest that military service may effectively compensate for the negative effect of not having a spouse on the health of the elderly.

Robustness Checks

We conducted several robustness checks, including placebo tests, instrumental variable (IV) estimations for mitigating potential endogenous problems, replacing dependent variables, and deleting some special observations. The empirical results are shown in Appendix Tables C1 and C2, and we found that the main result that

military service promoted the health outcomes of elderly individuals remained unchanged.[11]

Conclusions and Discussion

This article examined the effect of military service on health outcomes among the Chinese elderly. We drew three main conclusions from our findings. First, military service significantly promoted the physical health and cognitive abilities of the elderly, leading to an increase in the probability of self-rated health, which is supported by a series of robustness checks. Second, the results of mechanism tests show that positive influences of military service on the health can be explained by healthier habits, better education, higher income, and more advantageous social medical security of the elderly veterans. However, we found that the greater government support benefits enjoyed by veterans did not translate into improved health status. Third, we found that the positive effect of military service on health outcomes is more valid for older, rural, and spouseless elderly individuals.

At present, a few studies focused on health outcomes of elderly veterans in developing countries. We used the microdata of the Chinese elderly individuals to explore the effect of military service on health outcomes and the underlying mechanisms. In theory, this article supplemented and developed the existing literature. First, our results are similar to some literature from developed countries (Angrist, 1998; McLaughlin et al., 2008; Sheehan & Hayward, 2019; Waller & McGuire, 2011), which emphasized the positive effects of military service on veterans' health. Second, we expanded the research field about military service for human capital accumulation (e.g., Johansen et al., 2014; Lindqvist & Vestman, 2011; MacLean, 2017; Padavic & Prokos, 2017). Third, several studies using samples of Chinese young people found that military service contributes to increased earnings, access to formal employment and political resources (Hou et al., 2020; Wu & Treiman, 2004; Zhang, 2015), and these advantages may make veterans healthier. Thus, our findings further enriched the research related to military service in developing countries like China based on the perspective of older adults.

Our findings have important practical implications. First, considering that health is an important component of human capital, especially the health of the elderly, we confirmed that military service is an important way to promote human capital accumulation in China. Thus, for Chinese young people, joining the military may not only be a way to help them achieve higher socioeconomic status (Hou et al., 2020; Zhang, 2015), but there are also long-term positive effects from a health perspective. Second, this article provided some support for developing countries to develop a better social security system that is beneficial to veterans' health. We found that elderly veterans did not have an advantage in some social security methods in China (Zhao & Guo, 2022). Worse, mechanism analysis results implied that the positive effects of existing social security policies on veterans' health were very limited. Moreover, we concentrated on veterans who theoretically received more social welfare and social

security measures. Yet they did not obtain health benefits from the existing social security system. For those older adults who were conscripts or had a short-term service experience, their health safeguards may be of more concern (Diamant & O'Brien, 2015).

For the armed forces in China, taking some necessary measures may also further safeguard the health of elderly veterans (O'Brien & Diamant, 2015), especially in the context of the improvement of China's veterans protection system and the implementation of the civil-military integration strategy (Guo & Zhao, 2019). On one hand, it is important to continuously improve training practices to avoid injuries to military personnel due to training or overtraining. After all, physical injuries during military service may have a negative impact on the physical and mental health of elderly veterans. On the other hand, the armed forces should actively cooperate with the civil administration to improve health protection for elderly veterans (J. Wang, 2022). Diversified protection methods should be developed according to different types of soldiers, training experience, combat experience, and years of military service, which helps to meet the medical needs of elderly veterans in all aspects.

In addition, this article refutes some arguments that more misfortune has befallen Chinese veterans (Diamant, 2011; O'Brien & Diamant, 2015). At least, at the level of long-term health effects, Chinese elderly veterans indeed enjoyed more positive outcomes. We used large-scale survey data to obtain more reliable conclusions. Of course, regressions and statistical results describe overall trends and average effects and do not capture individual conditions. It may be unavoidable that individual veterans experience injustice (Zhang, 2015). Therefore, with the development of China's economy and improvement of the social security system, more efforts are still needed to enhance the treatment of veterans who have contributed to the country (Diamant, 2008; J. Wang, 2022), especially to improve the quality of life of older veterans.

There are some limitations to this article. First, due to survivor bias, we could only obtain the observations of the living elderly. Some soldiers might die in dangerous events and wars; there is no doubt that the inclusion of these samples could lead to a decline in the overall health of the veteran population, as death is one of the extreme outcome of unhealthiness. Therefore, there is a potential risk of overestimation in our results (White et al., 2019). Second, the analysis of social security for elderly veterans is not comprehensive. Limited by the CHARLS questionnaire, some special social security measures that may improve veterans' health outcome cannot be accurately captured, such as cash subsidies, political preferential treatment, employment resettlement, and so forth. Last, while we focused on elderly men over 60 years old, it remains to be further revealed by future studies whether military service has a positive effect on the health of young and middle-aged groups and women in developing countries.

Appendices

Appendix A. Definition of Variables

Table A1. Definitions of Dependent Variables, Independent Variable, and Control Variables.

Variables	Definitions
Panel A. Dependent variables	
IADL	= 1 if the elderly individual could complete all eight daily activities independently; 0 otherwise
CA	= 1 if the cognitive ability test score of the elderly individual was not less than 20; 0 otherwise
SRH	Good or very good = 1; fair, poor, or very poor = 0
Panel B. Independent variables	
Military	= 1 if the elderly individual was the military personnel; 0 otherwise
Panel C. Control variables	
Married	= 1 if the elderly individual was married and had a spouse; 0 otherwise
Hukou	Urban = 1; rural = 0
Hungry in childhood	= 1 if often went hungry in childhood; 0 otherwise
Poverty in childhood	= 1 if the family of origin was poor during childhood; 0 otherwise
Parental death in childhood	= 1 if at least one parent died in childhood; 0 otherwise
Father farming	Farmer = 1; other occupations = 0
Education of mother	*Unschooled* = 0; *elementary school* = 1; *middle school* = 2; *high school or vocational school* = 3; *college or bachelor degree* = 4
Education of father	*Unschooled* = 0; *elementary school* = 1; *middle school* = 2; *high school or vocational school* = 3; *college or bachelor degree* = 4
Number of siblings	The number of siblings of the elderly individual

Note. IADL = instrumental activities of daily living; CA = cognitive abilities; SRH = self-rated health.

Table A2. Definitions of Mechanism Variables.

Variables	Definitions
Panel A. Healthy habits	
Smoking	= 1 if the respondent is a current or past smoker; 0 otherwise
Alcoholism	= 1 if the respondent is or has been a regular drinker; 0 otherwise
Exercises	= 1 if the respondent exercises regularly; 0 otherwise
Leisure activities	Sum score across eleven dimensions according to the CHARLS, such as volunteering, participating in club activities, playing cards, visiting friends.
Panel B. Education	
Schooled	= 1 if the respondent had formal education; 0 otherwise
Years of schooling	*Unschooled* = 0; *elementary school* = 6; *middle school* = 9; *high school or vocational school* = 15; *college* = 18; *bachelor degree* = 19
Panel C. Income	
Individual income	Respondents' total personal income in the last year (10,000 yuan, RMB)
Household income	Total income of all family members last year (10,000 yuan, RMB)
Panel D. Social security	
Social medical insurance	= 1 if the respondent has social medical insurance (not including commercial insurance); 0 otherwise
Social pension insurance	= 1 if the respondent has social pension insurance (not including commercial insurance); 0 otherwise
Government medical support	= 1 if the government pays the premium of medical insurance; 0 otherwise
Government pension support	= 1 if the government pays the premium of pension insurance; 0 otherwise

Note. CHARLS = China Health and Retirement Longitudinal Study.

Appendix B. Mechanisms Tests

Table B1. Healthy Habits and Health Outcomes of the Elderly: Probit Estimations.

	(1)	(2)	(3)
Variables	IADL	CA	SRH
Panel A. Exercises			
Exercises	0.0625***	0.1465***	0.0549***
	(0.0134)	(0.0146)	(0.0113)
N	11,369	11,176	11,559
Panel B. Leisure activities			
Leisure activities	0.0457***	0.0284***	0.0206***
	(0.0081)	(0.0081)	(0.0074)
N	12,915	13,273	13,166

Note. Marginal effects are reported for probit method in all columns. Baseline control variables and fixed effects are added in all regressions. IADL = instrumental activities of daily living; CA = cognitive abilities; SRH = self-rated health.
***, **, and * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are shown in parentheses.

Table B2. Education and Income and Health Outcomes of the Elderly: Probit Estimations.

	(1)	(2)	(3)
Variables	IADL	CA	SRH
Panel A. Schooled			
Schooled	0.0825***	0.0205*	0.0606***
	(0.0150)	(0.0115)	(0.0128)
N	12,915	13,273	13,166
Panel B. Years of schooling			
Years of schooling	0.0068***	0.0040**	0.0031*
	(0.0018)	(0.0018)	(0.0017)
N	12,406	12,695	12,651
Panel C. Individual income			
Individual income	0.0170***	0.0577***	0.0183***
	(0.0053)	(0.0062)	(0.0049)
N	6,782	6,884	7,034
Panel D. Household income			
Household income	0.0104*	0.0225**	0.0325***
	(0.0054)	(0.0102)	(0.0103)
N	12,896	13,233	13,152

Note. Marginal effects are reported for probit method in all columns. Baseline control variables and fixed effects are added in all regressions. IADL = instrumental activities of daily living; CA = cognitive abilities; SRH = self-rated health.
***, **, and * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are shown in parentheses.

Table B3. Social Security and Health Outcomes of the Elderly: Probit Estimations.

	(1)	(2)	(3)
Variables	IADL	CA	SRH
Panel A. Social medical insurance			
Social medical insurance	0.0806**	0.0286**	0.0131***
	(0.0372)	(0.0118)	(0.0035)
N	12,915	13,273	13,166
Panel B. Government medical support			
Government medical support	0.0190	0.0105	0.0003
	(0.0142)	(0.0148)	(0.0128)
N	12,915	13,273	13,166
Panel C. Government pension support			
Government pension support	0.1633***	0.0023	0.0056
	(0.0278)	(0.0291)	(0.0254)
N	12,915	13,273	13,166

Note. Marginal effects are reported for probit method in all columns. Baseline control variables and fixed effects are added in all regressions. IADL = instrumental activities of daily living; CA = cognitive abilities; SRH = self-rated health.
***, **, and * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are shown in parentheses.

Appendix C. Robustness Checks

We conducted some robustness checks for the main results. First, we performed a placebo test. If the main results always hold whether or not elderly individuals had military service, the treatment effect of the health improvement in the elderly veterans loses statistical senses. Therefore, we constructed counterfactuals for placebo testing, and the specific methods are as follows. We randomly assigned a dummy value (0 or 1) to each elderly person, and generate a simulated independent variable *Military**. The total number of elderly veterans is consistent with the original sample. Then, we exploited *Military** to perform 1,000 regressions on the three variables of measuring elderly health outcomes (including all control variables and fixed effects). Table C1 reports the statistical results of placebo tests and we found that regardless of the dependent variables, the proportion of nonsignificant in *Military** is close to 90%. In general, the placebo tests suggested that the simulated treatment effect does not hold, and the main results have great credibility.

Second, although we controlled for some early experience variables of the elderly, such as *Hungry in childhood*, to mitigate potential endogeneity, adopting instrumental variable (IV) method is necessary to further identify causal effects. According to Hou et al. (2020) and Zhao and Guo (2022), we likewise used one-year birth cohort-level military service rates as an IV. Moreover, to further ensure the exogeneity of IV,

Table C1. Placebo Tests.

Significance of *Military**	IADL		CA		SRH		Number of simulations
	Frequency	%	Frequency	%	Frequency	%	
Nonsignificant	887	88.7	902	90.2	912	91.2	1,000
Significant at the 10% level	47	4.7	50	5.0	55	5.5	1,000
Significant at the 5% level	43	4.3	34	3.4	21	2.1	1,000
Significant at the 1% level	23	2.3	14	1.4	12	1.2	1,000

Note. IADL = instrumental activities of daily living; CA = cognitive abilities; SRH = self-rated health.

Table C2. Robustness Checks.

Variables	(1) IADL	(2) CA	(3) SRH
Panel A. IV estimations			
Military	0.8006**	0.5629**	1.5108***
	(0.3122)	(0.2528)	(0.4424)
First-stage *F* value	22.5306	29.7618	18.1374
N	13,065	13,285	13,178
Panel B. Continuous dependent variables			
Military	0.1045**	0.4018***	0.1560***
	(0.0434)	(0.1244)	(0.0245)
N	14,653	14,867	14,746
Panel C. Controlling for some provincial macro variables			
Military	0.0493***	0.0479***	0.0501***
	(0.0121)	(0.0132)	(0.0143)
N	14,653	14,867	14,746
Panel D. Deleting observations that may be involved in the Counterattack against Vietnam in Self-defense			
Military	0.0499***	0.0493***	0.0489**
	(0.0134)	(0.0117)	(0.0201)
N	9,236	9,433	9,302
Panel E. Keeping one observation for each person			
Military	0.0408**	0.0605**	0.0350***
	(0.0187)	(0.0285)	(0.0126)
N	6,915	6,986	7,031

Note. Marginal effects are reported for probit method in all columns in Panels C–E. Baseline control variables and fixed effects are added in all regressions. IADL = instrumental activities of daily living; CA = cognitive abilities; SRH = self-rated health; IV = instrumental variable.
***, **, and * indicate significant at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are shown in parentheses.

we obtained it from another database, the China Household Panel Survey (CFPS), which also records information on respondents' military service. In Panel A of Table C2, we found that the first-stage *F* values are greater than the thresholds of the weak IV issue in all columns, indicating that the IV was selected reasonably (Angrist & Pischke, 2008). In addition, the estimated coefficients of *Military* remain significantly positive, although there is some increase in magnitude, but this does not prevent the analysis of causal effects (Angrist & Pischke, 2008; Zhao et al., 2022). Taken together, using an IV approach to mitigate potential endogeneity problem, our main results were not challenged.

Third, based on previous literature (Shen & Zeng, 2014; Zeng & Vaupel, 2002), we constructed a number of binary variables to describe health outcomes in the elderly. To test that our dependent variables were not sensitive to the way they were measured, we used initial continuous variables for further robustness tests. More specifically, we adopted the IADL and CA scores to measure physical health and cognitive abilities, respectively, using five levels of self-rated health (5 = *very good*; 4 = *good*; 3 = *fair*; 2 = *poor*; 1 = *very poor*). Clearly, in Panel B of Table C2, after replacing the measurement of the dependent variables, our main specification remains unchanged, indicating that main results are insensitive to the measurement of the dependent variables.

Fourth, although we controlled for the provincial fixed effect, we further added several variables to conduct a robustness test considering that some macro variables may also affect the health of elderly individuals. These macro variables at the provincial level included economic growth rate, government social security spending, number of health care facilities, number of health care workers, and number of hospital beds. In Panel C of Table C2, we found that this change had little effect on our results, suggesting that the main specification is satisfactory for controlling these macro characteristic variables.

Fifth, we eliminated extreme observations. According to previous literature, participation in war had a strong negative impact on the health of soldiers (e.g., Bedard & Deschênes, 2006; Dobkin & Shabani, 2009; Landes et al., 2017). We selected the recent main large-scale conflict of China, Counterattack against Vietnam in Self-defense, as a representative. Moreover, we considered 18 to 25 as a possible age range for joining the army (Zhao & Guo, 2022), which means observations born in 1954–1961 were removed. These observations were excluded to eliminate the effect of possible war experience on the main empirical results. In the Panel D of Table C2, we found that the coefficients on *Military* kept significantly positive.

Last, we pooled multiple waves of the CHARLS in the main analyses, and each wave of the survey with the same elderly individual was treated as an independent sample. However, whether the elderly had military service experience before 60 was fixed in each wave, while their health outcomes were likely to change over time. To test the sensitivity of our results, we only kept the first observation for each elderly person (Hou et al., 2020), finding that the results were still robust in Panel E of Table C2.

Acknowledgments

The authors thank the Editor-in-Chief Professor Patricia M. Shields, and three anonymous referees for comments that considerably improved the quality of this paper. The authors appreciate the Institute of Social Science Survey, Peking University for providing the China Health and Retirement Longitudinal Study (CHARLS) data.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: the Fundamental Research Funds for the Central Universities (No. QCDC-2020-21).

ORCID iDs

Chunkai Zhao https://orcid.org/0000-0003-1918-9598
Jianhao Guo https://orcid.org/0000-0003-1470-5960
Jinchen Yan https://orcid.org/0000-0003-2895-2698

Notes

1. China's Law on the Protection of the Rights and Interests of the Elderly stipulates that the starting age standard for the elderly is 60 years old. It is also common internationally to regard an elderly person in developing countries as being 60 years of age or older. Although in some countries, this criterion is 65 years old. In this paper, following Chinese and international practice, we defined individuals aged 60 or older as the elderly.
2. According to unofficial information (relevant official information is not available due to confidentiality), the percentage of passing the medical examination for enlisting in the military is close to 50% in China. In addition, 20% of enlisted men are eliminated during the pre-service education phase for reasons including weak political thinking, misdeed experiences, and psychological vulnerability. Data source: https://www.163.com/dy/article/HHPQRU5R0553BEVE.html and https://www.sohu.com/a/582232738_121187951.
3. According to the Regulations of the Chinese People's Liberation Army on Active Duty Officers' Service, these conditions include holding a division-level position or higher or a senior technical position, having served (or participated in service) for more than 30 years, being 50 or older, and being substantially incapacitated although not having reached the maximum age for active service.
4. These lagging compensations are mainly reflected in special preferential treatment and monetary compensation such as retirement payments. The main reason is that active-duty military personnel in China earn relatively low incomes, while veterans, especially some military officers and those who have served for a long time, often enjoy better benefits after retirement (W. Li & Wang, 2021). From a historical perspective, at the enlarged meeting of the Central Military Commission in 1985, Deng Xiaoping proposed that the

army should be patient, that is, the army should obey and serve the general situation of national modernization. This policy also led to a decline in the treatment of active military personnel (Gao, 2011). This phenomenon was alleviated with the rapid increase in China's economic strength after the reform and opening up.

5. There are 44 military colleges and universities in China, and each year they enroll a proportion of students from both the military and localities. Local students into these schools means enlisting in the military and enjoying military treatment, and all studying and living expenses are free in these military schools.
6. The baseline wave in 2008 was investigated in only two provinces, Zhejiang and Gansu. Since the information on military service in CHARLS was only in the special life course survey in 2014 (Zhao & Guo, 2022). Many observations in 2018 could not be matched. Therefore, we selected three waves of the survey in 2011, 2013, and 2015.
7. In China, people over 60 years old are considered to be elderly. However, international regulations identify those over the age of 65 as the elderly. In the heterogeneity analysis, we follow the international practice, finding that our results keep unchanged for the observations aged 65 years and older.
8. These eight activities include lifting a weight of 5 kg, stooping, kneeling, or crouching, walking continuously for 1 km, getting up from a chair after sitting for a long period, cooking, doing household chores, shopping for groceries, and making phone calls.
9. A *hukou* is a legal document created by the household administration to record and retain basic information about the residents. In China, there are two main categories of *hukou*: rural and urban.
10. The presence of a high number of missing values for some important mechanism variables makes it difficult to adopting the traditional mediating effect model. According to previous literature (e.g., Duflo et al., 2011; Gong et al., 2018), estimating Equations 3 and 4 can also provide evidence for confirming mechanisms.
11. More detailed explanations are shown in Appendix C.

References

Angrist, J. D. (1990). Lifetime earnings and the Vietnam era draft lottery: Evidence from social security administrative records. *American Economic Review*, *80*(3), 313–336.

Angrist, J. D. (1998). Estimating the labor market impact of voluntary military service using social security data on military applicants. *Econometrica*, *66*(2), 249–228.

Angrist, J. D., & Chen, S. H. (2011). Schooling and the Vietnam-era GI Bill: Evidence from the draft lottery. *American Economic Journal*, *3*(2), 96–118.

Angrist, J. D., Chen, S. H., & Frandsen, B. R. (2010). Did Vietnam veterans get sicker in the 1990s? The complicated effects of military service on self-reported health. *Journal of Public Economics*, *94*(11–12), 824–837.

Angrist, J. D., & Krueger, A. B. (1994). Why do World War II veterans earn more than nonveterans? *Journal of Labor Economics*, *12*(1), 74–97.

Angrist, J. D., & Pischke, J. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Bachman, J. G., Freedman-Doan, P., Segal, D. R., & O'Malley, P. M. (2000). Distinctive military attitudes among US enlistees, 1976-1997: Self-selection versus socialization. *Armed Forces & Society*, *26*(4), 561–585.

Barberger-Gateau, P., Fabrigoule, C., Helmer, C., Rouch, I., & Dartigues, J. F. (1999). Functional impairment in instrumental activities of daily living: An early clinical sign of dementia? *Journal of the American Geriatrics Society*, *47*(4), 456–462.

Bedard, K., & Deschênes, O. (2006). The long-term impact of military service on health: Evidence from World War II and Korean War veterans. *American Economic Review*, *96*(1), 176–194.

Benmelech, E., & Frydman, C. (2015). Military CEOs. *Journal of Financial Economics*, *117*(1), 43–59.

Black, W. G. Jr. (1993). Military-induced family separation: A stress reduction intervention. *Social Work*, *38*(3), 273–280.

Bound, J., & Turner, S. (2002). Going to war and going to college: Did World War II and the GI Bill increase educational attainment for returning veterans? *Journal of Labor Economics*, *20*(4), 784–815.

Brockmann, H., & Klein, T. (2004). Love and death in Germany: The marital biography and its effect on mortality. *Journal of Marriage and Family*, *66*(3), 567–581.

Brown, C., & Routon, P. W. (2016). Military service and the civilian labor force: Time-and income-based evidence. *Armed Forces & Society*, *42*(3), 562–584.

Cao, W., Li, L., Zhou, X., & Zhou, C. (2015). Social capital and depression: Evidence from urban elderly in China. *Aging and Mental Health*, *19*(5), 418–429.

Cao, Y., Meng, K., Chen, Z., Wang, P., Li, S., & Tang, N. (2019). From decentralization to integration: The development of social welfare services for disabled veterans in China. *Disability & Society*, *34*(6), 998–1003.

Chase, R. P., McMahon, S. A., & Winch, P. J. (2015). Injury careers after blast exposure among combat veterans deployed to Iraq or Afghanistan. *Social Science & Medicine*, *147*, 309–316.

Cheng, L., Liu, H., Zhang, Y., Shen, K., & Zeng, Y. (2015). The impact of health insurance on health outcomes and spending of the elderly: Evidence from China's new cooperative medical scheme. *Health Economics*, *24*(6), 672–691.

Cutler, D. M., & Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics*, *29*(1), 1–28.

Davies, P. S., Purcell, P. J., & Engelhardt, G. V. (2015). Vietnam-era military service and DI participation. *Economics Letters*, *136*, 5–8.

De Tray, D. (1982). Veteran status as a screening device. *American Economic Review*, *72*(1), 133–142.

Diamant, N. J. (2008). Veterans, organization, and the politics of martial citizenship in China. *Journal of East Asian Studies*, *8*(1), 119–158.

Diamant, N. J. (2011). Conspicuous silence: Veterans and the depoliticization of war memory in China. *Modern Asian Studies*, *45*(2), 431–461.

Diamant, N. J., & O'Brien, K. J. (2015). Veterans' political activism in China. *Modern China*, *41*(3), 278–312.

Dobkin, C., & Shabani, R. (2009). The health effects of military service: Evidence from the Vietnam draft. *Economic Inquiry*, *47*(1), 69–80.

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, *101*(5), 1739–1774.

Elder, G. H. Jr., Clipp, E. C., Brown, J. S., Martin, L. R., & Friedman, H. S. (2009). The life-long mortality risks of World War II experiences. *Research on Aging*, *31*(4), 391–412.

Eynde, O. V. (2015). Military service and human capital accumulation: Evidence from Colonial Punjab. *Journal of Human Resources*, *51*(4), 1003–1035.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.

Fredland, J. E., & Little, R. D. (1985). Socioeconomic status of World War II veterans by race: An empirical test of the bridging hypothesis. *Social Science Quarterly*, *66*(3), Article 533.

Gade, P. A., Lakhani, H., & Kimmel, M. (1991). Military service: A good place to start? *Military Psychology*, *3*(4), 251–267.

Gao, Z. (2011). Army building and nation building: Explorations and insights from the Communist Party of China. *Comparative Economic and Social Systems*, *27*(4), 12–17. (in Chinese)

Gong, J., Lu, Y., & Song, H. (2018). The effect of teacher gender on students' academic and noncognitive outcomes. *Journal of Labor Economics*, *36*(3), 743–778.

Grimard, F., Laszlo, S., & Lim, W. (2010). Health, aging and childhood socio-economic conditions in Mexico. *Journal of Health Economics*, *29*(5), 630–640.

Grossman, M. (2000). The human capital model. In A. Culyer & J. Hewhouse (Eds.), *Handbook of health economics* (pp. 347–408). Elsevier.

Guo, J., & Zhao, C. (2019). Impacting factors of the Chinese military enterprises' capital structure and approaches of importing private capital. *Defence and Peace Economics*, *30*(7), 858–876.

Heaton, W. R., & Lovejoy, C. D. (1986). *The reform of military education in China: An overview, in China's Military Reforms*. Routledge.

Hou, B., Liu, H., & Wang, S. X. (2020). Returns to military service in off-farm wage employment: Evidence from rural China. *China Economic Review*, *59*, 101361.

Johansen, R. B., Laberg, J. C., & Martinussen, M. (2014). Military identity as predictor of perceived military competence and skills. *Armed Forces & Society*, *40*(3), 521–543.

Johnston, D. W., Shields, M. A., & Siminski, P. (2016). Long-term health effects of Vietnam-era military service: A quasi-experiment using Australian conscription lotteries. *Journal of Health Economics*, *45*, 12–26.

Kleykamp, M. (2009). A great place to start? The effect of prior military service on hiring. *Armed Forces & Society*, *35*(2), 266–285.

Laaksonen, M., Rahkonen, O., Martikainen, P., & Lahelma, E. (2005). Socioeconomic position and self-rated health: The contribution of childhood socioeconomic circumstances, adult socioeconomic status, and material resources. *American Journal of Public Health*, *95*(8), 1403–1409.

Landes, S. D., Wilder, J., & Williams, D. (2017). The effect of race and birth cohort on the veteran mortality differential. *Social Science & Medicine*, *179*, 36–44.

Li, T., & Zhang, Y. (2015). Social network types and the health of older adults: Exploring reciprocal associations. *Social Science & Medicine*, *130*, 59–68.

Li, W., & Wang, J. (2021). The improving thinking of the rights and interests guarantee system for veterans in the new era. *Law Science Magazine*, *42*(7), 112–123. (in Chinese)

Lindqvist, E., & Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal*, *3*(1), 101–128.

Liu, X., Engel, C. C., Kang, H., & Cowan, D. (2005). The effect of veteran status on mortality among older Americans and its pathways. *Population Research and Policy Review*, *24*(6), 573–592.

MacLean, A. (2017). Skills mismatch? Military service, combat occupations, and civilian earnings. *Sociological Perspectives*, *60*(2), 229–250.

MacLean, A., & Kleykamp, M. (2021). Generations of veterans: Socioeconomic attainment from World War II to the contemporary era. *Social Science History*, *45*(1), 83–110.

McKinney, W. P., McIntire, D. D., Carmody, T. J., & Joseph, A. (1997). Comparing the smoking behavior of veterans and nonveterans. *Public Health Reports*, *112*(3), 212–218.

McLaughlin, R., Nielsen, L., & Waller, M. (2008). An evaluation of the effect of military service on mortality: Quantifying the healthy soldier effect. *Annals of Epidemiology*, *18*(12), 928–936.

O'Brien, K. J., & Diamant, N. J. (2015). Contentious veterans: China's retired officers speak out. *Armed Forces & Society*, *41*(3), 563–581.

Oh, D., & Berry, F. S. (2021). The individual health insurance mandate and veterans health coverage. *Armed Forces & Society*, *47*(2), 248–275.

Padavic, I., & Prokos, A. (2017). Aiming high: Explaining the earnings advantage for female veterans. *Armed Forces & Society, 43*(2), 368–386.

Routon, P. W. (2014). The effect of 21st century military service on civilian labor and educational outcomes. *Journal of Labor Research*, *35*(1), 15–38.

Sheehan, C. M., & Hayward, M. D. (2019). Black/White differences in mortality among veteran and non-veteran males. *Social Science Research*, *79*, 101–114.

Shen, K., & Zeng, Y. (2014). Direct and indirect effects of childhood conditions on survival and health among male and female elderly in China. *Social Science & Medicine*, *119*, 207–214.

Smith, I. III, Marsh, K., & Segal, D. R. (2012). The World War II veteran advantage? A lifetime cross-sectional study of social status attainment. *Armed Forces & Society*, *38*(1), 5–26.

Spiro, A. III, Settersten, R. A., & Aldwin, C. M. (2016). Long-term outcomes of military service in aging and the life course: A positive re-envisioning. *Gerontologist*, *56*(1), 5–13.

Tan, H., Guo, C., & Zhou, Y. (2017). Cadre children and cognitive function of parents in China: The value of political connection. *Chinese Sociological Review*, *49*(4), 382–406.

Teachman, J. D. (2007). Military service and educational attainment in the all-volunteer era. *Sociology of Education*, *80*(4), 359–374.

Teachman, J. D. (2011). Are veterans healthier? Military service and health at age 40 in the all-volunteer era. *Social Science Research*, *40*(1), 326–335.

Teachman, J. D., & Call, V. R. (1996). The effect of military service on educational, occupational, and income attainment. *Social Science Research*, *25*(1), 1–31.

Torun, H., & Tumen, S. (2016). The effects of compulsory military service exemption on education and labor market outcomes: Evidence from a natural experiment. *Economics of Education Review*, *54*, 16–35.

Wakenhut, R. (1979). Effects of military service on the political socialization of draftees. *Armed Forces & Society*, *5*(4), 626–641.

Waller, M., & McGuire, A. (2011). Changes over time in the "healthy soldier effect." *Population Health Metrics*, *9*(1), 7–15.

Wang, J. (2022). Behind veterans' protests: Passive and piecemeal policy-making in China. *Modern China*, *48*(2), 253–288.

Wang, L., Elder, G. H. Jr., & Spence, N. J. (2012). Status configurations, military service and higher education. *Social Forces*, *91*(2), 397–422.
White, L. O., Schulz, C., Klein, A. M., & von Klitzing, K. (2019). Declining effects of child maltreatment on mental health in the elderly: Another nail in the coffin of the deficit model of aging or a healthy survivor bias? A comment on Rapsey, Scott and Patterson's 25-year longitudinal study. *Journal of Affective Disorders*, *255*, 180–181.
Williamson, V., Harwood, H., Greenberg, K., Stevelink, S. A., & Greenberg, N. (2019). The impact of military service on the mental health of older UK veterans: A qualitative study. *International Journal of Geriatric Psychiatry*, *34*(10), 1412–1420.
Wilmoth, J. M., London, A. S., & Parker, W. M. (2010). Military service and men's health trajectories in later life. *Journals of Gerontology Series B, 65*(6), 744–755.
Wu, X., & Treiman, D. J. (2004). The household registration system and social stratification in China: 1955-1996. *Demography*, *41*(2), 363–384.
Ye, L., & Zhang, X. (2021). The association mechanism between social network types and health-related behaviours among the elderly in rural Hubei Province, China. *International Journal of Health Planning and Management*, *36*(3), 826–846.
Zeng, Y., & Vaupel, J. W. (2002). Functional capacity and self-evaluation of health and life of oldest old in China. *Journal of Social Issues*, *58*(4), 733–748.
Zhang, C. (2015). Military service and life chances in contemporary China. *Chinese Sociological Review*, *47*(3), 230–254.
Zhao, C., & Guo, J. (2022). Are veterans happy? Long-term military service and the life satisfaction of elderly individuals in China. *Journal of Happiness Studies*, *23*(2), 477–508.
Zhao, C., Wu, Y., & Guo, J. (2022). Mobile payment and Chinese rural household consumption. *China Economic Review*, *71*, 101719.

Author Biographies

Chunkai Zhao, PhD in Economics, is an associate professor in the College of Economics and Management, South China Agricultural University. His areas of interest are empirical microeconomics, human capital, and defense economics. He recently focused on assessing the impact of military service on individuals and families in China. His research has been published in such journals as *Journal of Business Research*, *China Economic Review, Journal of Happiness Studies, Child Abuse & Neglect*, and *Teaching and Teacher Education*.

Jianhao Guo has a master's degree in defense economics, graduated from the Army Logistics University of PLA. He is an active-duty officer in China. His research areas of interest are the socioeconomic situation of Chinese veterans and the military industry in China. His research has been published in such journals as *Journal of Happiness Studies, Defence and Peace Economics*, and *China Economic Review.*

Jinchen Yan, majoring in economics, is a PhD candidate of economics at School of Economics, Renmin University of China. His research interests include labor economics and political economies. He has recently focused on military benefits and family economic behaviors in China. He has published an academic paper in Management World (in Chinese).

Air Quality, Atmosphere & Health (2024) 17:1917–1935
https://doi.org/10.1007/s11869-024-01551-0

Understanding government pollution control actions and well-being: Evidence from micro individuals

Chunkai Zhao[1] · **Xing Li**[2]

Received: 18 January 2024 / Accepted: 12 March 2024 / Published online: 3 April 2024
© The Author(s), under exclusive licence to Springer Nature B.V. 2024

Abstract
The Chinese government took many measures to improve air quality in recent years, and the Atmosphere Ten Articles Plan (ATAP) was one of the most significant air pollution control policies. We used the staggered difference-in-differences method to determine the causal link between air quality improvement and people's subjective well-being (SWB) by utilizing the quasi-natural experiment of different Chinese cities implementing the ATAP. We found that ATAP implementations significantly improved people's SWB, as supported by multiple robustness tests, such as excluding spillover effects and self-selection effects and placebo tests. Moreover, potential mechanisms by which ATAP implementations affected people's SWB included improving physical and mental health status, alleviating emotional problems, reducing insomnia, and promoting exercise and outdoor activities. Furthermore, heterogeneity effect results suggested that ATAP implementations were more beneficial to some socially vulnerable groups and this SWB improvement effect was greater for residents living in urban areas. These findings added to the body of literature on how environmental policies affect people's welfare and had important implications for better air quality improvement policies in China and other developing countries.

Keywords Air quality improvement · Subjective well-being (SWB) · Atmosphere Ten Articles Plan (ATAP) · Quasi-natural experiment · China

JEL Classification I31 · Q58 · Q53 · I12

Introduction

Air pollution has become a major environmental challenge for Chinese society. To improve air quality, the Chinese government adopted numerous command-and-control and market-based environmental regulations that were proven to be effective in reducing air pollution (Greenstone et al. 2021; Li & Zhao 2023; Luo & Li 2018; Li et al. 2023; Tanaka 2015). Among them, the most representative measure is the Atmosphere Ten Article Plan (ATAP).[1] In 2013, the State Council of China officially issued the ATAP document to strengthen air pollution control and regulation. The ATAP specified air pollution emission reduction targets for key regions of all provinces, requiring local governments at all levels to make a concentrated effort to reduce air pollution (Greenstone et al. 2021). In addition, the ATAP proposed to significantly improve air quality in key regions of China and heavily polluted cities by 2017 (Song et al. 2019). Similar to other Chinese environmental regulation policies, a considerable body of studies affirmed the effectiveness of the ATAP (e.g., Feng et al. 2019; Fu et al. 2022; Gao et al. 2016; Huang et al. 2018; Luo & Li 2018; Maji et al. 2020; Yu et al. 2022). However, fewer studies examined the relationship between the air quality improvement effect from the ATAP and people's subjective well-being (SWB), and it is noteworthy in an

✉ Xing Li
xingli277@126.com

Chunkai Zhao
ckz_zhao@163.com

[1] College of Economics and Management, South China Agricultural University, No. 483, Wushan Road, Tianhe District, Guangzhou, Guangdong, China

[2] College of Economics and Management, Southwest University, No. 2, Tiansheng Road, Beibei District, Chongqing, China

[1] The ATAP is also known as the "Air Pollution Prevention and Control Action Plan" in China (Luo & Li 2018; Yu et al. 2022).

era when the nations worldwide are focusing on happiness, healthy living, and sustainable development.

There are two categories of literature that are closely related to our study. The first category focused on the effect of environment factors, such as air pollution, on people's SWB (e.g., Gunby & Coupé, 2023; Levinson 2012; Menz 2011; Welsch 2006; Sanduijav et al. 2021). In particular, some literature based on Chinese settings also concluded that air pollution is not conducive to the enhancement of individual SWB (e.g., Dong et al. 2022; Li et al. 2014; Song et al. 2019; Zhang et al. 2017a, b; Zheng et al. 2019). Moreover, the second type of literature relevant to this paper aimed to assess the individual micro-welfare effect of environmental policies, such as environmental regulations, environmental information disclosure, and pollution control measures (Clancy et al. 2002; Liu et al. 2024). In recent years, academics' interest in how environmental policies affect people's SWB has grown (Guo et al. 2020; Song et al. 2019; Wang et al. 2021). However, the SWB enhancing effect due to ATAP implementations, the most central administrative tool for combating air pollution in China (Greenstone et al. 2021), remains little mentioned in previous studies. It would be interesting to explore whether and how the air quality enhancement effect from the ATAP affects Chinese people's SWB.

We expected that the air quality improvement brought by ATAP might enhance people's SWB from the following three mechanisms. First, clearly, air quality improvements contribute to physical health, such as improving overall self-rated health and curbing obesity (Deschenes et al. 2020; He et al. 2016; Liu et al. 2024), which are among the key elements affecting the SWB (Li et al. 2014; Stutzer & Meier 2016; Zhao & Guo 2022). Second, in terms of mental health, reducing air pollution from the ATAP may help decrease negative emotions (Zhang et al. 2017b), and improve insomnia (Chen et al. 2019; Liu et al. 2020). Third, by implementing ATAP, air quality improvements have also facilitated more active participation in outdoor activities (Deschenes et al. 2020; Mittleman 2007; Neidell 2010). In previous studies, outdoor activities and physical exercise were important factors in enhancing individual SWB (Chang et al. 2019; Rasmussen & Laumann 2014).

By using the ATAP as a proxy for air quality improvement and using a difference-in-differences (DID) model, we were able to clearly detect the causal link between air quality improvement and the increase in residents' SWB in China. The ATAP implementation was not consistent across Chinese cities,[2] which provides a highly convincing opportunity to illustrate the effectiveness of Chinese air pollution control policies and underlying micro effects by using this quasi-natural experiment. In addition, the four-waves tracking data, namely the China Labor-force Dynamics Survey (CLDS), not only provided us with matchable city codes, but also recorded individual SWB and a lot of personal and family information, allowing us to explore the mechanisms by which the ATAP affected residents' SWB.

We came to some interesting conclusions and the key contributions of this study can be summed up as follows. First, as far as we were aware, our study is one of the first to look at the effect of ATAP implementations on people's SWB. The baseline results suggested that ATAP implementations resulted in an increase in residents' SWB score. Thus, our study further enriched the literature in two research fields as mentioned above. One is on the determinants of individual SWB from an environmental perspective (e.g., Levinson 2012; Welsch & Kühling 2009; Sanduijav et al. 2021); the second is on the subjective welfare effects of environmental policies (e.g., Guo et al. 2020; Wang et al. 2021).

Second, we revealed potential mechanisms by which the ATAP enhanced individual SWB. In addition to physical and mental health mentioned in prior research (Deschenes et al. 2020; Heyes & Zhu 2019), we further examined several channels, such as obesity, sleep quality, and outdoor activities. More importantly, the unique Chinese setting helped us thoroughly understand how some environmental policies pose stable positive impacts on individuals. There are also important implications for other developing countries caught in severe air pollution.

Third, using ATAP implementations as a proxy variable for air quality enhancement allows us to use the DID model to mitigate potential endogeneity problems. More importantly, we further analyzed the spillover effects of ATAP and residents' self-selection effects to make the causal effects as convincing as possible. These measures may provide methodological insights for subsequent studies. Additionally, the heterogeneity analysis implied that the enhancement effect of ATAP implementations on the SWB of some socially disadvantaged groups is more pronounced, suggesting that the ATAP helped promote social equity and public welfare enhancement to a certain extent (Wang & Palazzo 2021). In particular, we were concerned about the heterogeneity across income levels.

The rest of our research is structured as follows. The backdrop is introduced in "The ATAP in China". The data and variables are described in "Data and empirical strategy" along with an explanation of the estimation method. The empirical findings are presented and explained in "Results". Examining potential mechanisms is "Mechanisms". The paper is concluded in the last Section.

[2] In this paper, cities are used to refer to administrative units like prefecture-level cities and provincial municipalities. For most cities in China, they include both the urban areas and rural areas under their jurisdiction.

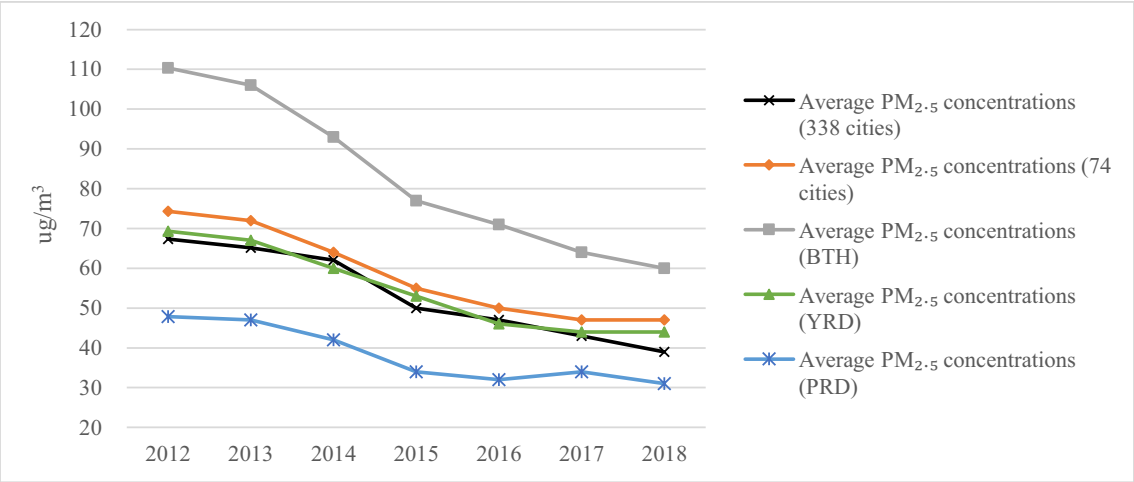


Fig. 1 Overall $PM_{2.5}$ concentrations during ATAP implementations in China (2013–2018). Source: China Ecological and Environmental Status Bulletin (2013–2018) and Atmospheric Composition Analysis Group of Dalhousie University

The ATAP in china

To control air pollution, on September 10, 2013, the State Council of China officially released the ATAP. This plan overall included ten key actions,[3] which covered not only source governance and strengthening comprehensive pollution control, but quantifiable targets for the major pollutant, $PM_{2.5}$. Specifically, by 2017, the average $PM_{2.5}$ concentrations would be reduced by no less than 10%; for the three core regions, Beijing-Tianjin-Hebei (BTH), Yangtze River Delta (YRD), and Pearl River Delta (PRD), the targets were 25%, 20% and 15%, respectively.

Due to the top-down administrative system, China fully implemented the ATAP. Soon, the overall air quality across the country subsequently improved. The average $PM_{2.5}$ concentrations in China between 2012 and 2018 are shown in Fig. 1. This data suggests that $PM_{2.5}$, the principal air pollutant, decreased quickly in all 338 cities, 74 main cities, and the three core regions.[4] Specifically, in 2017, the average $PM_{2.5}$ concentrations in 338 cities was 43 ug/m^3, a 30.6% decrease compared to 2014, further reducing to 39 ug/m^3 in 2018. Figure 1 indicates that the ATAP may contribute to the improvement of air quality in Chinese cities.

Further, we showed the air pollution situation during the ATAP in three core regions of China. The BTH in particular was one of China's most heavily contaminated areas. In Fig. 2, the average $PM_{2.5}$ concentrations in the three core regions decreased by 39.62%, 34.32%, and 27.66%, respectively, far exceeding the targets of 25%, 20%, and 15% set at the beginning of ATAP implementations. With relative decreases in average $PM_{2.5}$ concentrations of 34.83%, 37.10%, and 33.96%, their key cities Beijing, Shanghai, and Guangzhou likewise met the policy goals. Moreover, in Fig. 3, we provided the overall change between 2012 and 2018 in two other aspects: days with heavy pollution ($PM_{2.5}$ concentrations > 150 ug/m^3) and hours with severe pollution ($PM_{2.5}$ concentrations > 300 ug/m^3). Apparently, air pollution was significantly reduced in most Chinese cities.

Last, we depicted the levels of additional air pollutants over years in Fig. 4. Similar to $PM_{2.5}$, concentrations of other air pollutants also decreased during the ATAP implementations. Specifically, in 2018, concentrations for PM_{10}, SO_2, CO, and NO_2 were 74, 14, 1.5, and 29 ug/m^3, respectively. Compared to 2014, they decreased by 32.38%, 60.00%, 31.82%, and 23.68%, respectively. Taken together, according to these statistics, the ATAP's beneficial policy impact on $PM_{2.5}$ and other air pollutants was highly evident, which is consistent with the results of earlier research (e.g., Feng et al. 2019; Gao et al. 2016; Huang et al. 2018; Luo & Li 2018; Maji et al. 2020; Yu et al. 2022). The outstanding role of the ATAP in improving China's air pollution is difficult to be questioned, which provides a realistic basis for considering ATAP as a proxy for air quality improvement in this paper.

[3] This is the reason why this plan is called the "Atmosphere Ten Articles."

[4] The 2018 China Ecological and Environmental Status Bulletin no longer counted 74 main cities, but 161 cities instead. Thus, we manually counted the average $PM_{2.5}$ concentrations for 74 cities in 2018. These 74 main cities are the first batch of key monitoring cities, which cover province-level municipalities, sub-provincial cities, provincial capitals, planned cities, and regional key cities (like the BTH city cluster). The second batch of 87 cities were further included in the scope of key monitoring.

Fig. 2 Reduction in $PM_{2.5}$ concentrations in key regions and their core cities during ATAP implementations in China (2013–2018). Source: China Ecological and Environmental Status Bulletin

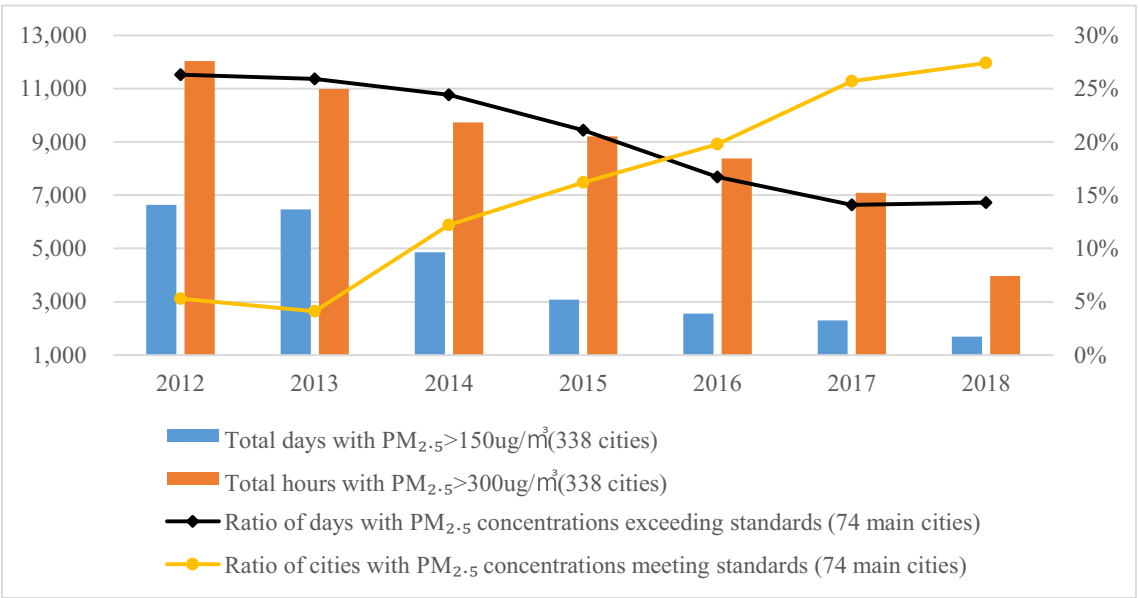


Fig. 3 Heavy air pollution and the ratio of cities meeting air quality standards during ATAP implementations in China (2012–2018). Source: China Ecological and Environmental Status Bulletin, daily city air quality data published by the Ministry of Ecology and Environmental of China, and Atmospheric Composition Analysis Group of Dalhousie University

Data and empirical strategy

Data source

The data for this paper came from four-waves of CLDS, which ran from 2012 to 2018. Based on a random stratified sampling method, the biennial CLDS takes the laborer aged 15–64 as its main target and profoundly reflects the socioeconomic changes in China (Liu et al. 2021; Zhao et al. 2023). Moreover, the CLDS contains data from three strata, namely individuals, households, and communities/villages. Taking the fourth wave in 2018 as an example, the CLDS sample covered 29 provinces in China, consisting of 16,537 individuals from 381 communities/villages. In summary, the CLDS offers a top-notch fundamental database for theoretical study that is empirically focused and for research on public policy.

Moreover, the CLDS provides matchable city names and international codes and the survey information on people's SWB in multiple aspects, which facilitates our study of the welfare effects of ATAP implementation (Lu et al. 2020). Also, the CLDS covered high-quality data on basic

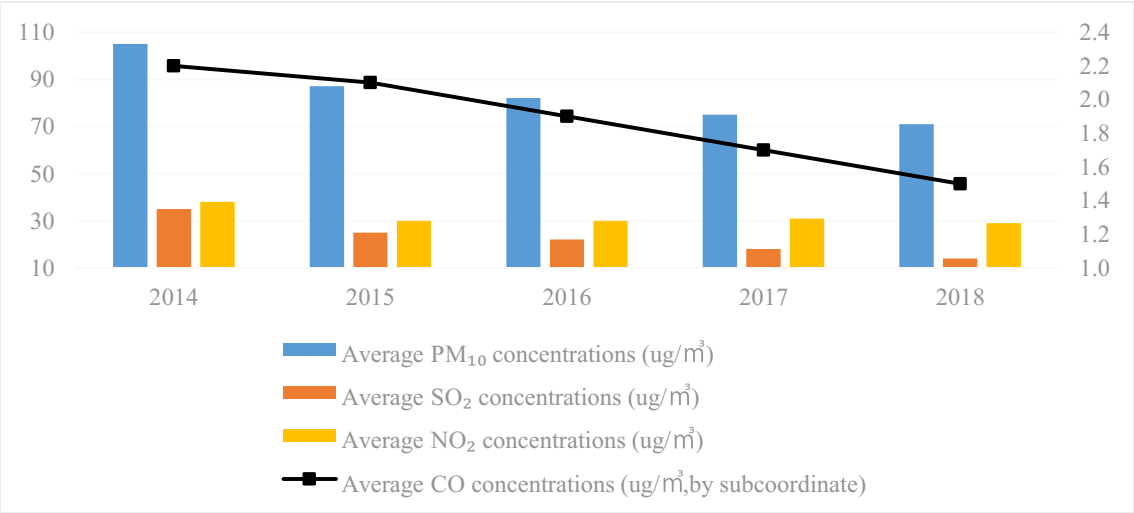


Fig. 4 Other air pollutants (338 cities) during ATAP implementations in China (2014–2018). Source: China Ecological and Environmental Status Bulletin

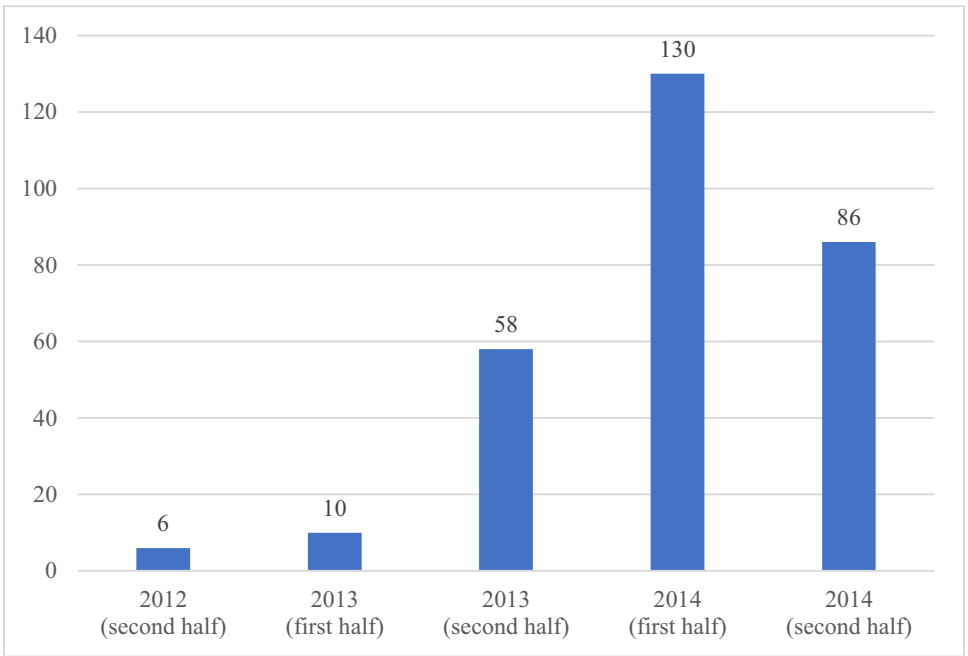


Fig. 5 The number of cities implementing the ATAP counted on a semi-annual basis

demographics, physical and mental health, economic status, and social networks, which is useful for determining the potential mechanisms by which the ATAP affected residents' happiness.[5]

The model and variables

Using the quasi-natural experiment of ATAP implementations, we were able to identify the causal effect of air quality improvements on individual SWB. Although the State Council of China issued the ATAP on September 10, 2013, the timing of formal enactment of specific implementation programs is not consistent across local governments, which provided an ideal opportunity to construct a DID design. Specifically, the BTH city cluster began formally implementing the ATAP in September 2013, while some southern cities formally implemented it in 2014. For example, Zhanjiang and Maoming in Guangdong Province implemented their plans on December 2014. As shown in Fig. 5, before the

[5] The SWB and happiness are synonymous expressions in this paper.

second half of 2013,[6] 58 cities implemented the ATAP; 130 cities did so in the first half of 2014 and 86 cities in the second half of that year.

Due to differences in the timing of policy implementations, while some cities have begun implementing the ATAP, the ATAP has not yet affected others, and residents living in non-pilots can serve as a control group.[7] As for the treatment group, it consisted of people living in pilot cities with formal implementation of the ATAP. As mentioned above, the ATAP was implemented in gradual batches, which led to the possibility that a previous control group could become a later treatment group. In this case, the traditional DID model (typically set to Post*treated) is not applicable (De Chaisemartin & d'Haultfoeuille 2020). In our study, the staggered DID model can be written as:

$$SWB_{ict} = \alpha_0 + \alpha_1 ATAP_{ct} + \alpha_2 X_{it} + \alpha_3 PM2.5_{ct-1} + \delta_c + \varphi_t + \mu_{ict} \quad (1)$$

where subscript i indicates individual, c represents city, and t denotes year. SWB_{ict} refers to the SWB of individual i from city c in survey year t. The CLDS database recorded the SWB status of respondents through the question, "Overall, do you think you are happy with your life." Following the conventional practice of the previous literature (e.g., Zhang et al. 2017a; Zhao & Guo 2022), we divided the answers into five levels from very unhappy to very happy, so we assigned the key explained variable *SWB* the integers 1 to 5 to the answers accordingly.

$ATAP_{ct}$ indicates that a city c has implemented the ATAP in year t. In the framework of the DID model, we defined the core explanatory variable *ATAP* as 1 if the city formally implemented the ATAP from a given year and 0 otherwise. The original data comes from the official implementation time of the ATAP promulgated by the official website of each province and city. In addition, the implementation time of the ATAP was between 2012 and 2014 and can be specific to months. Considering that the CLDS survey was conducted in July or August, it is easy to be clear whether ATAP implementation preceded the survey in each wave. Based on the practice of most of the previous literature and considering the policy lag effect (Acemoglu et al. 2016; Lipp 2007), cities that were implemented in the first half of the year were deemed to be in effect for the current year, while those in the second half of the year were considered to be in effect for the next year.[8] For instance, as depicted in Fig. 5, in cities implemented in the second half of 2013, we set the variable *ATAP* to 1 in 2014 and subsequent years.

X_{it} represents a series of individual-level variables to control for differences caused by individual characteristics. More specifically, based on previous studies (e.g., Levinson 2012; Li & Wang 2022; Wang et al. 2021; Zhang et al. 2017b; Zhao & Guo 2022), we controlled for gender, ethnicity, age, *hukou*, education, marriage, residence, and income. In particular, $PM2.5_{ct-1}$ was also included in the control variable set, which means the one-year lag average $PM_{2.5}$ concentrations of city c. This variable was added for the following reasons. On the one hand, since the implementation order for the ATAP may be related to local air quality (Gao et al. 2016; Li et al. 2022), and air pollution directly affects people's SWB (Chang et al. 2019; Levinson 2012; Zheng et al. 2019), controlling for one-year lag average $PM_{2.5}$ concentrations can effectively reduce endogeneity from omitted variable bias. On the other hand, $PM_{2.5}$ concentrations are the major air pollutants that local governments focused on in the ATAP and have good representativeness and data

Table 1 Descriptive statistics

Variables	Obs	Mean	S.D	Min	Max
Explained variable					
SWB	30,918	3.8191	1.0015	1	5
Explanatory variable					
ATAP	30,918	0.5686	0.4953	0	1
Control variables					
Gender	30,918	0.4662	0.4989	0	1
Ethnic	30,918	0.8883	0.3151	0	1
Age	30,918	45.7990	12.448	15	96
Age2	30,918	30.2673	14.3327	2.25	92.16
Hukou	30,918	0.1972	0.3979	0	1
Education	30,918	2.8481	1.1953	1	6
Marriage	30,918	0.9188	0.2731	0	1
Residence	30,918	0.3018	0.4591	0	1
Income	30,918	10.0857	5.3346	0	15.6012
$PM_{2.5}$	30,918	43.888	17.679	8.0788	94.7913

[6] These 16 cities were implemented the ATAP earlier than September 2013, because the air pollution control program had already started in the "12th Five-Year Plan for Prevention and Control of Air Pollution in Key Regions," which is considered as the predecessor of the ATAP. Therefore, according to Luo and Li (2018), these cities' actions for implementing the ATAP predated the release of the State Council of China. In the analysis below, we removed these cities for a robustness test.

[7] In the later robustness tests, we further confirmed the validity of the treatment and control groups by excluding the spillover and self-selection effects.

[8] One of the main advantages of assigning the *ATAP* in this way is that it allows for better parallel trend tests. Of course, this relatively crude setup also faces risks; therefore, in the empirical analysis, we perform two robustness tests to confirm that this setup method has no great impact on our main results. In addition, since the CLDS investigation event is in July or August, this setting does not cause a time mismatch between ATAP and people SWB.

availability.[9] The definitions of these variables are shown in Table 9 of the Appendix, and Table 1 reports the descriptive statistics.

δ_c and ϕ_t are city fixed effect (FE) and year FE, respectively, which allows us to use a reduced form of the DID model (e.g., Duflo 2001; Gruber et al. 1999). Since the ATAP is implemented at the city level, the inclusion of year and city FEs helps to control for city characteristics that do not vary over years and temporal characteristics that do not vary over regions. Last, ε_{ict} represents the random error term.

It is important to emphasize that since our explained variable *SWB* is an ordered variable, we exploited both linear methods and nonlinear ordered logit models in baseline results, according to a common practice in most previous studies (e.g., Dumludag & Gokdemir 2022; Zhang et al. 2017a; Zhao & Guo 2022). However, since different model adoptions do not bias estimates on happiness, we mainly used linear estimates in the subsequent regressions (Wang et al. 2021; Zhang et al. 2017b).

Results

The effect of the ATAP on people's SWB

Using the quasi-natural experiment of ATAP implementations in Chinese cities, we conducted empirical tests on the causal association between an increase in air quality and people's SWB in Table 2. Due to the fact that our key explanatory variable is city-specific, the city-level clustering standard errors were adopted. In column (1), some individual characteristics and city and year FEs were controlled for, and we discovered that the estimated coefficient on *ATAP* was significantly positive. Further, in column (2), we additionally controlled for *Income*, a vital variable affecting residents' SWB, finding that the results remain unchanged. The one-year lag average $PM_{2.5}$ concentrations at the city level were added to column (3), and after doing so, we found that the result that the ATAP promoted residents' SWB remains robust. Moreover, considering that *SWB* is an ordinal variable, we further used the ordered logit model. The marginal effects when *SWB* is set to 5 are reported in Column (4) of Table 2,[10] and the conclusion stays the same.

Our results are not only statistically significant, but also quite convincing in terms of economic significance. Specifically, according to our preferred column (3), we found that air quality improvement effects from ATAP implementations could increase people's SWB score by an average of 0.5771 points, which amounts to a 15.11% increase in individual SWB relative to the average value (3.8191). Converting the magnitude using standard deviations, ATAP implementations enhanced the SWB score by 0.2854 standard deviations. In addition, we found that our estimations were fairly economically significant when compared to earlier studies examining the effect of environmental regulations on people's happiness (Guo et al. 2020; Pan 2023; Wang et al. 2021). Thus, our baseline results confirmed the welfare-enhancing effect of ATAP implementations in China.

Table 2 The effect of air quality improvement on individual SWB

	OLS			Ordered logit
	(1)	(2)	(3)	(4)
ATAP	0.8414***	0.8142***	0.5771***	0.2140***
	(0.0314)	(0.0320)	(0.0830)	(0.0003)
Gender	0.0364***	0.0365***	0.0365***	0.0108***
	(0.0129)	(0.0129)	(0.0129)	(0.0035)
Ethnic	-0.0456	-0.0455	-0.0455	-0.0182
	(0.0456)	(0.0456)	(0.0456)	(0.0136)
Age	-0.0001	-0.0001	-0.0001	0.0001
	(0.0007)	(0.0007)	(0.0007)	(0.0002)
Age2	0.0002	0.0002	0.0002	0.0178**
	(0.0004)	(0.0004)	(0.0004)	(0.0080)
Hukou	0.0540**	0.0541**	0.0541**	0.0272***
	(0.0273)	(0.0272)	(0.0272)	(0.0026)
Education	0.0977***	0.0977***	0.0977***	0.0019***
	(0.0089)	(0.0089)	(0.0089)	(0.0005)
Marriage	0.1616***	0.1617***	0.1617***	0.0062
	(0.0278)	(0.0278)	(0.0278)	(0.0098)
Residence	0.0333	0.0333	0.0333	0.0428***
	(0.0354)	(0.0354)	(0.0354)	(0.0077)
Income		0.0056***	0.0056***	0.4502***
		(0.0017)	(0.0017)	(0.0460)
$LPM_{2.5}$			-0.0184***	-0.0016*
			(0.0041)	(0.0009)
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.1033	0.1171	0.1171	
Observations	30,918	30,918	30,918	30,918

When *SWB* equal to 5, marginal effects are reported in column (4) using the ordered logit model. Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering

Parallel trend tests and spillover effects

To satisfy the parallel trend assumption is the primary need for applying DID models. However, the parallel trend test of staggered DID is different from the traditional DID model

[9] Average $PM_{2.5}$ concentration data for each city is from the Washington University in St. Louis.

[10] The estimation results for the other values of SWB are shown in Table 10 in the Appendix.

Table 3 Parallel trend tests

	OLS	Ordered logit
	(1)	(2)
pre3	-0.1786	-0.0027
	(0.1506)	(0.0154)
pre2	0.1871	0.0644
	(0.1799)	(0.0720)
current	0.2067***	0.1016***
	(0.0470)	(0.0344)
post1	0.1189***	0.0948***
	(0.0333)	(0.1016)
post2	0.1711***	0.0061***
	(0.0434)	(0.0019)
post3	0.5325***	0.2534**
	(0.1783)	(0.1078)
post4 +	0.7945**	0.1984
	(0.3218)	(0.1393)
Control variables	Yes	Yes
City FE	Yes	Yes
Year FE	Yes	Yes
Observations	30,918	30,918

When *SWB* equal to 5, marginal effects are reported in column (2) using the ordered logit model. Pre means the period before the ATAP was implemented and post indicates the period after the ATAP was implemented. The pre 1 period was used as a benchmark to avoid full covariance and period post 5 are summarized to post 4, marked as post 4 + . Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering

(Chaisemartin & d'Haultfoeuille 2020). Since the time of policy implementations is not unique, it is not possible to use a fixed policy implementation period as the base point. In this paper, the time axis is divided into the periods before and after the ATAP implementations, and the period when the policy is implemented is the critical. The change in the average treatment effect before and after the ATAP implementations is used to test whether the parallel trend is met.

With the help of event analysis method, results for parallel trend tests are shown in Table 3. It is clear that the estimated coefficients on pre1 and pre2 were insignificant, indicating that there was no significant difference in SWB scores between the treatment and control groups prior to ATAP implementations. In other words, the prerequisites for the use of the DID model in this study are valid. In addition, after ATAP implementations, people in the treatment group had significantly higher SWB scores compared to these in the control group, which means with the implementation of the ATAP and the improvement of air quality, people's well-being has further increased.

Another central assumption of the DID approach is that there are no spillover effects, i.e., the control group cannot be affected by policy implementations (Lu et al. 2019; Miguel & Kremer 2004). Spillover effects lead to the stable unit treatment values assumption of the DID no longer holds (Callaway and Sant'Anna 2021). Similar to Lu et al. (2019), we adopted two identification strategies to examine the potential spillover effects. First, we explored the impact of ATAP implementations on neighboring cities. In columns (1) and (3) of Table 4, we found that ATAP coefficients are statistically insignificant, showing that ATAP implementations had a minimal impact on the air quality and SWB of neighboring city residents. Second, we studied the impact of ATAP on other cities in the same province and the results in columns (2) and (4). Similarly, we found that there is little correlation between ATAP implementations and $PM_{2.5}$ or overall happiness in these cities, implying that spillover effects are non-existent. Thus, these estimates further suggest that it is quite reasonable to use the DID approach by excluding potential spillover effects in this study.

Robustness tests

Although baseline empirical results support the conclusion that ATAP implementations enhanced people's SWB, they

Table 4 Spillover effects

	Average $PM_{2.5}$ concentrations in neighboring cities	Average $PM_{2.5}$ concentrations of other cities in the same province	The SWB of residents in neighboring cities	The SWB of residents in other cities in the same province
	(1)	(2)	(3)	(4)
ATAP	-0.3332	-0.0023	0.6417	0.6910
	(0.4763)	(0.0134)	(0.7076)	(0.7166)
Control variables	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	1,442	1,442	19,213	17,354

Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering

are still influenced by a number of confounding factors. To the greatest extent possible, we performed certain robustness tests to show the dependability of our baseline findings. The results are shown in Tables 11, 12, 13 and 14 in the Appendix.

Alternative specifications

We conducted robustness tests by using some alternative specifications. First, since the meaning of life satisfaction and SWB are very similar and are interchangeable in some studies (e.g., Levinson 2012; Welsch 2006; Zhang et al. 2017b; Zhao & Guo 2022), we used life satisfaction as a proxy in the robustness test. Similarly, according to the answers to the question, "Overall, are you satisfied with your living situation," we set the variable *Life satisfaction* as integers from 1 to 5, respectively, with a larger number indicating higher life satisfaction. The coefficient on *ATAP* is still significantly positive at the 1% level in column (1) of Table 11. Therefore, the choice of the core explained variable has little impact on our primary findings.

Second, considering that air pollution and people's SWB are closely related to seasons, climate, and weather (Li & Zhao 2023; Welsch 2006; Zhang et al. 2017b), we controlled for the survey month and weather FEs. Related information was recorded by CLDS investigators. The results in column (2) of Table 11 show that even after controlling for these FEs, the ATAP estimations are still reliable. Moreover, in column (3), we used the individual FE instead of city FE for further robustness tests. We found that the sample size decreases by 85% after controlling for individual FE, suggesting that some samples that do not vary by year may have been deleted. Fortunately, the estimated coefficient of *ATAP* remains significantly positive.

Third, in the main specification, we set the core explanatory variable by whether the month in which the policy is implemented falls in the first or second half of the year. Although this approach is common in the previous literature (e.g., Li & Zhao 2023; Lipp 2007), its crudeness cannot be denied. To avoid the potential risk in our baseline regressions, we performed two robustness tests. First, those cities that implemented ATAP in the last quarter were considered as the treatment group for the following year. Second, we used total treatment months/12, rather than treatment dummies (0 or 1), to measure the treatment effects of ATAP.[11] Results in columns (4) and (5) of Table 11 indicated that these changes do not have an impact on our main results.

Fourth, although city FE are added in prior regressions, we further controlled for some time-varying city characteristic factors for further robustness tests because our DID estimations are based on the city level. Specifically, the macro variable set covers economic growth rate, GDP per capita, fiscal expenditure, population density, industrial structure, financial development, fiscal revenue, and average housing prices.[12] In column (6) of Table 11, we can see that controlling for these macro variables has a small effect on the baseline results, implying that our main specification remains satisfactory after capturing macro city characteristics.

Last, to isolate the impact of air quality improvement on people's happiness and obtain a more precise estimate of the effect of ATAP policy on people's SWB, we used the instrumental variable (IV) method for regression. Specifically, we used ATAP as an IV for air quality. In terms of the correlation assumption, the implementation of ATAP can directly affect air quality as highlighted above; in terms of exclusion restriction assumption, the ATAP is implemented by government departments to improve air quality and has no direct relationship with residents' SWB. Columns (7) and (8) of Table 11 report the results of two stages. We can see that ATAP implementations did significantly improve the air quality, which is similar to previous literature (e.g., Feng et al. 2019; Gao et al. 2016; Maji et al. 2020; Yu et al. 2022). Additionally, the second stage results indirectly suggest that the ATAP is effective in enhancing residents' SWB by improving air quality.

Placebo tests and expectation effects

To rule out the possibility that respondents' increased SWB was due to other pre-implemented policies, we set implementations of the ATAP for each city one year earlier to conduct a placebo test (Chen et al. 2020). Column (1) of Table 12 shows that in the counterfactual scenario when the time of the policy is advanced by one year, the impact of the ATAP on SWB is no longer significant. These results indicate that other prior policies may not cause the increase in residents' SWB, and our baseline results are robust. Another placebo test came from ignoring the specific month of ATAP implementations and directly considering the year as the time of policy implementations (i.e., disregarding the first and second half of the year and the policy time lag). Similarly, in column (2) of Table 12, we noticed that the coefficient on *ATAP* is also not significant, implying that this artificially constructed treatment effect is also not valid.

[11] In this case, only those cities that happen to implement ATAP in that year will be treated as set to 1 (full treatment). For example, if a city implemented ATAP in July 2013, the treatment variable for this city in 2013 would be 0.5.

[12] Industrial structure is measured by the share of tertiary industry in the secondary industry; financial development is determined by the ratio of total deposits and loans to GDP. Housing prices data are obtained from https://www.gotohui.com, and the China City Statistical Yearbook is where the remaining macro variables.

Self-selection and expectation effects

In addition to the spillover effect verified earlier, another concern is from residents' self-selective behaviors (Callaway and Sant'Anna 2021; Huber & Steinmayr 2021). More specifically, the air quality improvement due to ATAP implementations may cause some people to migrate to the pilot cities (Chen et al. 2022; Fu et al. 2024), leading to a possible "contamination effect" of the treatment group. To exclude this potential self-selection effect, we firstly examined the relationship between ATAP implementations and total population of cities. Column (1) of Table 13 told us that ATAP implementations did not result in a mass migration of population from other cities to pilot cities. In addition, we further removed all floating populations in column (2). Clearly, there is no shock to the conclusion that ATAP implementations promoted people's SWB. Taken together, we considered that our main results are very little influenced by residental self-selection effects.

Second, although we controlled for the lagged one-year average $PM_{2.5}$ concentrations at the city level, the air pollutant of greatest concern to the ATAP, ATAP implementations in different batches are not random, which may still interfere with our main DID estimates. To make the treatment and control groups more comparable (Delfgaauw & Dur 2010; Wang et al. 2022), we constructed the DID model based on the propensity score matching (PSM) method. Results in column (3) of Table 13 using the PSM-DID approach suggest that the ATAP still improved individual SWB.

Third, to control for the self-selection effect of policy implementations, we added an interaction term between the air quality characteristics ($PM_{2.5}$) in the year before policy implementation (in 2012) and the year trend in column (4) of Table 13, further excluding the non-randomness of ATAP implementation in cities. It can be observed that after controlling for the self-selection effect of policy implementation, the ATAP still significantly promotes people's SWB.

Fourth, ATAP implementations in different cities have been proposed after a long preparation process. Governments and residents in pilot cities may have expectations and preparations for this policy. Such expectation effect may also result in biased estimates. Therefore, we further included a dummy variable for the year prior to ATAP implementations. Not surprisingly, our estimates in column (5) of Table 13 are consistent with prior main results, suggesting that the baseline results are satisfactory in capturing the expectation effect.

Disruptive policies

Removing the effects of some disruptive policies is also very common in the previous literature by using DID methods (e.g., Chen et al. 2020; Qian 2008). We considered some other environmental policies that may have impacts on the air pollution of cities and residents' SWB. Given that China's carbon reduction and emissions policies are synergistic and that the sources of CO_2 and most air pollutants are the same, we firstly removed the two most important carbon reduction policies in China, low carbon cities and carbon market pilot policies (Wang et al. 2022). We found that our main estimates remain unchanged after removing these cities associated with low carbon policies in columns (1) and (2) of Table 14. Another study concluded that national forest city constructions also contributed to air quality in China (Li & Zhao 2023; Liao et al. 2021), so we further excluded the effect of the forest city policy in column (3). Clearly, the DID estimate closely matches the baseline findings in Table 2. Moreover, scholars have recently begun to focus on the environmental effects of innovation and digital construction in China (e.g., Chen 2023; Li & Zhao 2024). Here we selected two relevant disruptive policies, innovative-oriented cities and smart cities. In columns (4) and (5), we noted that the coefficients on *ATAP* remain unchanged after excluding the effects of these two policies. Finally, in column (6) of Table 14, we removed cities that implemented the ATAP before September 2013. As mentioned earlier, some cities started implementing the ATAP document before it was officially released. We found our results are also insensitive to this change.

Heterogeneity effects

The heterogeneous effects of the ATAP on people's SWB were examined from multiple aspects in this subsection. First, since there is a gradient in the SWB for urban and rural areas in China (Zhao & Guo 2022), we tested the urban–rural heterogeneity effect.[13] As shown in the first two columns of Table 5, we noted that ATAP implementations could improve the SWB for urban residents, but the impact on rural residents is insignificant. An intuitive reason is that air pollution is more severe in urban China and ATAP implementations could be more effective in enhancing residents' happiness by improving air quality. Instead, better air quality in rural areas and more limited exposure of people to severe air pollution (Chen et al. 2019), and they might benefit less from the marginal effects of air quality improvements from the ATAP.

Second, we further investigated the age heterogeneity, taking into account that the middle-aged and elderly, as a vulnerable population, are more likely to be harmed by air pollution in terms of disease and unhappiness. In line with earlier literature (Lim et al. 2012; Zhao & Guo 2022), the youth and the middle-aged and elderly groups are divided, using 45 years

[13] We used the respondents' place of residence (urban or rural) as a criterion for classification.

Table 5 Heterogeneity effects

	Rural	Urban	The youth	The middle-aged and elderly	Relatively high income	Relatively low income
	(1)	(2)	(3)	(4)	(5)	(6)
ATAP	0.0028	0.2035***	0.0361	0.2915***	0.1247**	0.9392***
	(0.0365)	(0.0217)	(0.0376)	(0.0956)	(0.0659)	(0.1611)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,332	21,586	15,657	15,261	15,587	15,331

Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering

old as the cutoff.[14] Clearly, in columns (3) and (4) of Table 5, the coefficients on *ATAP* are only significant in the later column, meaning the positive impact is for the middle-aged and elderly residents. These results suggest that the air quality improvement from ATAP implementations may have an inclusive effect, which contributes more to the increase in welfare of the middle-aged and elderly vulnerable groups.

Third, we tested for income heterogeneity effects. Specifically, considering that respondents' median income was what we used to divide the sample into two groups (19,800 yuan, RMB). The last two columns of Table 5 show that *ATAP* estimates are both significantly positive, but again, the coefficient is larger in the low-income group. As before, the ATAP is more favorable to the SWB of socially disadvantaged groups, implying the potential benefits of this public environmental policy for low-income groups in China (Zhang et al. 2017b).

Mechanisms

Given our main findings that ATAP implementations effectively improved residents' SWB, we explored multiple potential mechanisms. We classified these channels into the following categories: physical health, emotion and mental health, and exercise and outdoor activities.

Physical health

The harm of air pollution on people's physical health has been widely recognized by the academics. As a core policy dedicated to improving air quality, the primary mechanism by which the ATAP promoted SWB stems from physical health effects. A valid measure of physical function and illness condition, self-rated health is a person's overall subjective estimate of his or her own health status (Grimard et al. 2010). Based on the responses to the question, "What do you think of your current physical health?," we gave values integers from 1 to 5, where 1 means very unhealthy and 5 refers to very healthy in sequence. Results in column (1) of Panel A in Table 6 indicated that ATAP implementations significantly improved individual self-rated health. As shown in Panel B, there is a positive correlation between self-rated health and SWB, as highlighted in prior research (e.g., Li et al. 2014; Liu et al. 2024; Song et al. 2019; Zhao & Guo 2022), Therefore, there are good reasons the ATAP enhanced residents' SWB by improving their physical health.

Table 6 The mechanism of physical health and obesity

	(1)	(2)	(3)	(4)
Panel A	Self-rated health	BMI	Overweight	Obesity
ATAP	0.3174***	-0.3669***	-0.1698***	-0.4913***
	(0.0618)	(0.0744)	(0.0027)	(0.0139)
Panel B	SWB			
Self-rated health	4.0817***			
	(0.0470)			
BMI		-4.0369***		
		(0.0294)		
Overweight			-3.9231***	
			(0.0584)	
Obesity				-3.9039***
				(0.0558)
Control variables	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	26,121	26,049	26,049	26,049

The explained variable is *SWB* for all columns in Panel B. Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering

[14] In the full sample, the mean age was similarly close to 45 years old.

Table 7 The mechanism of emotion and mental health

	(1)	(2)	(3)	(4)
Panel A	Emotional problems	Emotional stability	Mental health	Insomnia
ATAP	-0.1465***	-0.1610***	-0.7559***	-0.1693**
	(0.0047)	(0.0033)	(0.0094)	(0.0895)
Panel B	SWB			
Emotional problems	-1.1263***			
	(0.0176)			
Emotional stability		-1.6346***		
		(0.0115)		
Mental health			-1.6687***	
			(0.0155)	
Insomnia				-3.9350***
				(0.0343)
Control variables	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	25,468	25,039	30,918	8,821

The explained variable is *SWB* for all columns in Panel B. Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering

Additionally, since obesity is associated with numerous chronic diseases and unhealthy outcomes (Seaton et al. 1995; Deschenes et al. 2020), we further considered the effect of air quality improvements brought about by the ATAP on individual obesity. Referring to Deschenes et al. (2020), we constructed three indicators: *BMI*, *Overweight*, and *Obesity*. Specifically, we calculated the continuous variable *BMI* by dividing the respondent's weight (kg) by the square of height (m). Two dummy variables, *Overweight* and *Obesity*, were set to 1 if the BMI was over 25 and over 30, respectively, and 0 otherwise (Deschenes et al. 2020).[15] Not surprisingly, as reported in columns (2)-(4) of Panel A in Table 6, we could see that the ATAP significantly reduced BMI, leading to a lower probability of being overweight and obese. In Panel B of Table 6, we can see that that obesity may be detrimental to people's SWB, which is in line with some previous studies (Li et al. 2018; Stutzer & Meier 2016). In sum, our results suggest the ATAP promoted individual physical health and SWB by reducing obesity.

Emotion and mental health

Cohn et al. (2009) and Sanchez and Vazquez (2014) argued that positive emotions could promote people's SWB. Also, this extensively validates the adverse effect of air pollution on individual emotion (Lim et al. 2012; Zheng et al. 2019). We empirically tested how the ATAP improved individual SWB through positive emotions in two ways. The CLDS database detailed the emotional and psychological problems of respondents in several ways.

First, according to the question, "Are you affected by emotional problems at work or in other daily life?", we set five consecutive positive integers from 1 to 5, with higher values indicating greater vulnerability to emotional problems. Column (1) of Panel A and Panel B of Table 7 tells us that the ATAP reduced the respondents' emotional problems and thus enhanced their SWB. Second, the CLDS database provided information about respondents' emotional stability through nine questions, such as inattentiveness or feeling depressed, sad, and lonely. We summed these answers to construct another mechanism variable,[16] *Emotional stability*, with larger values indicating greater emotional instability. According to the findings in column (2) of Panel A and Panel B of Table 7, the ATAP improved respondents' emotional stability and SWB while stifling their negative feelings. Third, the CLDS questionnaire also included seven questions about the respondents' psychological problems, such as "feeling uncertain about the future" and "feeling like a failure in life." Similarly, we constructed the psychological problem variable by adding these dimensions, ranging from integers 0 to 7. The ATAP estimate is also significantly negative, as shown in column (3) of Panel A and Panel B of

[15] Since the normal value of BMI for individuals is between 20 and 25, we excluded samples with BMI less than 20 (too thin to be healthy); thus, the larger the value of BMI, the fatter (unhealthier) respondents are.

[16] For example, if the respondent does not match any of items for emotional stability, set to 0. If one of them matches, set to 1, and if all of them match, set to 9.

Table 8 The mechanism of exercise and outdoor activities

	(1)	(2)	(3)	(4)
Panel A	Exercise	Exercise frequency	Exercise time	Outdoor activities
ATAP	0.3465***	0.3954***	0.2896***	0.1982***
	(0.0198)	(0.0206)	(0.0297)	(0.0095)
Panel B	SWB			
Exercise	4.1974***			
	(0.0302)			
Exercise frequency		3.6037***		
		(0.0283)		
Exercise time			3.8621***	
			(0.0397)	
Outdoor activities				3.7920***
				(0.0427)
Control variables	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	24,489	5,321	5,321	30,918

The explained variable is *SWB* for all columns in Panel B. Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering

Table 7, demonstrating that improving air quality preserved mental health and increased SWB.

Moreover, prior research indicated that getting enough sleep enhances people's SWB (Otsuka et al. 2020), yet extreme air pollution may result in insomnia (Chen et al. 2019; Heyes & Zhu 2019; Liu et al. 2020). We further analyzed the mechanism of insomnia. As expected, the results in column (4) of Panel A in Table 7 show that the air quality improvement from the ATAP can lead to a 16.93% reduction in the probability of insomnia. Furthermore, results in column (4) of Panel B told us that the insomnia is indeed detrimental to individual SWB enhancement. Thus, our results support the idea that the ATAP improved SWB by reducing insomnia.

Exercise and outdoor activities

Exercise and outdoor activities are beneficial to increase SWB (Hallmann et al. 2013; Korpela et al. 2017), but air pollution leads to a decrease in outdoor trips (Deschenes et al. 2020; Mittleman 2007; Neidell 2010). Therefore, in this subsection, we explored whether the air quality improvement effect generated by the ATAP could enhance individual happiness through more exercise and outdoor activities.

According to three questions in the CLDS, "Do you exercise regularly?", "What is your average frequency of exercise per week?", and "What is the approximate average time you spend exercising each time?", we constructed three mechanism variables. The first is a dummy variable, with 1 indicating regular exercise and 0 otherwise; the last two are continuous variables, with larger values meaning more frequent or longer exercise. As shown in the first three columns of Panel A of Table 8, the improving air pollution not only improved residents' exercise habits but also enhanced their exercise frequency and duration. Also, in Panel B, results show that there is a positive correlation between exercising and people's SWB enhancement, which is consistent with the findings in the literature (Korpela et al. 2017). Taken together, these findings provide supportive evidence that the ATAP implementations could make people happier by promoting exercise.

In addition, the CLDS recorded respondents' outdoor activities in nine areas, such as volunteer activities, club activities, social work activities, and religious activities. We added up the number of all activities to reflect the frequency of respondents' participation in outdoor activities. The higher the value, the higher the frequency of participating in outdoor activities. It is evident from the findings in column (4) of Panels A and B of Table 8 that ATAP implementations could promote the frequency of respondents' participation in outdoor activities and outdoor activities also really help to improve the happiness of residents. Thus, we provided evidence that ATAP implementations enhanced individual SWB by promoting exercise and outdoor activities.

Conclusions and discussion

Many academics focused on detrimental impacts of air pollution on people's SWB in China (e.g., Dong et al. 2022; Li et al. 2014; Song et al. 2019; Zhang et al. 2017a, b; Zheng et al. 2019), but the benefits of environmental regulation and control measures on people' SWB have received little

consideration (Gao et al. 2021). The ATAP is as the strongest pollution control policy by the Chinese government in recent years (Greenstone et al. 2021), making it important to explore its micro welfare effects. By considering the ATAP as a proxy for air quality improvement and as a quasi-natural experiment, we used the DID approach to confirm the subjective welfare-enhancing effects of air quality enhancement.

We drew several conclusions from our findings. First, we found that the air quality improvement due to the ATAP significantly improved residents' SWB, which was confirmed by a number of robustness checks. In addition to placebo tests, using alternative specifications, and excluding policy interference, we also excluded the impact of potential spillover effects and self-selection effects to further ensure the reliability of the results. According to the baseline findings, ATAP implementations raised residents' happiness scores by 0.5771 points. Second, we explored several potential mechanisms. Results show that improving physical and mental health status, alleviating emotional problems, reducing insomnia, and promoting exercise and outdoor activities can explain the positive impact from the ATAP. Third, the results of heterogeneity analysis showed that this SWB enhancement effect created by ATAP are more beneficial to some socially disadvantaged groups, such as the middle-aged and elderly and low-income people. Also, for residents living in urban areas with more severe air pollution, they gained a greater marginal effect from the ATAP implementation from the perspective of SWB enhancement.

At present, the effectiveness of environmental regulation and pollution control actions in China have been widely recognized, and several scholars elaborated on the benefits of these policies at the individual micro-level (Huang et al. 2018; Song et al. 2019; Tanaka 2015; Wang et al. 2021), yet they paid little attention to the emerging field of individual subjective welfare. Therefore, our study further complements the research field on the role of air pollution control policies in China, as well as welfare growth effects of environmental quality improvements. In addition, we found that the ATAP had a stronger effect on enhancing SWB for some socially vulnerable groups, indicating that these environment-related public policies helped promote social equity and sustainable development in China. More importantly, our findings provide insights not only for further improving China's air pollution control policy system, but they also offer some lessons for other developing countries facing air pollution.

In terms of practical implications, China should further strengthen air pollution control. Although air quality has moderated in some highly polluted areas in China (e.g., BTH), winter air pollution events still occurred in some cities after the ATAP. Some possible follow-up pollution control measures include further promotion of clean energy and new energy vehicles, reduction of winter coal supply, and enhanced transparency of pollution information. Additionally, our study implies that pollution control measures also need to consider the public's subjective perceptions and willingness, which helps to motivate residents to improve air quality. Residents are not only the victims of air pollution but are also the participants and monitors of pollution control. Taken together, this research suggests that air quality improvements and air pollution control policies could be an important and effective strategy to improve people's SWB and could have large benefits in terms of public welfare and sustainable economic and social development.

There are some limitations of this paper and recommendations for further research. First, limited by the CLDS questionnaire, we could not determine the exact date of the interview because respondents' recent experiences may influence their happiness (Menz 2011; Zhang et al. 2017b). Follow-up studies may design better SWB scales to continuously observe people's responses to air quality improvements. Second, our main specification captures the impact of city-level air quality improvement effects from the ATAP on micro-individuals, yet individuals living in different areas of the same city may also face different levels of air pollution. In particular, Chinese cities focus on functional zoning (Liu et al. 2018), such as industrial, commercial, and agricultural zones. Therefore, more specific and detailed geographic information in follow-up surveys would better confirm the SWB effects of ATAP implementations. Third, some mechanisms in this paper are imperfect. For example, in physical health, our main mechanism variable is self-rated health, while information on respiratory or lung disease closely linked to air pollution is lacking (Seaton et al. 1995). Thus, future research could complement mechanisms of ATAP implementations on individual SWB. Last, we aimed to assess the micro-welfare effects of the ATAP, a major air pollution control policy in recent years in China, yet ignored the potential costs. Therefore, the cost–benefit of ATAP implementations should be given attention in subsequent studies.

Appendix

Table 9 Variable definitions

Variable	Definition
SWB	1 = extremely unhappy; 2 = somewhat unhappy; 3 = indifferent; 4 = somewhat happy; 5 = extremely happy
ATAP	= 1 if a city implemented ATAP from a certain year; 0 otherwise
Gender	Female = 1; male = 0
Ethnic	Han = 1; minority = 0
Age	Age of the respondent
Age2	Age squared divided by 100
Hukou	Rural = 1; urban = 0
Education	1 = unschooled; 2 = primary school; 3 = middle school; 4 = general or vocational high school; 5 = post-secondary or bachelor's degree; 6 = master's or doctoral degree
Marriage	= 1 if the respondent is married and has a spouse; 0 otherwise
Residence	Rural = 1; urban = 0
Income	Total income of the respondent (yuan, logarithmic)
$LPM_{2.5}$	Average $PM_{2.5}$ concentrations with a one-year lag at city level (ug/m^3)

Table 10 Marginal effects of SWB in column (4) of Table 2

	Extremely unhappy	Somewhat unhappy	Indifferent	Somewhat happy
	(1)	(2)	(3)	(4)
ATAP	-0.0492***	-0.0977***	-0.2514***	0.0667***
	(0.0012)	(0.0028)	(0.0071)	(0.0029)
Control variables	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	30,918	30,918	30,918	30,918

When *SWB* equal to 1, 2, 3, or 4, marginal effects are reported in each of the columns using the ordered logit model. Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering

Table 11 Robustness checks: alternative specifications

	Replacing the explained variable	Weather and survey month FEs	Individual FE	ATAP cities in the fourth quarter are considered the treatment group for the following year	Using total treatment months/12 to measure treatment effect	Controlling for more macro control variables	IV estimates	
							First-stage	Second-stage
	(1)	(2)	(3)	(1)	(2)	(3)	(7)	(8)
ATAP	0.4495***	0.2864***	0.6049*	0.4283***	0.3876***	0.3549***	-0.1352***	
	(0.0423)	(0.0094)	(0.3530)	(0.0628)	(0.0689)	(0.0076)	(0.0094)	
$PM_{2.5}$								-3.5348***
								(0.0648)
Control variables	Yes	Yes	Yes	Yes			Yes	Yes
Individual FE	No	No	Yes	No	No	No	No	No
City FE	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8821	30,909	4362	30,918	30,918	28,312	30,918	30,918
Kleibergen-Paap rk LM							0.0000	
Kleibergen-Paap Wald rk F							924.7842	
Anderson-Rubin Wald test							0.0372	

Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering. In the last two columns, $LPM_{2.5}$ is not included in the control variable set. Kleibergen-Paap rk LM reports the results of the underidentification test; Kleibergen-Paap Wald rk F reports the results of the weak identification test; Anderson-Rubin Wald test reports the results of the weak-instrument-robust inference

Table 12 Robustness checks: placebo tests

	(1)	(2)
ATAP	0.9476	0.0378
	(1.6542)	(0.3945)
Control variables	Yes	Yes
City FE	Yes	Yes
Year FE	Yes	Yes
Observations	30,918	30,918

Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering

Table 13 Robustness checks: considering self-selection effects and expectation effects

	Total population of cities	Deleting floating population	PSM-DID	$2012PM_{2.5} \times$ Year	Controlling for expectation effects
	(1)	(2)	(3)	(4)	(5)
ATAP	-0.1289	0.5844***	0.5913***	0.6581***	0.6472***
	(0.5885)	(0.1752)	(0.0744)	(0.03318)	(0.0479)
$2012PM_{2.5} \times$ Year	No	No	No	Yes	No
Control variables	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	1,442	27,275	30,895	30,918	19,213

Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering

Table 14 Robustness checks: excluding other policy interference

	Carbon market pilots	Low-carbon cities	National forest cities	Innovation-oriented cities	Smart cities	Cities implemented before September 2013
	(1)	(2)	(3)	(4)	(5)	(6)
ATAP	0.5942***	0.4932***	0.7541***	0.6658***	0.9152***	0.6395***
	(0.0568)	(0.0169)	(0.1226)	(0.0955)	(0.0935)	(0.0749)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	23,809	5,121	18,530	16,618	16,015	27,641

Significant levels of 1%, 5%, and 10% are indicated by the symbols ***, **, and *, respectively. Parentheses show standard errors for city-level clustering

Funding This work was supported by the National Natural Science Foundation of China (No. 72304102), the Ministry of education of Humanities and Social Science project (No. 23YJCZH110), and the Philosophy and Social Science Program Discipline Co-construction Project of Guangdong Province (No. GD23XYJ68).

Data availability The authors do not have permission to share data.

Declarations

Ethical approval Not suitable.

Competing interests We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Consent to participate Not suitable.

Consent to Publish Agree.

References

Acemoglu D, Moscona J, Robinson JA (2016) State capacity and American technology: Evidence from the nineteenth century. Am Econ Rev 106(5):61–67

Callaway B, Sant'anna PH (2021) Difference-in-differences with multiple time periods. J Econ 225(2):200–230

Chang PJ, Song R, Lin Y (2019) Air pollution as a moderator in the association between leisure activities and well-being in urban China. J Happiness Stud 20(8):2401–2430

Chen J (2023) Mitigating nitrogen dioxide air pollution: The roles and effect of national smart city pilots in China. Energy 263:125652

Chen G, Xiang H, Mao Z, Huo W, Guo Y, Wang C, Li S (2019) Is long-term exposure to air pollution associated with poor sleep quality in rural China? Environ Int 133:105205

Chen Y, Fan Z, Gu X, Zhou LA (2020) Arrival of young talent: The send-down movement and rural education in China. Am Econ Rev 110(11):3393–3430

Chen S, Oliva P, Zhang P (2022) The effect of air pollution on migration: Evidence from China. J Dev Econ 156:102833
Clancy L, Goodman P, Sinclair H, Dockery DW (2002) Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. The Lancet 360(9341):1210–1214
Cohn MA, Fredrickson BL, Brown SL, Mikels JA, Conway AM (2009) Happiness unpacked: Positive emotions increase life satisfaction by building resilience. Emotion 9(3):361–368
De Chaisemartin C, d'Haultfoeuille X (2020) Two-way fixed effects estimators with heterogeneous treatment effects. Am Econ Rev 110(9):2964–2996
Delfgaauw J, Dur R (2010) Managerial talent, motivation, and self-selection into public management. J Public Econ 94(9–10):654–660
Deschenes O, Wang H, Wang S, Zhang P (2020) The effect of air pollution on body weight and obesity: Evidence from China. J Dev Econ 145:102461
Dong X, Zhang M, Wang W (2022) Research on the influence of haze pollution on Chinese residents' happiness based on Baidu Index data. Air Qual Atmos Health 15(11):1981–1992
Duflo E (2001) Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. Am Econ Rev 91(4):795–813
Dumludag D, Gokdemir O (2022) Income aspiration, income comparison and life satisfaction: The case of Turkish migrants in the Netherlands. J Happiness Stud 23(4):1359–1378
Feng Y, Ning M, Lei Y, Sun Y, Liu W, Wang J (2019) Defending blue sky in China: Effectiveness of the "Air Pollution Prevention and Control Action Plan" on air quality improvements from 2013 to 2017. J Environ Manage 252:109603
Fu J, Xiao G, Wu C (2022) Regional performance of air pollution prevention and control in China: A gap between empirical evidence and perception. Air Qual Atmos Health 15(10):1789–1805
Fu J, Fu H, Zhu C, Sun Y, Cao H, Deng X (2024) The impact of air pollution on residents' health in China: The mediating effect of population migration. Air Qual Atmos Health. https://doi.org/10.1007/s11869-024-01506-5
Gao J, Yuan Z, Liu X, Xia X, Huang X, Dong Z (2016) Improving air pollution control policy in China—A perspective based on cost–benefit analysis. Sci Total Environ 543:307–314
Greenstone M, He G, Li S, Zou EY (2021) China's war on pollution: Evidence from the first 5 years. Rev Environ Econ Policy 15(2):281–299
Grimard F, Laszlo S, Lim W (2010) Health, aging and childhood socioeconomic conditions in Mexico. J Health Econ 29(5):630–640
Gruber J, Levine P, Staiger D (1999) Abortion legalization and child living circumstances: Who is the "marginal child"? Quart J Econ 114(1):263–291
Gunby N, Coupé T (2023) Weather-related home damage and subjective well-being. Environ Resource Econ 84(2):409–438
Guo S, Wang W, Zhang M (2020) Exploring the impact of environmental regulations on happiness: New evidence from China. Environ Sci Pollut Res 27(16):19484–19501
Hallmann K, Breuer C, Kühnreich B (2013) Happiness, pride and elite sporting success: What population segments gain most from national athletic achievements? Sport Manag Rev 16(2):226–235
He G, Fan M, Zhou M (2016) The effect of air pollution on mortality in China: Evidence from the 2008 Beijing Olympic Games. J Environ Econ Manag 79:18–39
Heyes A, Zhu M (2019) Air pollution as a cause of sleeplessness: Social media evidence from a panel of Chinese cities. J Environ Econ Manag 98:102247
Huang J, Pan X, Guo X, Li G (2018) Health impact of China's Air Pollution Prevention and Control Action Plan: An analysis of national air quality monitoring and mortality data. Lancet Planetary Health 2(7):e313–e323
Huber M, Steinmayr A (2021) A framework for separating individual-level treatment effects from spillover effects. J Bus Econ Stat 39(2):422–436
Korpela K, De Bloom J, Sianoja M, Pasanen T, Kinnunen U (2017) Nature at home and at work: Naturally good? Links between window views, indoor plants, outdoor activities and employee well-being over one year. Landsc Urban Plan 160:38–47
Levinson A (2012) Valuing public goods using happiness data: The case of air quality. J Public Econ 96(9–10):869–880
Li X, Wang D (2022) Does transfer payments promote low-carbon development of resource-exhausted cities in China? Earth's Future 10(1):e2021EF002339
Li X, Zhao C (2023) Can national forest city construction mitigate air pollution in China? Evidence from a quasi-natural experiment. Environ Geochem Health 45(6):3003–3024
Li X, Zhao C (2024) Did innovative city constructions reduce carbon emissions? A quasi-natural experiment in China. Environ Dev Sustain 26(3):6315–6340
Li Z, Folmer H, Xue J (2014) To what extent does air pollution affect happiness? The case of the Jinchuan mining area, China. Ecol Econ 99:88–99
Li S, Chen Y, He G (2018) Laugh and grow fat: Happiness affects body mass index among urban Chinese adults. Soc Sci Med 208:55–63
Li JJ, Massa M, Zhang H, Zhang J (2021a) Air pollution, behavioral bias, and the disposition effect in China. J Financ Econ 142(2):641–673
Li X, Hu Z, Cao J (2021b) The impact of carbon market pilots on air pollution: Evidence from China. Environ Sci Pollut Res 28(44):62274–62291
Li X, Hu Z, Cao J, Xu X (2022) The impact of environmental accountability on air pollution: A public attention perspective. Energy Policy 161:112733
Li X, Zhao C, Cao J (2023) Official Turnover and Sustainable Development in China. J Chin Polit Sci. https://doi.org/10.1007/s11366-023-09864-0
Liao L, Zhao C, Li X, Qin J (2021) Towards low carbon development: The role of forest city constructions in China. Ecol Ind 131:108199
Lim YH, Kim H, Kim JH, Bae S, Park HY, Hong YC (2012) Air pollution and symptoms of depression in elderly adults. Environ Health Perspect 120(7):1023–1028
Lipp J (2007) Lessons for effective renewable electricity policy from Denmark, Germany and the United Kingdom. Energy Policy 35(11):5481–5495
Liu Y, Fu B, Zhao W, Wang S, Deng Y (2018) A solution to the conflicts of multiple planning boundaries: Landscape functional zoning in a resource-based city in China. Habitat Int 77:43–55
Liu J, Wu T, Liu Q, Wu S, Chen JC (2020) Air pollution exposure and adverse sleep health across the life course: A systematic review. Environ Pollut 262:114263
Liu Y, Liu Y, Lin Y (2021) Upward or downward comparison? Migrants' socioeconomic status and subjective wellbeing in Chinese cities. Urban Studies 58(12):2490–2513
Liu Y, Yang S, Li C, Fang Q, Han C, Cheng C, Li X (2024) Changes in years of life lost and life expectancy by reducing air pollution. Air Qual Atmos Health 17(1):89–102
Lu Y, Wang J, Zhu L (2019) Place-based policies, creation, and agglomeration economies: Evidence from China's economic zone program. Am Econ J Econ Pol 11(3):325–360
Lu H, Tong P, Zhu R (2020) Longitudinal evidence on social trust and happiness in China: Causal effects and mechanisms. J Happiness Stud 21(5):1841–1858
Luo Z, Li H (2018) The impact of "Atmosphere Ten Articles" policy on air quality in China. China Ind Econ 32(9):136–154 ((**in Chinese**))

Maji KJ, Li VO, Lam JC (2020) Effects of China's current Air Pollution Prevention and Control Action Plan on air pollution patterns, health risks and mortalities in Beijing 2014–2018. Chemosphere 260:127572

Menz T (2011) Do people habituate to air pollution? Evidence from international life satisfaction data. Ecol Econ 71:211–219

Miguel E, Kremer M (2004) Worms: Identifying impacts on education and health in the presence of treatment externalities. Econometrica 72(1):159–217

Mittleman MA (2007) Air pollution, exercise, and cardiovascular risk. N Engl J Med 357(11):1147–1149

Neidell M (2010) Air quality warnings and outdoor activities: Evidence from Southern California using a regression discontinuity design. J Epidemiol Community Health 64(10):921–926

Otsuka Y, Kaneita Y, Itani O, Jike M, Osaki Y, Higuchi S, Yoshimoto H (2020) The relationship between subjective happiness and sleep problems in Japanese adolescents. Sleep Medicine 69:120–126

Pan H (2023) Effect of the ecological environment on the residents' happiness: The mechanism and an evidence from Zhejiang Province of China. Environ Dev Sustain 25(3):2716–2734

Qian N (2008) Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance. Quart J Econ 123(3):1251–1285

Rasmussen M, Laumann K (2014) The role of exercise during adolescence on adult happiness and mood. Leis Stud 33(4):341–356

Sanchez A, Vazquez C (2014) Looking at the eyes of happiness: Positive emotions mediate the influence of life satisfaction on attention to happy faces. J Positive Psychol 9(5):435–448

Sanduijav C, Ferreira S, Filipski M, Hashida Y (2021) Air pollution and happiness: Evidence from the coldest capital in the world. Ecol Econ 187:107085

Seaton A, Godden D, MacNee W, Donaldson K (1995) Particulate air pollution and acute health effects. The Lancet 345(8943):176–178

Song Y, Li M, Zhang M, Sun X (2019) Study on the impact of air pollution control on urban residents' happiness from microscopic perspective. J Clean Prod 229:1307–1318

Stutzer A, Meier AN (2016) Limited self-control, obesity, and the loss of happiness. Health Econ 25(11):1409–1424

Tanaka S (2015) Environmental regulations on air pollution in China and their impact on infant mortality. J Health Econ 42:90–103

Wang S, Palazzo E (2021) Sponge City and social equity: Impact assessment of urban stormwater management in Baicheng City, China. Urban Climate 37:100829

Wang J, Wang Y, Sun C, Chen X (2021) Does mandatory air quality information disclosure raise happiness? Evidence from China. Energy Economics 94:105094

Wang L, Wang Z, Ma Y (2022) Does environmental regulation promote the high-quality development of manufacturing? A quasi-natural experiment based on China's carbon emission trading pilot scheme. Socioecon Plann Sci 81:101216

Welsch H (2006) Environment and happiness: Valuation of air pollution using life satisfaction data. Ecol Econ 58(4):801–813

Welsch H, Kühling J (2009) Using happiness data for environmental valuation: Issues and applications. J Econ Surv 23(2):385–406

Yu Y, Dai C, Wei Y, Ren H, Zhou J (2022) Air pollution prevention and control action plan substantially reduced PM2.5 concentration in China. Energy Economics 113:106206

Zhang X, Zhang X, Chen X (2017a) Valuing air quality using happiness data: The case of China. Ecol Econ 137:29–36

Zhang X, Zhang X, Chen X (2017b) Happiness in the air: How does a dirty sky affect mental health and subjective well-being? J Environ Econ Manag 85:81–94

Zhao C, Guo J (2022) Are veterans happy? Long-term military service and the life satisfaction of elderly individuals in China. J Happiness Stud 23(2):477–508

Zhao C, Deng M, Cao X (2021) Does haze pollution damage urban innovation? Empirical evidence from China. Environ Sci Pollut Res 28(13):16334–16349

Zhao C, Chen B, Li X (2023) Rising housing prices and marriage delays in China: Evidence from the urban land transaction policy. Cities 135:104214

Zheng S, Wang J, Sun C, Zhang X, Kahn ME (2019) Air pollution lowers Chinese urbanites' expressed happiness on social media. Nat Hum Behav 3(3):237–243

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Electronic Commerce Research (2024) 24:69–104
https://doi.org/10.1007/s10660-022-09549-5

The effect of digital finance on Residents' happiness: the case of mobile payments in China

Chunkai Zhao[1] · **Xing Li**[2] · **Jianfeng Yan**[3]

Accepted: 5 March 2022 / Published online: 15 June 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

With the popularization of digital finance in China, mobile payments have penetrated into all aspects of residents' daily life. However, few studies have examined the potential impact of mobile payments on people's happiness in China. Using the nationally representative data from the China Household Finance Survey (CHFS), this study adopts the ordered probit regression with endogenous treatment to adjust for possible endogeneity to assess the effect of mobile payments on residents' happiness. The results suggest an association between mobile payment usage and increases in happiness, which is supported by several robustness checks, such as using an alternative instrumental variable (IV), replacing the explained variable, and removing some extreme observations. In addition, we explore the mechanisms by which mobile payments affect residents' happiness from multiple perspectives. Positive mechanisms include promoting quality of life, reducing transaction costs, stimulating entrepreneurship, and increasing social interaction. However, as a non-cash payment method, mobile payments may also lead to over-consumption, which is detrimental to residents' happiness. Furthermore, the heterogeneous analysis shows inclusive attributes of mobile payments. We find mobile payments have a greater positive effect on happiness of some socially disadvantaged groups, such as elderly individuals, rural residents, the low-educated, and low-income households. These findings supplement the literature on online happiness and financial inclusion and refer to the possible negative impact of mobile payments. Therefore, it is necessary to actively promote mobile payments to benefit more socially vulnerable groups and prevent potential risks from over-consumption.

Keywords Digital finance · Mobile payments · Happiness · Ordered probit regression with endogenous treatment · China

JEL Classification I31 · G51 · D12

✉ Xing Li
Xingli277@126.com

Extended author information available on the last page of the article

1 Introduction

In recent years, with the increasingly close integration of information technology and financial services, China is moving into the digital financial era. China's digital financial system, consisting of mobile payments, online banking, financial service outsourcing, and online loans, has not only changed the financial formats but also people's daily life [38, 41]. Among them, mobile payment, as the payment tool closely connected with people's daily economic activities, has achieved great results in China [29, 73]. Mobile payment refers to the use of electronic devices such as smartphones by mobile clients to pay by electronic money [48, 58]. Mobile payment effectively unites the Internet, terminal devices and financial institutions to form a new type of payment system [11, 80]. Nowadays, China is the world's most developed and widespread country for mobile payments. A report showed that in 2017, more than 47% of consumers use mobile payments in China, ranking first in the world, and Alipay and WeChat Pay[1] have 810 million users.[2] The more convenient life brought by mobile payment is gradually turning China into a cashless society.

The success of mobile payments has inevitably attracted increasing discussion for related topics in the literature. However, most studies focused on the acceptances or drivers of mobile payments [1, 11, 39, 42, 48, 50, 55, 80] or impacts on household economic behaviors [29, 30, 35, 41, 73, 79]. Only a few analyses paid attention to household welfare-enhancing effects of mobile payments, such as household risk-sharing improvements and financial exclusion reductions [29, 72]. To date, to our best knowledge, few studies have explored how mobile payments affect residents' happiness. Given the inclusive financial attributes of digital finance [6, 21, 29, 56], as well as the convenience that mobile payments provide for family everyday life [41, 80], we supposed that Chinese residents, especially some socially vulnerable groups, are benefited by the popularity and development of mobile payments.

Specifically, as a portable and fast payment method, the convenience of mobile payments makes people no longer rely on traditional financial transaction tools in economic activities, which saves unnecessary transaction costs and improves transaction efficiency [1, 11]. In addition, mobile payments may improve household consumption structure and enhance people's quality of life [41, 51, 83]. Furthermore, mobile payment tools, such as WeChat Pay and Alipay, can also be used as online lending and online social platforms, lowering thresholds of access to funds, accelerating information transmission, and providing better conditions for starting businesses and increasing social interaction [29, 73]. Considering the strong link between factors such as lower transaction costs and improved quality of life and residents' happiness, mobile payments are likely to enhance people's welfare and happiness through multiple channels.

Based on a nationally representative sample from the China Household Finance Survey (CHFS) in 2017, we tried to identify the causal effect between mobile payment usage and the happiness. However, previous literature suggested that using

[1] Alipay and WeChat Pay are the most dominant mobile payment tools in China.

[2] Data resource: https://www.statista.com/statistics/744944/mobile-payment-platforms-users/.

mobile payment is not random but selective for households, and the resulting endogeneity problems may lead to biased estimates [73, 79]. Therefore, we employed the ordered probit regression with endogenous treatment to address potential endogeneity issues [57, 82, 86]. In addition, this study has important practical implications as the inclusiveness of mobile payments benefits more socially vulnerable groups and make them happier. In recent years, the United Nations has been active in making people realize the importance of happiness in their lives, and advocated reducing inequality and improving mental health and well-being of all mankind [86]. The results of this research offer some insights for policymakers to further promote digital finance and mobile payments, enhance financial accessibility, and improve happiness of residents.

This paper contributes to the existing literature in three aspects. First, it is the first to employ nationally representative sampling survey data to evaluate the impact of mobile payments on residents' happiness. More importantly, using the ordered probit regression with endogenous treatment helps us to identify the causal relationship between mobile payment usage and people's happiness. Second, we analyzed and verified in detail the mechanisms by which mobile payments promote people's happiness in several ways. On the positive side, mobile payments can reduce transaction costs, improve quality of life, promote entrepreneurship, and enhance social interactions; on the negative side, given the close relationship between mobile payments and consumption [41, 51], mobile payments may lead to a decrease in happiness due to excessive household consumption. The findings supplement the literature on the potential adverse effects of mobile payments [34, 42]. Third, our heterogeneity focused on the inclusive financial attributes of digital finance and mobile payments in terms of age, urban and rural areas, education level, and household income. We find that mobile payment usage is more helpful to improve the happiness of some socially disadvantaged groups, such as elderly individuals, rural residents, the low-educated, and low-income households, which provides evidence for inclusive finance to promote residents' welfare from the perspective of happiness.

The rest of this paper is organized as follows. Section 2 provides a brief literature review and describes the institutional background. Section 3 outlines theoretical mechanisms of mobile payments on happiness. Section 4 explains our estimation strategy. Section 5 introduces the data and variables. Section 6 presents empirical results. Section 7 concludes this paper.

2 Literature review and institutional background

2.1 Literature review

There exist three strands of literature related to our study. First, some scholars explored the function of mobile payments and its impact on household economic behaviors. With the mobile Internet, mobile payment can fill the shortcomings of traditional financial payment methods and reduce the high transaction costs caused by backward infrastructure, which plays an important role in enhancing the welfare of households in developing countries [2, 29, 56, 87]. In particular,

[87] found that rural residents and women were rewarded with greater happiness through mobile payments. In addition, mobile payments eliminate the need for cash, bank cards, or checks in daily economic activities, which helps to save transaction time, reduce transaction friction, and improve convenience and security [30, 72, 83]. Moreover, the role of mobile payments in providing small loans and broadening social networks makes it effective in stimulating entrepreneurship [73, 79]. Also, more convenient financial services provided by mobile payments may smooth consumption risks and change savings behavior [29, 35, 41].

Second, one type of literature focused on people's online happiness, but reached two opposing views. Early research supported a hypothesis of social displacement, that people's happiness was weakened by the Internet use, as it reduces the offline communication and narrowed the social circle [28, 33, 53]. Subsequent studies provided more support for the idea that the Internet enhances people's happiness. On the one hand, the Internet increases the efficiency of information and knowledge dissemination and using Internet tools helps improve people's quality of life [9, 22, 81]. On the other hand, as a kind of social capital, it provides more opportunities and resources [64, 68], which can create more employment opportunities, enhance social status and psychological empowerment, and help create greater happiness [9, 88].

In recent years, with the popularity of online shopping and mobile Internet, more and more research discussed whether they would enhance people's happiness. Not surprisingly, these studies highlighted the positive role of online shopping and mobile Internet in enhancing residents' happiness [17, 34, 45, 57, 81, 86]. Online shopping improves shopping efficiency, saves time and enhances users' shopping experience with the advantages of the Internet platform [17, 34, 86]. Nevertheless, the potential downside is that online shopping may induce impulsive buying and conspicuous consumption [4, 24], or even damage consumer well-being by over-consumption [66].

Similarly, mobile Internet enhances the convenience of Internet use for residents, strengthening online interaction and expanding social networks [45, 47, 57, 76]. However, [81] stressed the potential negative impact of mobile Internet from the perspective of privacy risk and task risk and argued that there is an inverted U-shaped relationship between mobile Internet and users' happiness. Furthermore, [34] found that although online shopping increased happiness, mobile Internet use and mobile shopping instead decreased users' happiness, due to potential privacy concerns. The above literature is summarized in Table 1.

Third, another type of literature relevant to our research is to analyze the effects of financial inclusion and financial accessibility on people's happiness. In terms of the economic, inclusive finance provides low-cost financial services that can effectively alleviate conflicts between household income and consumption, increase the risk-resistance capacity [21, 36], and maximize the household's utilities to enhance their happiness. At the psychological level, financial inclusion can alleviate poverty, financial exclusion, and income inequality [56, 61], promote social equity and improve residents' confidence in their future lives [6, 13], and ultimately enhance people's happiness.

Table 1 Prior studies on online happiness

Topics	Authors	Findings
Internet use and happiness	Huang [28], Kraut et al. [33], Morgan & Cotten [53]	Happiness was weakened by the Internet use
	Carter et al. [9]; Grieve et al. [22], Pénard & Poussing [64], Shapira et al. [68], Zhu et al. [88]	Internet use enhanced people's happiness
Mobile Internet and happiness	Lu & Kandilov [45], Nie et al. [57], Yang et al. [76]	Mobile Internet use was positively associated with happiness
	Zhan & Zhou [81]	An inverted U-shaped relationship between mobile Internet use and happiness
	Kwak et al. [34]	Mobile Internet use damaged the happiness, due to privacy concerns
Online shopping and happiness	Duan & Dholakia [17], Kwak et al. [34], Zheng & Ma [86]	Online shopping helped promote consumers' happiness
	Boonchoo & Thoumrungroje [4], He et al. [24], Rose & Dhandayudham [66]	Potential negative effects of online shopping on happiness
Mobile payment and happiness	Zheng & Ma [87]	Mobile payments promoted the happiness, especially rural residents and women

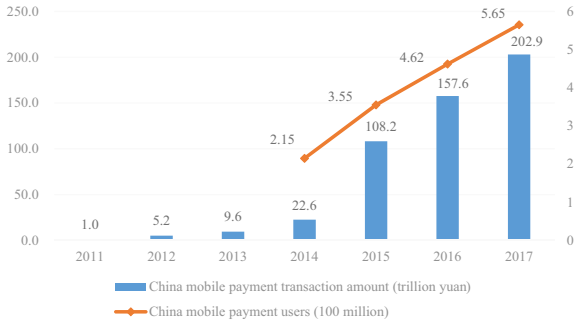


Fig. 1 Transaction and users scale of China Mobile Payment. *Data source*: The People's Bank of China, the Payment & Clearing Association of China and iiMedia Research

2.2 Institutional background

As early as 1999, China Mobile, one of the three largest telecom operators in China, in cooperation with some financial institutions, started mobile payment pilots in some cities in Guangdong. However, due to the lagging financial technology and mobile Internet technology at that time, mobile payments were not promoted nationwide. Thanks to the popularity of smartphones, mobile Internet, and mobile e-commerce, things took a turn for the better in 2011 when users in some major cities began to use mobile payment tools. More importantly, the People's Bank of China issued third-party payment licenses to Alipay, Tenpay, UnionPay, UnionPay Commerce, and 99Bill in June 2011, which provided some mobile payment platforms with legal status and laid a solid foundation for the promotion of mobile payments in China.

In addition to the efforts of the Chinese government, telecom operators, financial institutions and third-party payment platforms, the use of QR codes has further accelerated the development of China's mobile payments since 2014. With QR codes, mobile payment processes become more convenient and secure, being used and promoted on a large scale in China. Benefit from this advancement, non-bank mobile payment tools such as Alipay and WeChat Pay are gradually becoming the mainstay of China's mobile payment market. Furthermore, mobile payment platforms were actively engaged in other financial services. For example, in 2013, Alipay accessed the Yu'E Bao money fund featuring easy operation, low threshold, zero fees, and anytime access [29, 40]. In addition to earning financial income, Yu'E Bao can also be used directly for shopping, money transfers, bill payments, and repayments, quickly becoming the largest money fund in the world [40]. What's more, WeChat Pay, based on China's largest social networking software, launched the WeChat red packet in 2014, which fits in with Chinese cultural traditions and is equally popular with the majority of users.

Figure 1 shows the total number of mobile payment transactions from 2011 to 2017 and the number of mobile payment users from 2014 to 2017 in China. In 2013, the transaction value of mobile payments in China was only 9.6 trillion yuan, but it grew 135.42% in 2014 and saw an explosive growth in 2015, reaching 108.2 trillion yuan. In 2017, statistics from the People's Bank of China show that the number of mobile payment users in China has reached 565 million and the total transaction

value is more than 200 trillion is 90 times of the United States in the same period[3]. The number of mobile payment users exceeded 565 million, which means that more than one-third of Chinese people are using mobile payments.

With the further development of digital finance and mobile payment technology in China, fingerprint payment, unmanned supermarket, face scanning payment and iris payment are also gradually applied to mobile payment, which also makes mobile payment in China more convenient, efficient, and secure. Nowadays, the application scenarios of mobile payment are more diversified, including online and offline shopping, social networking, financial management, borrowing and lending, transportation, travel, entertainment, catering and takeaway, and so forth, which have penetrated into every aspect of life in China. The new payment revolution created by mobile payments is helping to improve the consumer experience, reduce transaction costs, and create greater digital financial benefits.

3 Theoretical analysis

We classified the possible mechanisms by which mobile payment usage affects residents' happiness into the following four categories: improving the quality of life, reducing transaction costs, stimulating entrepreneurship, and increasing social interaction.

3.1 Improving the quality of life

Consumption is an important way to enhance personal utility and welfare, and consumption upgrade and promotion of household consumption structure are important ways to improve people's quality of life. However, China had the problems of high saving and low consumption for a long time [52, 74]. From the perspective of preventive savings, Chinese residents are used to reducing their consumption to cope with future uncertain events and risks [10, 75], which leads to a decline in quality of life. Besides, factors such as traditional culture, consumption habits, and insufficient supply of financial consumer tools attract the attention of scholars [14, 26]. As a multi-functional comprehensive financial instrument, the convenience and practicality of mobile payment may change consumption habits [42], improve the consumer experience [41], and promote household consumption upgrade [83].

First of all, mobile payment as a non-cash payment method can effectively reduce people's psychological account losses during consumption. Skipping the psychological process of cash payment helps to speed up consumption decision-making and stimulate the increase of household consumption [70, 83]. Next, mobile payment makes it easier for residents to get the required goods and services through their smartphones, which may change their consumption habits. Mobile payment usually cooperates with some shopping platforms to push promotional information for users,

[3] Data resource: https://www.sohu.com/a/232624359_100157588.

which brings better consumer experience and make consumers more accustomed to using the mobile payment for shopping [42, 80]. Finally, mobile payment may also optimize the household consumption structure. For instance, mobile payment provides information to make people pay more attention to the high-level consumption field in the health, education, culture, and entertainment and electronic products, thereby optimizing household consumption structure and enhancing their utilities and happiness.

3.2 Reducing transaction costs

In China, due to the incompletion of the financial system and credit system, traditional non-cash payment methods such as credit card or cheque payment have large transaction costs for common residents or households [60, 78]. Taking credit cards as an example, the penetration rate of credit cards in China is only 34.4%, while in developed countries in Europe and America, the ratio exceeds 85%[4]. Additionally, the non-popularity of credit cards makes them generally only available in high-end consumer places like upscale department stores and ritzy restaurants, which means many people are excluded from credit card payments, especially for socially disadvantaged groups. Furthermore, traditional payment methods limit economic activities across regions [65], and residents have to shop, transfer money, or pay utility bills offline or at bank branches, which incurs more additional transaction costs from wasting time, traffic, communication, and information acquisition.

As a more convenient and alternative tool for traditional financial payment methods, mobile payment may effectively lower the threshold of financial services and reduce the transaction costs of economic activities [56, 80]. First, compared with traditional financial institutions, mobile payment uses Internet technology to realize electronic money payment, simplifying the transaction process and improving payment efficiency [11, 23]. For consumers, these reduced costs include travel time and fees to and from banks, waiting time at bank counters, cash safekeeping costs, and the potential risk of receiving counterfeit currency. Second, mobile payment breaks through the physical limitations and geographical isolation of conventional payment ways, and realizes online remote payment, saving household transport and time costs [54, 77]. Third, mobile payment based on the internet platform also helps to improve people's ability to obtain information, alleviate information asymmetry in transactions [83], and help residents to obtain required goods and services at lower prices [49, 86]. In summary, mobile payment makes financial services more convenient and makes up for the deficiencies of traditional financial tools and brings a better trading experience feeling.

[4] Data source: https://www.quora.com/Why-aren%E2%80%99t-credit-cards-popular-in-China-the-worlds-second-largest-economy.

3.3 Stimulating entrepreneurship

We considered the third possible mechanism of mobile payments affecting residents' happiness is stimulating entrepreneurship. Previous studies have found an inextricable relationship between mobile payment and Chinese households or individual entrepreneurship [29, 73, 79]. First, mobile payments, relying on the Internet and big data, effectively reduce the cost of financial services and broaden the access to funding for entrepreneurs [29, 30]. For example, with the credit points generated by mobile payment usage, entrepreneurs may obtain small loans from Ant Cash Now or Weilidai.[5] These microcredits are simpler and do not require collateral, saving entrepreneurs the cost of obtaining a bank loan, reducing financial constraints on families and providing incentives to start a business [79]. Second, mobile payments, provide an effective channel for information dissemination to explore and grasp more business opportunities through the Internet [39]. Third, mobile payments make online entrepreneurship more possible. Through mobile payments and e-commerce platforms, entrepreneurs may sell their goods or services to all over the world and have a wider source of customers [67, 69]. A cluster effect can be formed among Internet entrepreneurs to realize the optimal allocation of resources.

Given that the positive impact of entrepreneurship or self-employment on happiness has been widely recognized [63, 85], we regarded entrepreneurship would be one of the channels by which mobile payments promote residents' happiness.

3.4 Increasing social interaction

The last possible channel of mobile payments affecting Chinese residents' happiness is social interaction. First, mobile payments increase the frequency with which people participate in social networking. Relying on the Internet platform, mobile payment provides people with an effective means of communication and social contact. For example, WeChat Pay is developed by relying on WeChat, China's largest online social platform. By combining the custom of WeChat red packets with traditional Chinese characteristics, which greatly enhances residents' online social interaction experience [50].

Second, mobile payment provides more convenience for people to participate in online entertainment. At present, China has ushered in the era of online entertainment charges. Watching movies and TV dramas, listening to music, and playing online games on smartphones sometimes need to pay. There is no doubt that mobile payment helps people to have fun online easily.

Third, in addition to online activities, mobile payments also make it easier for offline social interaction [29, 48]. In terms of transportation, people can use mobile payments to take buses, subway, and shared bicycles, and there are discounts on highway payment. Concerning dining out and hotel accommodation, mobile payment can be used to make reservations. In the aspect of communication, mobile

[5] Ant Cash Now and Weilidai are microfinance businesses under Alipay and WeChat Pay.

payment can improve the service of quick inquiry and pay the phone and Internet bills.

Taken together, these conveniences provided by mobile payment create good external conditions for social interaction. Moreover, social interaction, as a social behavior that meets the needs of individual autonomy, promotes interpersonal relationships and increase residents' happiness, which has been verified by many previous studies (e.g., [3, 45, 59]).

4 Empirical strategy

To identify the causal relationship between mobile payment usage and residents' happiness, we must rule out the endogenous problem. On the one hand, there may be a two-way causal relationship between mobile payment usage and residents' happiness. For example, higher happiness may indicate a higher level of education or financial literacy, which makes them more willing to try new digital payment methods. On the other hand, individuals' mobile payment behavior may be self-selected. Whether residents use mobile payments is confounded with a range of factors, such as perceptions, personality, habits, risk awareness, and thinking inertia. Although we reduce potential endogeneity from bidirectional causality and self-selection by controlling for individual characteristics, household characteristics, and city fixed effects, some confounders are not observable or measurable.

Undoubtedly, it is better to use the instrumental variable (IV) method to solve the endogenous problems caused by two-way causality and self-selection bias. Specifically, we utilize the mobile payment usage rate of the others within a community (village) as an IV. According to the theory of peer group effect [7, 18], the mobile payment use of individuals within a community (village) is influenced by others, due to the pervasive nature of mobile payment usage. When the more people within a community (village) choose to use mobile payments, individuals are more likely to follow them and adopt this payment method as well [57]. Therefore, our IV is consistent with the relevant hypothesis. Furthermore, since the IV only includes the mobile payment use of others within a community (village), the happiness of individual residents is unlikely to be influenced by such data aggregated through other channels [5, 27, 82]. In other words, the use of mobile payments by others only affects their own happiness, and it is difficult to make a connection with the individual's well-being.

Another concern is whether to use linear or nonlinear approaches for the estimation of happiness. Ferrer-i-carbonell & Frijters [19] concluded that estimates were relatively unaffected by the assumption that happiness was cardinal or ordinal. In subsequent studies, most of the literature was likely to use both linear and nonlinear estimates to guarantee the reliability of results e.g., [32, 37, 82]. Similar to these studies, in main specifications, we employed ordered probit models and performed robustness tests using linear estimations. In addition, given that our core explanatory variable is binary, we estimated the ordered probit regression with endogenous

treatment [57, 82, 86], including a ordered probit model for residents' happiness and a probit model for mobile payment. Models can be written as follows:

$$Happiness_{ic} = \alpha_0 + \alpha_1 MP_{ic} + \alpha_2 X_{ic} + \alpha_3 City_p + \mu_{ic} \tag{1}$$

$$Happiness_{ic} = \begin{cases} 1 \; if \; Happiness^*_{ic} < \lambda_1; \\ k \; if \; \lambda_{k-1} < Happiness^*_{ic} < \lambda_k, k = 2, 3, 4 \\ 5 \; if \; Happiness^*_{ic} > \lambda_4; \end{cases} \tag{2}$$

$$MP^*_{ic} = \beta_0 + \beta_1 Z_c + \beta_2 X_{ic} + \beta_3 City_p + \varepsilon_{ic} \tag{3}$$

$$MP_{ic} = \begin{cases} 1 \; if \; MP^*_{ic} > 0; \\ 0 \; otherwise \end{cases} \tag{4}$$

where $Happiness_{ic}$ denotes the happiness score of the householder i in the community (village) c. MP^*_{ic} is a latent variable representing householder i's continuously varying propensity of using mobile payment. MP_{ic} is a dummy variable that takes the value of 1 if the household i in the community (village) c uses mobile payment, and 0 otherwise. Z_c refers to the IV measured by the mobile payment usage rate of the others within a community (village) c. X_{ic} denotes control variables, and $City_p$ is a vector of city-level dummies, capturing the prefecture-level city fixed effects. ε_{ic} and μ_{ic} are error terms.

5 Data and variables

5.1 Data source

Our data comes from the CHFS in 2017, which covers 1417 communities (villages) of 326 cities in 29 provinces, with microdata for 40,013 households. The CHFS is conducted by the Southwestern University of Finance and Economics and it's nationally representative of Chinese households due to its stratified sample design. The CHFS collects detailed information on residents' demographic characteristics, household income and expenditure, assets and liabilities, employment and entrepreneurship, and so forth. Especially in this wave, the CHFS investigated whether households use mobile payment. In addition, in the subjective attitude part of the CHFS questionnaire, it recorded the information about residents' happiness. Therefore, the CHFS provides an accessible and reliable source of data for our research on mobile payment usage and residents' happiness.

5.2 Variable description

We assigned the explained variable *Happiness* based on the response of the CHFS questionnaire, "Overall, do you feel happy now?", which was widely used in previous studies (e.g., [19, 31, 32, 45, 85]. We assigned a value of 5 if their answer is "very happy", In turn, values 4 to 1 indicate "happy", "indifferent", "unhappy", and "very unhappy", respectively.

The householders were asked about the payment methods commonly used. The five options are cash payment, card payment (including bank card, credit card, etc.), computer payment (including online banking, computer Alipay, etc.), mobile terminal payment such as mobile phones, iPad (including Alipay, WeChat payment, mobile banking, Apple pay, etc.), and others. Based on responses to this question, we defined the explanatory variable *Mobile payment* equals to 1 if the householders used mobile terminal payment, and 0 otherwise.

As mentioned earlier, we utilized the mobile payment usage rate of the others within a community (village) as an IV. We use the household number provided by CHFS to match households living in the same community (village) and calculate the proportion of other householders who use mobile payments.

Additionally, we included two levels of control variables in this paper. First, at the householder level, considering the potential nonlinear effects, we chose *Age* and *Age squared* [57, 85]. The remaining demographic variables include *Gender*, *Urban*, *Party member*, *Education*, *Marriage Health*, which we considered the positive effects of the female householders, living in urban areas, political identity, good education, and self-rated health on residents' happiness in China [45, 46, 79, 82, 87]. Moreover, since insurances, social trust, and risk preferences may affect both individual mobile payment use and happiness [15, 43, 45], *Commercial insurance*, *Social trust*, and *Risk attitude* are also controlled for.

Second, in the household characteristics, we first controlled for two variables of household demographics, *Child dependency ratio*, *Elderly dependency ratio*, because of the influence of child and elderly dependency ratio on residents' happiness [85]. In addition, among the economic characteristics, we selected a series of variables [12, 15, 20, 37, 43, 57], such as *Household income*, *Fixed deposit*, *Current deposit*, *Household debt*, *Housing ownership*, *Shop ownership*, and *Car ownership*. Last, given the effects of social security on individual happiness [62], *Pension insurance* and *Medical insurance* were further included.

5.3 Descriptive statistics

Variable definitions and descriptive statistics are represented in Table 2, and all continuous variables (e.g., *Child dependency ratio*) are winsorized by replacing samples among the top 1% and the bottom 1% to avoid the influence of extreme values on the estimates. In addition, 18 observations that use mobile payments but not smartphones were removed. To confirm that householders' responses were acceptable and reliable, 27 householders under the age of 18 were also excluded. After excluding missing values for the above variables, our final observation was 36,998. We find

Table 2 Variable definitions and descriptive statistics (N = 36,998)

Variables	Definition	Mean	SD
Happiness	Householder happiness. An ordinal variable with values of 1, 2, 3, 4, and 5, with higher values indicating greater happiness	3.8628	0.8301
Mobile payment	Use = 1; nonuse = 0	0.2667	0.4422
Mobile payment usage rate	Mobile payment usage rate within a community (village)	0.2739	0.2013
Age	Householder age	55.385	14.064
Gender	Householder gender. Female = 1; male = 0	0.2073	0.4054
Hukou	Urban = 1; rural = 0	0.6816	0.4659
Party member	Householder politics status. Member of Communist Party of China = 1; others = 0	0.1254	0.3312
Education	Householder education status. 1 = unschooled, 2 = primary school, 3 = junior middle school, 4 = vocational high school, 5 = senior high school, 6 = junior college, 7 = bachelor degree, 8 = master's degree, PhD degree = 9	3.4324	1.6733
Marriage	Householder marital status. Married = 1; others = 0	0.8528	0.3543
Health	Householder self-rated health. An ordinal variable with values of 1, 2, 3, 4, and 5, with higher values being healthier	3.3824	1.0149
Commercial insurance	Owned = 1; not owned = 0	0.0881	0.2834
Social trust	An ordinal variable with values of 1, 2, 3, 4, and 5, with higher values indicating higher trust in strangers	3.9779	0.9100
Risk attitude	An orderly variable with values of 1, 2, 3, 4, and 5, with higher values indicating a greater preference for risk	1.8517	1.2310
Child dependency ratio	The ratio of the number of the elderly over 65 divided by the number of people aged 15–64 in the household	0.1055	0.1748
Elderly dependency ratio	The ratio of the number of the children less than 14 divided by the number of people aged 15–64 in the household	0.2304	0.3560
Household income	Household per capita income, 10,000 yuan	4.0449	2.2376
Fixed deposit	Household fixed deposit, 10,000 yuan	1.7588	5.9317
Current deposit	Household current deposit, 10,000 yuan	2.1919	5.5271
Household debt	Household total debt, 10,000 yuan	4.6672	14.423
Pension insurance	Proportion of people aged 16 years or older in the household with social pension insurance	0.7476	0.3463

Table 2 (continued)

Variables	Definition	Mean	SD
Medical insurance	Proportion of people aged 16 years or older in the household with social medical insurance	0.8923	0.3777
Housing ownership	Owned = 1; not owned = 0	0.8440	0.3628
Shop ownership	Owned = 1; not owned = 0	0.0357	0.1856
Car ownership	Owned = 1; not owned = 0	0.2477	0.4317

"yuan" is the unit of the RMB, which is the official currency of China

Table 3 The effect of mobile payments on residents' happiness

	(1)	(2)	(3)
	Ordered probit	IV (Ordered probit regression with endogenous treatment)	
	Happiness	Mobile payment (Probit)	Happiness (Ordered probit)
Mobile payment	0.0010		0.1780***
	(0.0045)		(0.0239)
Mobile payment usage rate		0.2226***	
		(0.0135)	
Age	− 0.0079***	− 0.0112***	− 0.0030***
	(0.0009)	(0.0012)	(0.0012)
Age squared	0.0102***	0.0045***	0.0071***
	(0.0009)	(0.0011)	(0.0010)
Gender	0.0133***	0.0078	0.0098**
	(0.0042)	(0.0048)	(0.0043)
Hukou	− 0.0101**	0.0651***	− 0.0271***
	(0.0045)	(0.0054)	(0.0051)
Party member	0.0220***	− 0.0145**	0.0254***
	(0.0047)	(0.0058)	(0.0048)
Education	− 0.0045***	0.0262***	− 0.0110***
	(0.0013)	(0.0013)	(0.0016)
Marriage	0.0746***	− 0.0417***	0.0796***
	(0.0056)	(0.0065)	(0.0056)
Health	0.0594***	0.0111***	0.0569***
	(0.0019)	(0.0020)	(0.0020)
Commercial insurance	0.0092	0.0528***	− 0.0057
	(0.0057)	(0.0060)	(0.0061)
Social trust	0.0075***	− 0.0225***	0.0123***
	(0.0019)	(0.0020)	(0.0020)
Risk attitude	− 0.0064***	0.0177***	− 0.0102***
	(0.0014)	(0.0015)	(0.0015)
Child dependency ratio	− 0.0033	0.0189	− 0.0096
	(0.0108)	(0.0119)	(0.0109)
Elderly dependency ratio	0.0358***	− 0.1054***	0.0466***
	(0.0071)	(0.0090)	(0.0072)
Household income	0.0007	− 0.0134***	0.0024**
	(0.0010)	(0.0012)	(0.0011)
Fixed deposit	0.0012***	0.0020***	0.0007**
	(0.0003)	(0.0003)	(0.0003)
Current deposit	0.0015***	0.0039***	0.0005
	(0.0003)	(0.0003)	(0.0003)
Household debt	− 0.0003***	0.0003**	− 0.0005***
	(0.0001)	(0.0001)	(0.0001)

Table 3 (continued)

	(1)	(2)	(3)
	Ordered probit	IV (Ordered probit regression with endogenous treatment)	
	Happiness	Mobile payment (Probit)	Happiness (Ordered probit)
Pension insurance	0.0381***	− 0.0059	0.0352***
	(0.0054)	(0.0056)	(0.0054)
Medical insurance	0.0101*	0.0898***	− 0.0013
	(0.0052)	(0.0060)	(0.0056)
Housing ownership	0.0228***	− 0.0164***	0.0285***
	(0.0046)	(0.0054)	(0.0047)
Shop ownership	0.0383***	0.0548***	0.0268***
	(0.0088)	(0.0088)	(0.0090)
Car ownership	0.0326***	0.0881***	0.0057
	(0.0041)	(0.0042)	(0.0055)
$\rho_{\epsilon\mu}$		− 0.3824***	
		(0.0495)	
City fixed effects	Yes	Yes	Yes
N	36,998	36,998	36,998

Column (1) reports the marginal effects when the explained variable *Happiness* is equal to 5. The ordered probit regression with endogenous treatment are used in the last two columns. Among them, column (2) shows the estimated results of the treatment equation and reports the marginal effects and column (3) presents the results of the outcome equation and reports the marginal effects when the explained variable *Happiness* is equal to 5

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the community (village) level are reported in parentheses

that the average happiness scores were 3.8628 and the percentage of residents using mobile payments was 26.67%, indicating that mobile payment services were not nationally widespread in 2017. The high standard deviation of *Mobile payment* suggests that the development of mobile payments and digital finance in China may be uneven.

6 Results

6.1 Baseline results

Table 3 reports the baseline results of the impact of mobile payment use on residents' well-being, with all standard errors have been cluster-corrected at the community (village) level. In column (1), the results by using the ordered probit method show that there is no association between mobile payment usage and residents'

happiness. However, the ordered probit estimation results may be biased because of potential endogeneity problems.

Columns (2) and (3) in Table 3 present results by using the ordered probit regression with endogenous treatment, where column (2) shows the results of the treatment equation and reports the marginal effects. The correlation coefficient ($\rho_{\varepsilon\mu}$), presented in the lower part of Table 3, is significant, indicating the presence of selection bias from unobserved factors. That is, it is more appropriate to use IV estimations [87, 88]. The IV, the mobile payment usage rate of the others within a community (village), is strongly positively correlated with household mobile payment usage status at the 1% significance level, indicating that the IV meets the correlation restriction. Column (3) presents the results of the outcome equation and reports the marginal effects when *Happiness* is equal to 5. We find that the coefficient on mobile payment is significantly positive at the 1% level, suggesting that mobile payment usage did increase Chinese residents' happiness after excluding potential endogeneity. Specifically, the results reveal that mobile payment usage increased the possibility that people rated their happiness as "very happy" by 17.80%. Furthermore, Table 12 in the appendix shows the results for happiness assignments of 1–4 and suggests that mobile payments reduced the probability that residents rated their happiness as "very unhappy", "unhappy", and "indifferent" by 2.00%, 4.69%, and 13.96%, respectively, and increased the probability of feeling "happy" by 2.85%. In summary, we caught a significant causal relationship between mobile payment usage and residents' happiness by adopting an IV method.

Among the control variables in column (3), it is clear that the coefficient on *Age squared* is significantly positive, suggesting that there is a U-shaped relationship between the age of householders and happiness. These results indicate that younger and older people are happier in China [85]. In addition, party membership, marriage, health, and social trust were positively associated with individual happiness, which is consistent with the findings in previous studies [20, 46, 79]. For household economic variables, the significantly positive coefficients on *Household income*, *Fixed deposit*, *Housing ownership* and *Shop ownership* indicate that householders' happiness will rise when they have higher assets, income, and savings [12, 15, 37, 57]. A significantly negative coefficient is found on *Household debt*, implying excessive household debt undermines people's happiness [43]. Moreover, there is a positive association between household pension insurance ownership rates and householders' happiness, which is consistent with the conclusions obtained in most of the prior literature [62, 88].

6.2 Mechanism analysis

In this subsection, we examined multiple potential mechanisms through which mobile payment usage increases residents' happiness. These channels include improving quality of life, reducing transaction costs, stimulating entrepreneurship, and increasing social interaction. Moreover, estimation methods are adjusted according to the types of mechanism variables.

Table 4 Mobile payments, quality of life, and residents' happiness

	(1) Per capita household consumption	(2) Engel's coefficient	(3) Enjoyable consumption	(4) Leisure consumption	(5) Educational consumption
Panel A. Quality of life (IV estimations)					
Mobile payment	0.5757***	-0.1691***	0.1595***	0.1738***	0.2940***
	(0.0540)	(0.0099)	(0.0084)	(0.0132)	(0.0184)
Control variables	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes
N	36,998	36,998	36,998	36,998	36,998
Panel B. Happiness (Ordered probit)					
Quality of life	0.0024***	-0.0441***	0.0020***	0.0011***	0.00036*
	(0.0009)	(0.0090)	(0.0004)	(0.0003)	(0.00021)
Control variables	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes
N	36,998	36,998	36,998	36,998	36,998

All columns in Panel A show the outcome equation results for the endogenous treatment effect model. Columns in Panel B report the marginal effects when *Happiness* is equal to 5. Baseline control variables and fixed effects are added in all regressions

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the community (village) level are reported in parentheses

6.2.1 Improving the quality of life

We considered whether mobile payments improve the quality of life from the perspective of consumption upgrading. The reason why consumption is used as a measure of quality of life is plenty of studies confirmed the close relationship between consumption and household quality of life [31, 88], and some classical economic models widely assumed that consumption is a central element in maximizing household utility [8, 16]. According to the CHFS questionnaire, mechanism variables we selected include *Per capita household consumption*, *Engel's coefficient*, *Enjoyable consumption*, *Leisure consumption,* and *Education consumption*. The specific definitions of these variables are shown in Panel A of Table 13 in the Appendix.

Panel A of Table 4 reports the results of the impact of mobile payments on quality of life, using the endogenous treatment effect model. In column (1), the estimates show that mobile payments contribute to improving residents' happiness, with statistical significance at the 1% level, which is similar to the results of [38]. In column (2), we can see that the coefficient on *Mobile payment* is significantly negative, implying that mobile payments reduce the proportion of food consumption expenditures and enhances the quality of life of the household. More importantly, in the last three columns, mobile payments have positive effects on various types of consumption that are closely related to individual happiness, indicating that mobile payments enhance people's willingness to consume in enjoyment, leisure, and education. In Panel B of Table 4, as expected, consumption is positively correlated with residents' happiness, and there is a negative association between Engel's coefficient and happiness. To sum up, these results indicate that mobile payments facilitate the consumption upgrading and improve the consumption structure for more utility and happiness.

6.2.2 Reducing transaction costs

As highlighted in Sect. 3, mobile payments provide convenience and enhance residents' happiness by reducing transaction costs. High transaction costs are reflected in the payment methods offered by traditional financial institutions, such as cash payments and card payments. First, we aimed to confirm that mobile payments save the high costs associated with these traditional payment methods, and results using the endogenous bivariate probit model are shown in columns (1) and (2) of Table 5. In Panel A, the coefficient on *Mobile payment* is significantly negative in column (1), indicating that mobile payment usage reduces the frequency of cash transactions in daily life. In column (2), the results suggest that mobile payments facilitate card payments. Additionally, in Panel B, although there is a negative relationship between cash payments and residents' happiness, the effect of card payments on happiness is insignificant. These results indicate that mobile payments reduce transaction costs to enhance people's happiness mainly by reducing cash payments, possibly thanks to the psychological account loss effect of mobile payments and the avoidance of the risk of receiving counterfeit bills and cash theft [23, 42].

Second, we tested whether mobile payments promote residents' happiness through online shopping, which also brings more convenience and less transaction

Table 5 Mobile payments, transaction costs, and residents' happiness

	(1) Cash payment	(2) Card payment	(3) Online shopping (dummy)	(4) Online shopping consumption
Panel A. Other payment methods and online shopping (IV estimations)				
Mobile payment	− 0.1405***	0.2035***	0.2444***	0.9187***
	(0.0089)	(0.0129)	(0.0248)	(0.1681)
Control variables	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes
N	36,998	36,998	36,894	12,442
Panel B. Happiness (Ordered probit)				
Other payment methods and online shopping	− 0.0394***	0.0047	0.0144***	0.0094**
	(0.0076)	(0.0044)	(0.0042)	(0.0049)
Control variables	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes
N	36,998	36,998	36,894	12,442

The first three columns of Panel A show the outcome equation results for the endogenous bivariate probit model, reporting the marginal effects, and the endogenous treatment effect model is used in column (4) of Panel A. All columns in Panel B report the marginal effects when *Happiness* is equal to 5. Baseline control variables and fixed effects are added in all regressions

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the community (village) level are reported in parentheses

costs. Given that the CHFS not only investigated whether respondents shopped online, but also recorded the consumption of online shopping, we accordingly constructed two mechanism variables to measure online shopping. The specific definitions of these variables are shown in Panel B of Table 13 in the Appendix. Columns (3) to (4) of Table 5 report the estimated results, and we find whether online shopping or online shopping consumption, mobile payments both have positive effects in Panel A, implying that mobile payment usage is a vital driving force for more active household online economic activities. Moreover, in Panel B, two online shopping variables are also significant at the 5% and above level, suggesting that mobile payments reduce transaction costs through online shopping to increase residents' happiness.

6.2.3 Stimulating entrepreneurship

Next, we considered whether mobile payments stimulate household entrepreneurship. We constructed three dummies to measure household entrepreneurial behaviors, *Entrepreneurship*, *Entrepreneurial intention*, and *Online business*, of which the details are in Panel C of Table 13 in the Appendix.

In columns (1), (3) and (5) of Table 6, the estimates from the endogenous bivariate probit model suggest that mobile payment usage significantly increase the

Table 6 Mobile payments, entrepreneurial behaviors, and residents' happiness

	(1) Entrepreneurship	(2) Happiness	(3) Entrepreneurial intention	(4) Happiness	(5) Online Business	(6) Happiness
	IV	Ordered probit	IV	Ordered probit	IV	Ordered probit
Mobile payment	0.0272***		0.0286***		0.0477***	
	(0.0033)		(0.0032)		(0.0066)	
Entrepreneurial behavior		0.0102**		0.0108**		0.0243***
		(0.0042)		(0.0043)		(0.0089)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	36,998	36,998	36,778	36,778	36,986	36,986

Columns (1), (3), and (5) show the outcome equation results for the endogenous bivariate probit model, reporting the marginal effects. Columns (2), (4), and (6) report the marginal effects when *Happiness* is equal to 5. Baseline control variables and fixed effects are added in all regressions

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the community (village) level are reported in parentheses

likelihood of entrepreneurship and having entrepreneurial intention and online businesses. Consistent with previous studies, mobile payments are certainly helping Chinese families start businesses [29, 73, 79]. In addition, the results of columns (2), (4), and (6) show that there are positive relationships between these entrepreneurship-related variables and residents' happiness, which are consistent with the previous studies [63, 85]. All in all, we provided evidence that mobile payments can promote residents' happiness by stimulating household entrepreneurship.

6.2.4 Social interaction

As noted above, mobile payment may increase residents' happiness by social interaction. According to the CHFS questionnaire, we considered three important categories of social interaction that are closely associated with individual happiness, including network interaction, human relationship spending, and daily communication. In terms of network interaction, two binary variables *Online social activities* and *Online entertainment* were selected. *Money gift receive* was used to measure the household human relationship spending. For daily communication, we chose *Traffic expenditure* and *Communication expenditure* for proxies. These variables are defined as shown in Panel D of Table 13 in the Appendix.

Table 7 explored the effects of mobile payment on household social interaction. In Panel A, we find that the coefficients on *Mobile payments* are significantly positive in all columns, which means households that use mobile payments are more active in social interaction than those who don't. For example, mobile payment has increased the probability of residents' online social activities by 7.21%, and the probability of online entertainment has also increased by 5.64%. In Panel B, the ordered probit estimates suggest a positive association between social interaction and individual happiness, with the exception of traffic expenditure, which is consistent with previous studies [3, 59]. Therefore, we concluded that mobile payments bring happiness by facilitating social interactions.

6.3 Further studies

6.3.1 Do mobile payments lead to over-consumption and over-indebtedness?

Previous literature found that the painless shopping and income illusion generated by non-cash payments may cause people to over-consume and fall into debt cages [44]. With the increasing convenience of online shopping and non-cash payments, the phenomenon of unplanned consumption, irrational consumption, and compulsive buying may be more and more serious [66, 77]. As an emerging non-cash payment method and a bridge to online shopping, whether mobile payments lead Chinese households to over-consumption or over-indebtedness to damage happiness needs further study.

We constructed two binary variables, *Over-consumption* and *Over-indebtedness*, of which definitions are reported in Pane E of Table 13 in the Appendix. In columns (1) and (2) of Table 8, we find that mobile payments significantly increase the

Table 7 Mobile payments, social interaction, and residents' happiness

	(1) Online social activities	(2) Online entertainment	(3) Money gift receive	(4) Traffic expenditure	(5) Communication expenditure
Panel A. Social interaction (IV estimations)					
Mobile payment	0.0721***	0.0564**	0.3059***	1.6820***	0.5522***
	(0.0268)	(0.0264)	(0.0270)	(0.1090)	(0.0437)
Control variables	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes
N	36,986	36,986	36,887	36,998	36,998
Panel B. Happiness (Ordered probit)					
Social interaction	0.0074*	0.0097**	0.0191***	0.0001	0.0041***
	(0.0040)	(0.0043)	(0.0032)	(0.0003)	(0.0009)
Control variables	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes
N	36,986	36,986	36,887	36,998	36,998

The first three columns of Panel A show the outcome equation results for the endogenous bivariate probit model, reporting the marginal effects, and the endogenous treatment effect model is used in the last two columns of Panel A. All columns in Panel B report the marginal effects when *Happiness* is equal to 5. Baseline control variables and fixed effects are added in all regressions

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the community (village) level are reported in parentheses

Table 8 Mobile payments and over-consumption and over-indebtedness

	(1) Over-consumption	(2) Happiness	(3) Over-indebtedness	(4) Happiness
	IV	Ordered probit	IV	Ordered probit
Mobile payment	0.2788***		− 0.0263	
	(0.0247)		(0.0383)	
Over-consumption		− 0.0290***		
		(0.0035)		
Over-indebtedness				− 0.0565***
				(0.0055)
Control variables	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes
N	36,998	36,998	36,998	36,998

Columns (1) and (3) show the outcome equation results for the endogenous bivariate probit model, reporting the marginal effects. Columns (2) and (4) report the marginal effects when *Happiness* is equal to 5. Baseline control variables and fixed effects are added in all regressions

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the community (village) level are reported in parentheses

probability of household over-consumption, and over-consumption is indeed negatively related to the happiness of householders. These results imply that over-consumption caused by mobile payments leads to a decrease in residents' happiness. In addition, we find that although over-indebtedness is negatively related to individual happiness in column (4), the coefficient on *Mobile payment* is insignificant in column (3), suggesting that the use of mobile payment is not sufficient to cause over-indebtedness among Chinese households. Overall, our estimates indicate that mobile payments may cause a decline in happiness through over-consumption, implying a potential mechanism by which mobile payments impair residents' happiness.

6.3.2 Heterogeneity analysis

Considering that different individual and household characteristics may have impacts on residents' mobile payment usage status and happiness, we further explored the heterogeneity effect from multiple perspectives. First, based on the traditional definition, we divided the sample into two groups, the young and middle-aged and the elderly, using age 60 as the cutoff. Columns (1) and (2) of Table 9 report estimation results from the ordered probit regression with endogenous treatment. We find that mobile payments only increase the happiness of elderly residents, while having a significant negative effect on younger people instead. These results may be related to inclusive financial attributes of digital finance and mobile payments [6, 30, 79]. As a typically socially disadvantaged group, elderly individuals are more likely to experience financial exclusion [13]. The popularity of mobile payments has provided greater convenience for some older adults to access financial services, which in turn enhances their happiness.

Table 9 Heterogeneity effects: IV estimations

	(1) Young and middle-aged	(2) Elderly	(3) Urban	(4) Rural	(5) Better educated	(6) Less educated	(7) High-income households	(8) Low-income households
Mobile payment	− 0.1001*	0.2220***	0.1380***	0.2657***	0.0196	0.2187***	0.0533	0.2145***
	(0.0544)	(0.0639)	(0.0272)	(0.0912)	(0.0416)	(0.0275)	(0.0328)	(0.0338)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	21,873	15,125	25,126	11,872	13,105	23,893	18,489	18,509

All columns show the results of the outcome equation by using the ordered probit regression with endogenous treatment, reporting the marginal effects when *Happiness* is equal to 5. Baseline control variables and fixed effects are added in all regressions

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the community (village) level are reported in parentheses

Second, China's dualistic urban–rural economic model may not only make urban and rural households differ in their acceptance and frequency of mobile payment use [79, 87], but also in their happiness [32]. Therefore, we divided the sample into two subsamples according to the *hukou*. In columns (3) and (4) of Table 9, the coefficients on *Mobile payment* are both significantly positive, but the values are larger in the rural subsample. Similarly, rural areas in China face the problems of insufficient financial supply and poor financial accessibility for residents [25, 71]. The rise of digital finance, represented by mobile payments, has effectively broadened the scope of financial services, allowing more rural residents to enjoy financial services. Therefore, mobile payments have a higher marginal utility in improving the happiness of rural households.

Third, we sought to examine the educational heterogeneity. In our sample, the median education level of householders is junior high school, so we defined the less educated group as those who have not attended school or only attended primary and junior high school. The columns (5) and (6) of Table 9 show that in the subsample of householders with relatively low education levels, the coefficient on *Mobile payment* is significantly positive, suggesting that mobile payments have a greater impact on enhancing the happiness of socially disadvantaged groups with low education.

Last, due to the dual impact of household income on residents' mobile payment usage and happiness [12, 48], heterogeneity effects by income deserve to be further explored. The full sample is divided into two groups respectively, based on the median of total household income (54,000 yuan), and results by using the ordered probit regression with endogenous treatment are shown in the last two columns of Table 9. The positive effect of mobile payments on residents' happiness is only found in the low-income group, indicating that mobile payments promoted happiness of residents with relatively disadvantaged incomes. These estimates also reflect the inclusive financial characteristics of mobile payments in terms of individual happiness.

6.4 Robustness checks

We conducted several robustness checks. First, we used the provincial mobile Internet access ratio as an alternate IV, and the data come from the China Statistical Yearbook. On the one hand, mobile Internet is the basis for residents to use smart phones, thus, provincial network access ratio is positively correlated with residents' mobile payment usage. On the other hand, the proportion of provincial mobile network access as a macro aggregation variable is difficult to directly individual happiness when city fixed effects are controlled [27]. In column (1) of Table 10, we find that the results are unchanged.

Second, in our main specifications, we employed the ordered probit regression with endogenous treatment. Now we utilized the endogenous treatment effect model (linear) to perform robustness tests [82]. In column (2) of Table 10, it is clear that our estimates are similar with the main results in column (3) of Table 3, indicating that our results are insensitive to the cardinal and ordinal assumptions of happiness [19, 37].

Table 10 Robustness checks of using alternative specifications: IV estimations

	(1) Using the alternative IV	(2) Endogenous treatment effect model	(3) Controlling for age and education fixed effects	(4) Adding macro control variables	(5) Binary variable for *Happiness*
Mobile payment	0.1246***	0.2867***	0.1901***	0.1764***	0.0921***
	(0.0260)	(0.0454)	(0.0234)	(0.0269)	(0.0264)
Control variables	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes
N	36,998	36,998	36,998	36,998	36,998

Columns (1), (3), and (4) show the results of the outcome equation by using the ordered probit regression with endogenous treatment, reporting the marginal effects when *Happiness* is equal to 5. Column (2) shows the outcome equation results for the endogenous treatment effect model. Columns (5) presents the outcome treatment equation results for the endogenous bivariate probit model, reporting the marginal effects. Baseline control variables and fixed effects are added in all regressions

***, **, * indicate significant at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the community (village) level are reported in parentheses

Third, according to the heterogeneity analysis, age and education are strongly associated with individuals' use of mobile payments, so we attempted to further control for age and education fixed effects. In column (3) of Table 10, the estimates show the conclusion that mobile payments enhance residents' happiness remains unchanged, suggesting that our main specification is satisfactory for capturing age and education fixed effects.

Forth, although we controlled for city fixed effects in main regressions, the spread of digital finance and mobile payments is associated with regional economic development and financial development, and previous studies also found that these macro variables could affect residents' happiness [37, 82]. We added some city characteristics variables for robustness tests, including per capita GDP, savings, public expenditure, and air pollution[6]. In column (5) of Table 10, the estimated coefficient on *Mobile payment* keeps significantly positive at the 1% level.

Fifth, we changed the definition of the core explained variable by converting it into a binary variable [85]. In other words, we assigned the value of 1 if householders' answer is "very happy" or "happiness", and 0 otherwise. We employed the probit method to estimate the outcome method of the endogenous treatment effect [27], and results are presented in column (5) of Table 10. We can see that a similar positive effect of mobile payment usage on residents' happiness, suggesting that our results are also insensitive to change the definition of our core explained variable.

[6] The unit of per capita GDP is 10,000 yuan. Public expenditure is measured by the proportion of prefecture-level city government expenditure to GDP. Savings is defined as the radio of savings of total financial institutions in GDP, and air pollution is the amount of smoke and sulfur dioxide emissions per unit of GDP.

Last, we excluded some extreme observations, including those over 80 years old, because the oldest old may have greater difficulty using mobile payments. Householders that are very unhealthy[7] and have not attended school were excluded, due to the potential adverse impact of the health and education of the use of mobile payment according to previous estimated results. Observations with very low social trust were removed[8] because they may be less receptive to emerging payment methods. Moreover, we removed the sample that used other electronic payment tools, like computer payments, so that the control group that did not use mobile payments would only use traditional payment methods. In Table 11, we find that excluding these extreme observations has little effect on our main results.

7 Conclusion and discussions

7.1 Main findings

Previous studies focused on online happiness and the potential benefits of inclusive finance [6, 21, 57, 81, 86], but rarely addressed the happiness effects of digital finance or mobile payments. In this study, adopting the nationally representative microdata from the CHFS, we attempted to investigate the effects of mobile payment usage on people's happiness in China. Considering the selectivity of mobile payment use, we employed the ordered probit regression with endogenous treatment to overcome potential endogeneity issues. Our main results suggested that mobile payments significantly improve residents' happiness, which is supported by a series of robustness checks, such as using an alternative IV, replacing the explained variable, and removing some extreme observations.

Moreover, we explored several potential mechanisms by which mobile payments affect residents' happiness. We found that the positive impact from mobile payment usage can be explained by promoting quality of life, reducing transaction costs, stimulating entrepreneurship, and increasing social interaction. Interestingly, our results confirmed a mechanism by which mobile payments undermined people's happiness, over-consumption. Furthermore, we analyzed the heterogeneity effects from several aspects and tried to verify the digital financial inclusion characteristics of mobile payments. Our estimates showed that the happiness effect of mobile payments was valid for some socially vulnerable groups, such as elderly individuals, rural residents, the low-educated, and low-income households. These results indicated that digital finance, represented by mobile payments, can reach more socially disadvantaged people and give them more access to financial services to enhance their welfare.

[7] Observations with an assigned value of 1 for self-rated health are deleted.
[8] Here, the householders whose social trust is assigned to 1 are removed.

Table 11 Robustness checks of removing some extreme observations: IV estimations

	(1) Excluding 80 years old and above	(2) Excluding very unhealthy householders	(3) Excluding unschooled householders	(4) Excluding households with low social trust	(5) Excluding samples using other electronic payments
Mobile payment	0.1724***	0.1338***	0.1649***	0.1879***	0.1947***
	(0.0254)	(0.0257)	(0.0243)	(0.0226)	(0.0391)
Control variables	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes
N	35,393	29,956	34,340	34,781	27,620

All columns show the results of the outcome equation by using the ordered probit regression with endogenous treatment, reporting the marginal effects when *Happiness* is equal to 5. Baseline control variables and fixed effects are added in all regressions

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the community (village) level are reported in parentheses

7.2 Theoretical implications

Even though the benefits of digital finance and mobile payments received some attention in prior literature (e.g., [2, 29, 30, 56, 73, 79, 83], they ignored the impact on residents' happiness [87]. Our findings further confirmed the role of mobile payments in enhancing people's happiness by using an IV method. Moreover, we reconciled previous theories about online happiness. The results from mechanism analysis showed that mobile payments improve people's happiness through multiple channels,however, over-consumption behaviors caused by mobile payments may harm users' happiness, which complemented previous literature emphasizing the potential risks of the mobile Internet [34, 42, 81].

Additionally, the findings of this paper add the research on the field of inclusive finance. Although the inclusive properties of digital finance have received attention from scholars [6, 36, 56, 61], it is still unknown how mobile payments enhance the happiness of socially disadvantaged groups. Through heterogeneity analysis, this study reveals the inclusive characteristics of digital finance from the perspective of subjective well-being.

7.3 Managerial implications

This research has several practical implications. To raise residents' happiness, it is necessary to develop a better mobile payment system to reduce the risk of over-consumption. Mobile payment practitioners should increase over-consumption alerts for their products and services. While it may be impossible to completely eliminate risks from over-consumption, making users aware of potential risks will lower their expectations and improve their happiness. In addition, some additional features could be implemented in the design of mobile payment-related products, such as consumption assessment and optimization of consumption information push.

Given the inclusive nature of mobile payments, there is a need to actively promote digital finance to reach more socially disadvantaged groups to improve their happiness. Feasible methods include enhancing publicity through electronic advertisements, on-site publicity, sitcoms, and short videos, introducing personalized mobile payment products and services, and providing appropriate business tilts and subsidies to rural and remote areas. In addition, mobile payment can be further promoted in conjunction with online shopping platforms, telecommunication companies, e-commerce sites, and social software to increase mobile payment usage among socially disadvantaged groups.

Moreover, considering the negative impact of the global outbreak of COVID-19 on people's offline economic activities, adopting flexible ways to promote mobile payments might help more households compensate for the loss of welfare caused by COVID-19 through online activities [41]. These measures are not only applicable to China, but also useful for other countries suffering from the COVID-19. Taken

together, we highlighted the potential positive effect of digital finance and mobile payments on people's happiness during the epidemic.

7.4 Limitations and future research

Some limitations of this study warrant discussion. First, in the linear IV method, two-stage least squares (2SLS), we can identify the validity of an IV by observing multiple statistics, such as the F value of the first stage, Anderson-Rubin Wald test, and Kleibergen-Paap rk LM statistic [38, 84]. However, considering the type of our key explanatory and explained variables, these statistics are not available using the ordered probit regression with endogenous treatment. Second, even though our study addressed one potential mechanism by which mobile payments are detrimental to people's happiness, other adverse channels, such as privacy concerns, digital divide, and potential risks [34, 42, 81], could not be verified due to unavailability of data. These factors that cause digital finance or mobile payments to harm people's welfare should be supplemented in follow-up research.

Appendix

Tables 12 and 13.

Table 12 Marginal effects of happiness in column (3) of Table 3

	Happiness = 4	Happiness = 3	Happiness = 2	Happiness = 1
	(1)	(2)	(3)	(4)
Mobile payment	0.0285***	− 0.1396***	− 0.0469***	− 0.0200***
	(0.0033)	(0.0162)	(0.0069)	(0.0039)
Control variables	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes
N	36,998	36,998	36,998	36,998

All columns show the results of the outcome equation by using the ordered probit regression with endogenous treatment, reporting the marginal effects when the explained variable *Happiness* is 4, 3, 2, and 1, respectively. Baseline control variables and fixed effects are added in all regressions

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the community (village) level are reported in parentheses

Table 13 Definition of mechanism variables

Variables	Definition
Panel A. Quality of life	
Per capita household consumption	Total household expenditure divided by household size (10,000 yuan)
Engel's coefficient	The ratio of household meal expenditure to total consumption
Enjoyable consumption	Household spending on beauty and luxury goods (10,000 yuan)
Leisure consumption	Household spending on tourism, culture, and entertainment (10,000 yuan)
Educational consumption	Household spending on education and training (10,000 yuan)
Panel B. Other payment methods and online shopping	
Cash payment	Use = 1; nonuse = 0
Card payment	Use = 1; nonuse = 0
Online shopping (dummy)	= 1 if the householder shops online last year; others = 0
Online shopping consumption	Total household spending on online shopping last year (yuan, logarithmic)
Panel C. Entrepreneurial behaviors	
Entrepreneurship	= 1 if the household has entrepreneurial behaviors (running a business or project); others = 0
Entrepreneurial intention	= 1 if the householder has the desire to start a business in the future; others = 0
Online Business	= 1 if the householder starts a business or sells goods on the Internet; others = 0
Panel D. Social interaction	
Online social activities	= 1 if the householder conducts online social activities (e.g., making friends); others = 0
Online entertainment	= 1 if the householder conducts online recreational activities (e.g., playing games, listening to songs, and watching movies); others = 0
Money gift receive	= 1 if the householder receives more than 100 yuan in cash or non-cash income from non-family members; others = 0
Traffic expenditure	Total household transportation expenses last year (1000 yuan)
Communication expenditure	Total household telephone and Internet expenses last year (1000 yuan)
Panel E. Over-consumption and over-indebtedness	
Over-consumption	= 1 if the total household consumption is not less than income; others = 0
Over-indebtedness	= 1 if the total household debt is not less than income; others = 0

Funding This work was supported by the Fundamental Research Funds for the Central Universities (No. QCDC-2020-21), and the Shanghai Philosophy and Social Science Planning Annual Project (No. 2021BJB003).

Declarations

Conflict of interest The authors declare that there are no conflict of interest.

References

1. Acker, A., & Murthy, D. (2020). What is Venmo? A descriptive analysis of social features in the mobile payment platform. *Telematics and Informatics, 52*, 101429.
2. Beck, T., Pamuk, H., Ramrattan, R., & Uras, B. (2018). Payment instruments, finance and development. *Journal of Development Economics, 133*, 162–186.
3. Bekalu, M. A., McCloud, R. F., Minsky, S., & Viswanath, K. (2021). Association of social participation, perception of neighborhood social cohesion, and social media use with happiness: Evidence of trade-off. *Journal of Community Psychology, 49*(2), 432–446.
4. Boonchoo, P., & Thoumrungroje, A. (2017). A cross-cultural examination of the impact of transformation expectations on impulse buying and conspicuous consumption. *Journal of International Consumer Marketing, 29*(3), 194–205.
5. Bucher-Koenen, T., & Lusardi, A. (2011). Financial literacy and retirement planning in Germany. *Journal of Pension Economics and Finance, 10*(4), 565–584.
6. Butticè, V., & Vismara, S. (2021). Inclusive digital finance: The industry of equity crowdfunding. *Journal of Technology Transfer*. https://doi.org/10.1007/s10961-021-09875-0
7. Card, D., & Krueger, A. B. (1996). Labor market effects of school quality: Theory and evidence. NBER Working Papers, No. 5450.
8. Cass, D. (1965). Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies, 32*(3), 233–240.
9. Carter, S., & Yeo, A. C. M. (2016). Mobile apps usage by Malaysian business undergraduates and postgraduates: Implications for consumer behaviour theory and marketing practice. *Internet Research, 26*(3), 733–757.
10. Chamon, M., Liu, K., & Prasad, E. S. (2013). Income uncertainty and household savings in China. *Journal of Development Economics, 105*, 164–177.
11. Chen, L. D. (2008). A model of consumer acceptance of mobile payment. *International Journal of Mobile Communications, 6*(1), 32–52.
12. Cheng, Z., Prakash, K., Smyth, R., & Wang, H. (2020). Housing wealth and happiness in urban China. *Cities, 96*, 102470.
13. Clifton, J., Fernández-Gutiérrez, M., & García-Olalla, M. (2017). Including vulnerable groups in financial services: Insights from consumer satisfaction. *Journal of Economic Policy Reform, 20*(3), 214–237.
14. Deidda, M. (2014). Precautionary saving under liquidity constraints: Evidence from Italy. *Empirical Economics, 46*(1), 329–360.
15. Delis, M. D., & Mylonidis, N. (2015). Trust, happiness, and households' financial decisions. *Journal of Financial Stability, 20*, 82–92.
16. Diamond, P. A. (1965). National debt in a neoclassical growth model. *American Economic Review, 55*(5), 1126–1150.
17. Duan, J., & Dholakia, R. R. (2017). Posting purchases on social media increases happiness: The mediating roles of purchases' impact on self and interpersonal relationships. *Journal of Consumer Marketing, 34*(5), 404–413.
18. Evans, W. N., Oates, W. E., & Schwab, R. M. (1992). Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy, 100*(5), 966–991.
19. Ferrer-i-Carbonell, A., & Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? *Economic Journal, 114*(497), 641–659.

20. Fong, E., Yuan, Y., & Gan, Y. (2021). Homeownership and happiness in urban China. *Journal of Housing and the Built Environment, 36*(1), 153–170.
21. Fungáčová, Z., & Weill, L. (2015). Understanding financial inclusion in China. *China Economic Review, 34*, 196–206.
22. Grieve, R., Indian, M., Witteveen, K., Tolan, G. A., & Marrington, J. (2013). Face-to-face or Facebook: Can social connectedness be derived online? *Computers in Human Behavior, 29*(3), 604–609.
23. Hancock, D., Humphrey, D. B., & Wilcox, J. A. (1999). Cost reductions in electronic payments: The roles of consolidation, economies of scale, and technical change. *Journal of Banking & Finance, 23*(2–4), 391–421.
24. He, H., Kukar-Kinney, M., & Ridgway, N. M. (2018). Compulsive buying in China: Measurement, prevalence, and online drivers. *Journal of Business Research, 91*, 28–39.
25. He, J., & Li, Q. (2020). Can online social interaction improve the digital finance participation of rural households? *China Agricultural Economic Review, 12*(2), 295–313.
26. Holm, M. B. (2018). Consumption with liquidity constraints: An analytical characterization. *Economics Letters, 167*, 40–42.
27. Hou, B., Liu, H., & Wang, S. X. (2020). Returns to military service in off-farm wage employment: Evidence from rural China. *China Economic Review, 59*, 101361.
28. Huang, C. (2010). Internet use and psychological well-being: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking, 13*(3), 241–249.
29. Huang, Y., Wang, X., & Wang, X. (2020). Mobile Payment in China: Practice and Its Effects. *Asian Economic Papers, 19*(3), 1–18.
30. Jack, W., & Suri, T. (2014). Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *American Economic Review, 104*(1), 183–223.
31. Kekäläinen, T., Wilska, T. A., & Kokko, K. (2017). Leisure consumption and well-being among older adults: Does age or life situation matter? *Applied Research in Quality of Life, 12*(3), 671–691.
32. Knight, J., & Gunatilaka, R. (2010). Great expectations? The subjective well-being of rural-urban migrants in China. *World Development, 38*(1), 113–124.
33. Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukophadhyay, T., & Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological well-being? *American Psychologist, 53*(9), 1017.
34. Kwak, Y., Kim, J., Kim, Y., & Choi, J. (2021). Understanding digital consumers' well-being in Asia: The moderating roles of digital natives and privacy concerns. *Journal of Consumer Affairs, 55*(4), 1274–1291.
35. Ky, S., Rugemintwari, C., & Sauviat, A. (2018). Does mobile money affect saving behaviour? Evidence from a developing country. *Journal of African Economies, 27*(3), 285–320.
36. Lai, J. T., Yan, I. K. M., Yi, X. J., & Zhang, H. (2020). Digital financial inclusion and consumption smoothing in China. *China & World Economy, 28*(1), 64–93.
37. Li, J. (2016). Why economic growth did not translate into increased happiness: Preliminary results of a multilevel modeling of happiness in China. *Social Indicators Research, 128*(1), 241–263.
38. Li, J., Wu, Y., & Xiao, J. J. (2020). The impact of digital finance on household consumption: Evidence from China. *Economic Modelling, 86*, 317–326.
39. Liébana-Cabanillas, F., Sánchez-Fernández, J., & Muñoz-Leiva, F. (2014). Antecedents of the adoption of the new mobile payment systems: The moderating effect of age. *Computers in Human Behavior, 35*, 464–478.
40. Liu, Q., Zhang, F., Mao, M., Xue, B., & Lin, Z. (2018). An empirical study on factors affecting continuance intention of using Yu'e Bao. *Tehnički Vjesnik, 25*(5), 1414–1420.
41. Liu, T., Pan, B., & Yin, Z. (2020). Pandemic, mobile payment, and household consumption: Micro-evidence from China. *Emerging Markets Finance and Trade, 56*(10), 2378–2389.
42. Liu, Z., Ben, S., & Zhang, R. (2019). Factors affecting consumers' mobile payment behavior: A meta-analysis. *Electronic Commerce Research, 19*(3), 575–601.
43. Liu, Z., Zhong, X., Zhang, T., & Li, W. (2020). Household debt and happiness: Evidence from the China Household Finance Survey. *Applied Economics Letters, 27*(3), 199–205.
44. Lo, H. Y., & Harvey, N. (2011). Shopping without pain: Compulsive buying and the effects of credit card availability in Europe and the Far East. *Journal of Economic Psychology, 32*(1), 79–92.
45. Lu, H., & Kandilov, I. T. (2021). Does mobile internet use affect the subjective well-being of older Chinese adults? An instrumental variable quantile analysis. *Journal of Happiness Studies, 22*(7), 3137–3156.

46. Lu, Y., Bao, K., & Zhuang, M. (2020). "Identity premium" on subjective well-being: Evidence from the communist party membership in China. *Social Science Journal*. https://doi.org/10.1080/03623319.2020.1727247
47. Ma, W., Grafton, R. Q., & Renwick, A. (2020). Smartphone use and income growth in rural China: Empirical results and policy implications. *Electronic Commerce Research, 20*(4), 713–736.
48. Mallat, N. (2007). Exploring consumer adoption of mobile payments - A qualitative study. *Journal of Strategic Information Systems, 16*(4), 413–432.
49. Martikainen, E., Schmiedel, H., & Takalo, T. (2015). Convergence of European retail payments. *Journal of Banking & Finance, 50*(1), 81–91.
50. Matemba, E. D., Li, G., & Maiseli, B. J. (2018). Consumers' stickiness to mobile payment applications: An empirical study of WeChat wallet. *Journal of Database Management, 29*(3), 43–66.
51. Maurer, B. (2012). Mobile money: Communication, consumption and change in the payments space. *Journal of Development Studies, 48*(5), 589–604.
52. Modigliani, F., & Cao, S. L. (2004). The Chinese saving puzzle and the life-cycle hypothesis. *Journal of Economic Literature, 42*(1), 145–170.
53. Morgan, C., & Cotten, S. R. (2003). The relationship between Internet activities and depressive symptoms in a sample of college freshmen. *Cyberpsychology & Behavior, 6*(2), 133–142.
54. Mosteller, J., Donthu, N., & Eroglu, S. (2014). The fluent online shopping experience. *Journal of Business Research, 67*(11), 2486–2493.
55. Musa, A., Khan, H. U., & Alshare, K. A. (2015). Factors influence consumers' adoption of mobile payment devices in Qatar. *International Journal of Mobile Communications, 13*(6), 670–689.
56. N'dri, L. M., & Kakinaka, M. (2020). Financial inclusion, mobile money, and individual welfare: The case of Burkina Faso. *Telecommunications Policy, 44*(3), 101926.
57. Nie, P., Ma, W., & Sousa-Poza, A. (2021). The relationship between smartphone use and subjective well-being in rural China. *Electronic Commerce Research, 21*(4), 983–1009.
58. OECD. (2012). *Report on consumer protection in online and mobile payments*. OECD Publishing.
59. Okabayashi, H., & Hougham, G. W. (2014). Gender differences of social interactions and their effects on subjective well-being among Japanese elders. *Aging & Mental Health, 18*(1), 59–71.
60. Pan, S., Roderick, M., & He, X. (2009). Does financial intermediation development increase per capita income in rural China? *China & World Economy, 17*(4), 72–87.
61. Park, C. Y., & Mercado, J. R. (2018). Financial inclusion, poverty, and income inequality. *Singapore Economic Review, 63*(1), 185–206.
62. Pak, T. Y. (2020). Social protection for happiness? The impact of social pension reform on subjective well-being of the Korean elderly. *Journal of Policy Modeling, 42*(2), 349–366.
63. Pathak, S., & Muralidharan, E. (2021). Consequences of cross-cultural differences in perceived well-being for entrepreneurship. *Journal of Business Research, 122*, 582–596.
64. Pénard, T., & Poussing, N. (2010). Internet use and social capital: The strength of virtual ties. *Journal of Economic Issues, 44*(3), 569–595.
65. Renkow, M., Hallstrom, D. G., & Karanja, D. D. (2004). Rural infrastructure, transactions costs and market participation in Kenya. *Journal of Development Economics, 73*(1), 349–367.
66. Rose, S., & Dhandayudham, A. (2014). Towards an understanding of Internet-based problem shopping behaviour: The concept of online shopping addiction and its proposed predictors. *Journal of Behavioral Addictions, 3*(2), 83–89.
67. Sebora, T. C., Lee, S. M., & Sukasame, N. (2009). Critical success factors for e-commerce entrepreneurship: An empirical study of Thailand. *Small Business Economics, 32*(3), 303–316.
68. Shapira, N., Barak, A., & Gal, I. (2007). Promoting older adults' well-being through Internet training and use. *Aging & Mental Health, 11*(5), 477–484.
69. Solaymani, S., Sohaili, K., & Yazdinejad, E. A. (2012). Adoption and use of e-commerce in SMEs. *Electronic Commerce Research, 12*(3), 249–263.
70. Sotiropoulos, V., & D'Astous, A. (2012). Social networks and credit card overspending among young adult consumers. *Journal of Consumer Affairs, 46*(3), 457–484.
71. Su, L., Peng, Y., Kong, R., & Chen, Q. (2021). Impact of E-Commerce Adoption on Farmers' Participation in the Digital Financial Market: Evidence from Rural China. *Journal of Theoretical and Applied Electronic Commerce Research, 16*(5), 1434–1457.
72. Suri, T. (2017). Mobile money. *Annual Review of Economics, 9*(1), 497–520.
73. Wang, X. (2020). Mobile payment and informal business: Evidence from China's household panel data. *China & World Economy, 28*(3), 90–115.

74. Wang, X., & Wen, Y. (2012). Housing prices and the high Chinese saving rate puzzle. *China Economic Review, 23*(2), 265–283.
75. Wei, S., & Zhang, X. (2011). The competitive saving motive: Evidence from rising sex ratios and savings rates in China. *Journal of Political Economy, 119*(3), 511–564.
76. Yang, L., Lu, H., Wang, S., & Li, M. (2021). Mobile Internet Use and Multidimensional Poverty: Evidence from A Household Survey in Rural China. *Social Indicators Research, 158*(3), 1065–1086.
77. Yang, X., Li, Y., & Liao, Q. (2016). Exploring continued use of mobile shopping channel in China: The effects of active coping and its antecedents. *Electronic Commerce Research, 16*(2), 245–267.
78. Yeung, G., He, C., & Zhang, P. (2017). Rural banking in China: Geographically accessible but still financially excluded? *Regional Studies, 51*(2), 297–312.
79. Yin, Z., Xue, G., Guo, P., & Tao, W. (2019). What drives entrepreneurship in digital economy? Evidence from China. *Economic Modelling, 82*, 66–73.
80. Yu, L., Cao, X., Liu, Z., Gong, M., Adeel, L., & Cheung, C. (2018). Understanding mobile payment users' continuance intention: A trust transfer perspective. *Internet Research, 28*(2), 456–476.
81. Zhan, G., & Zhou, Z. (2018). Mobile internet and consumer happiness: The role of risk. *Internet Research, 28*(3), 785–803.
82. Zhao, C., & Guo, J. (2021). Are Veterans Happy? Long-term military service and the life satisfaction of elderly individuals in China. *Journal of Happiness Studies*. https://doi.org/10.1007/s10902-021-00410-4
83. Zhao, C., Wu, Y., & Guo, J. (2022). Mobile payment and Chinese rural household consumption. *China Economic Review, 71*, 101719.
84. Zhao, C., & Li, X. (2022). Living under the shadow: Adverse childhood experiences and entrepreneurial behaviors in Chinese adults. *Journal of Business Research, 138*, 239–255.
85. Zhao, D., Jiang, J., & Yin, Z. (2020). Can entrepreneurship bring happiness? Evidence from China. *Economic Modelling, 91*, 679–686.
86. Zheng, H., & Ma, W. (2021). Click it and buy happiness: Does online shopping improve subjective well-being of rural residents in China? *Applied Economics, 53*(36), 4192–4206.
87. Zheng, H., & Ma, W. (2022). Scan the QR code of happiness: Can mobile payment adoption make people happier? *Applied Research in Quality of Life*. https://doi.org/10.1007/s11482-022-10036-0
88. Zhu, Z., Ma, W., Leng, C., & Nie, P. (2021). The relationship between happiness and consumption expenditure: Evidence from rural China. *Applied Research in Quality of Life, 16*(4), 1587–1611.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Chunkai Zhao[1] · Xing Li[2] · Jianfeng Yan[3]

Chunkai Zhao
ckz_zhao@163.com

[1] College of Economics and Management, South China Agricultural University, No.483, Wushan, Tianhe District, Guangzhou, Guangdong Province, China

[2] College of Economics and Management, Southwest University, 2 Tiansheng Road, Beibei, Chongqing, China

[3] Institute of Finance and Economics Research, Shanghai University of Finance and Economics, 777 Guoding Road, Yangpu District, Shanghai, China

通讯作者论文

Journal of Chinese Political Science (2024) 29:649–672
https://doi.org/10.1007/s11366-023-09864-0

RESEARCH ARTICLE

Official Turnover and Sustainable Development in China

Xing Li[1] · Chunkai Zhao[2] · Jianhua Cao[3]

Accepted: 28 July 2023 / Published online: 24 August 2023
© Journal of Chinese Political Science/Association of Chinese Political Studies 2023

Abstract

Achieving sustainable development is an important goal for China's economic transformation to high-quality development. Using the city-year panel data and constructing a green total factor productivity (GTFP) indicator, we implement the idea of an "event study" to explore the impact of official turnover on sustainable development in China. The results, supported by a series of robustness tests, show that official turnover significantly contributes to sustainable development. Furthermore, official turnover mainly promotes sustainable development through the deterrent effect and reducing government construction investment; the turnover of highly educated, local, promoted, and experienced officials could better promote sustainable development. These findings shed new light on the determinants of sustainable development and suggest that local official turnover plays an important role in China's high-quality development.

Keywords Official turnover · Sustainable development · Deterrent effect · Government construction investment · China

✉ Chunkai Zhao
ckz_zhao@163.com

Xing Li
xingli277@126.com

Jianhua Cao
cao.jianhua@mail.shufe.edu.cn

[1] College of Economics and Management, Southwest University, No. 2, Tiansheng Road, Beibei District, Chongqing, China

[2] College of Economics and Management, South China Agricultural University, No.483, Wushan Street, Tianhe District, Guangzhou, Guangdong Province, China

[3] Institute of Finance and Economics, Shanghai University of Finance and Economics, 777 Guoding Road, Yangpu District, Shanghai, China

Introduction

China's extensive development model has achieved relatively rapid economic growth; however, it has also exacted a huge price regarding resources and the ecological environment, obstructing the country from achieving sustainable development. According to the Global Environmental Performance Index, jointly published by Yale University, Columbia University, and the World Economic Forum, China ranked 120th of 180 countries and regions in 2018 and even ranked fourth from the bottom in air quality [38]. Worse, China's inefficient production mode leads to massive resource consumption and carbon emissions [35]. The deterioration of the ecological environment threatens people's health and brings a series of socio-economic consequences [53, 75].

China has proposed carbon peaking and carbon neutrality goals [63, 76, 77], constructed an ecological civilization, and transformed its economy to sustainable development models, all of which have attracted attention from policy and academic circles [48]. The economy's sustainable development does not depend solely on economic growth or environmental protection, but on reducing energy consumption, air pollutants, and carbon emissions while achieving economic growth [66]. This study uses green total factor productivity (GTFP) to characterize the degree of sustainable development of China's economy. On the one hand, GTFP comprises three aspects: input factors (such as labor, capital, and energy), desired output (such as economic output), and undesired output (such as air pollutants and carbon emissions). It systematically and comprehensively measures the coordination between resource consumption, economic output, and environmental pollution to represent the economy's degree of sustainable development [32, 41, 46, 67]. On the other hand, with the continuous improvement of the measurement method, GTFP is increasingly able to accurately estimate whether the economic development mode can simultaneously achieve energy conservation, emission reduction, and economic growth; therefore, GTFP is a fairly appropriate indicator of sustainable development [18, 34, 37].

One reason that "high economic growth" and "high environmental pollution" coexist in China is the mismatch between the aggressive central government policies and local political evaluation systems [44]. Under the political system of administrative centralization and fiscal decentralization, the performance appraisal of local officials is highly dependent on gross domestic product (GDP) growth [21, 36, 60, 65]. To meet the assessment requirements of GDP and achieve promotion, local government officials often invest significant resources in industrial enterprises with high energy consumption and pollution [55, 60]. Such officials may even condone environmental pollution by reducing regulation [5, 30] to quickly obtain significant economic output in the short term, increasing air pollutants and carbon emissions and leading to overcapacity [12, 51]. In this situation, polluting companies can maintain their sloppy operations to save costs and discharge pollutants [12, 15].

Official turnover will likely break the tacit understanding between government officials and high-polluting enterprises [53]. Official turnover means that the central government's evaluation of specific officials during specific period has ended, and they do not need to continue to increase investment and construction

to maintain high GDP growth. For enterprises with high energy consumption and high pollution, official turnover also increases the policy's uncertainty, which deters these enterprises from emitting carbon dioxide and air pollutants without restraint, at least in the short term [12, 68]. In other words, official turnover could reduce input factors or undesired output for and high-polluting enterprises, thus promoting sustainable development.

Given the above, we specifically discuss whether official turnover affects sustainable development. Based on information on the mayors and party secretaries of 277 prefecture-level and above cities in China from 2004 to 2016, we use the method of Afsharian and Ahn [1] to correct the efficiency frontier of GTFP and calculate a more accurate efficiency calibration value, and use the idea of "event study" to explore the causal relationship between official turnover and sustainable development. We find that official turnover significantly contributes to GTFP growth, indicating that official turnover positively promotes sustainable development in China, which is more pronounced for a municipal secretary change. Regarding official characteristics, changes in officials with higher education, local *hukou*, and older age are more likely to promote sustainable development. The mechanism analysis suggests that the deterrent effect on enterprise and the reduction of government construction investment primarily drives the effect of official turnover on sustainable development.

This paper contributes to the existing literature in the following ways. First, our study enriches the research field on environmental politics. The existing literature focuses on how officials' political cycles affect air pollution or carbon emissions [31, 42, 56], while other studies examine the impact of official turnover on economic growth [36, 40]. Based on the special perspective of sustainable development, we systematically study the impact of official turnover on the degree of coordination between the economy, resources, and environment.

Second, we measure the sustainable development level of China's economy to more accurately identify the causal relationship between official turnover and sustainable development. The traditional data envelopment analysis (DEA) model generally has some problems, such as incorrectly setting the frontier, which will lead to inaccurate measurement of efficiency [1, 50]. We improve the overall technology frontier DEA model. Under the condition of effectively solving the problems of technology frontier setting bias and efficiency measurement bias, we have conducted a more accurate calculation of the GTFP growth at the city level in China.

Third, we conduct an in-depth discussion on how officials' turnover affects sustainable development and further clarify the heterogeneity from the perspective of individual characteristics of officials. Therefore, these results provide ideas on advancing sustainable development and help deepen the theoretical cognition and empirical foundation of "better-giving play to the role of the government" in economic development [20]. Furthermore, given that political uncertainty involves official turnover, we may also provide some clues for this research field [2, 11, 29, 69].

The rest of this paper is arranged as follows. Section 2 mainly elaborates on the background, Section 3 outlines the theoretical hypothesis, and Section 4 introduces a more accurate measurement method for GTFP. Section 5 explains our data,

variables, and econometric models, Section 6 presents the main results, and Section 7 concludes this paper.

Official Turnover in China

Official turnover originates in the cadre exchange system that began in the early years of the People's Republic of China [12]. The cadre exchange aims to avoid the bureaucracy and corruption resulting from officials serving in the same place for a long time [25]. Since the reform and opening up in 1978, the Chinese government has been improving the system of turnover of local officials.

In 1980, China's central government began addressing the issue of tenure for senior officials; in 1982, the central government enacted a constitution stipulating that the retirement age for provincial and ministerial leaders would be 65 and 60, respectively, and for bureau-level leaders would be 60. Additionally, the 12th National Congress adopted a new party constitution, stipulating that the positions of leading cadres at all levels can be changed or removed and that each term of office is five years, with no more than two consecutive terms of office. These policies have led to the initial establishment of a system of official turnover in China.

The "Decision on Implementing the System for the Exchange of Leading Cadres in Party and State Organs" was implemented in 1990, further improving China's official turnover system and establishing the institutionalization of China's official turnover mechanism. In 1994, the Fourth Plenary Session of the 14th Central Committee further emphasized the need to conscientiously implement the exchange system of leading cadres, increase the turnover of cadres at the provincial and ministerial levels, and continue to promote turnover of cadres at the local, municipal, and county levels [57]. These measures further limited the time that leading cadres could serve in one place and gradually normalized the official turnover in China.

Since the 21st century, the "Regulations on the Selection and Appointment of Party and Government Leading Cadres" and the "Interim Provisions on the Term of Office of Leading Cadres of the Party and Government" have been promulgated to continue to institutionalize and standardize China's official turnover system [54]. They stipulate that party and government leading cadres who have served two consecutive terms in the same post may not be re-elected. Furthermore, government officials who have served in the same post for more than 10 years must be exchanged. More importantly, with the development of the construction of ecological civilization and the emphasis on sustainable development, the Chinese government issued the Decision on Implementing the Scientific Outlook on Development and Strengthening Environmental Protection in 2005. This issuance included environmental protection as an important part of the assessment of the leadership and leading cadres and took the assessment as a basis for the cadres' selection, appointment, rewards, and punishment. In 2013, the Chinese government issued the Notice on Improving the Performance Assessment of Local Party and Government Leading Groups and Leading Cadres, strengthening the weight of environmental protection and other indicators. The official turnover system in China has become an important way to govern officials, curb corruption, and improve the economic growth and sustainable development [21, 28].

Theoretical Hypotheses

The Effect of Official Turnover on Sustainable Development

In China, political promotion is the most important career incentive for local government officials [10, 36]. The development of the local economy is the main performance evaluation standard; thus, to quickly meet the evaluation requirements in the short term, local officials tend to concentrate resources and develop the economy in various ways, even at the expense of the ecological environment [36]. With GDP as the main promotion assessment target, local government officials invest more natural resources and government investment in city constructions and traditional industry development, resulting in significant carbon emissions and serious air pollution [33]. In other words, extending officials' terms in office increases air pollution [26]. Furthermore, to develop the local economy faster, local officials even form a government–enterprise association with local industrial enterprises and acquiesce in their pollution emission behaviors, generating substantial economic output from energy-consuming and high-polluting enterprises [24]. To pursue the goal of economic growth, some officials even reduced environmental regulations to attract more enterprises to increase local economic output [56]. The local economy can achieve rapid growth under this political promotion competition system; however, this growth comes at the cost of significant natural resource investments, serious air pollution, and carbon emissions [13]. Clearly, this economic growth model is inefficient and has seriously affected the local ecological environment, which is unsustainable.

Nonetheless, official turnover may inhibit this development model. On the one hand, the official turnover means that the central government's assessment of local officials has ended, and local government officials do not need to continue to invest many resources in inefficient production [76, 77]. On the other hand, the turnover of local officials also means that local enterprises with high energy consumption and high pollution would no longer be able to discharge pollutants at will [78]. In general, official turnover could reduce the input of resources, reduce the inefficient production mode, and reduce the emissions of carbon dioxide and air pollutants, thus achieving sustainable development. Therefore, we propose our first hypothesis:

Hypothesis 1 Official turnover likely promotes sustainable development.

Mechanism Analysis

Official Turnover and Deterrent Effect

Official turnover could disrupt the relationship network formed by officials and local industrial enterprises due to their long-term employment, reducing the possibility of collusion between the government and enterprises. Thus, official turnover can deter

industrial enterprises with high energy consumption and high pollution, reducing the inefficient production and emission behavior of local industrial enterprises [56].

In China, the regular term of office of the local official is five years [36]. Because technological innovation takes a long time [23, 72, 74], officials often form political relations with local energy-consuming and high-polluting enterprises to obtain a large amount of economic output within a short term and meet the central government's assessment requirements. This practice condones these enterprises' energy-consuming and high-polluting behaviors, ensuring significant output for the local economy [42]. With the shelter of government officials, these energy-consuming and high-polluting enterprises can maintain inefficient production to save production costs [27].

Official turnover could directly break the link between the government and enterprises, making enterprises lose the shelter of government departments [80]. The production mode of high energy consumption, high pollution, and low efficiency is no longer allowed [31]. More importantly, the politically sensitive period formed by official turnover would also deter the local government–enterprise association from engaging in potentially damaging practices [12]. For example, Julio and Yook [29] found that firms reduce investment expenditures by an average of 4.8% during election years. Furthermore, environmental protection and other relevant departments strengthen the supervision during official turnover, making the production behavior of industrial enterprises with high energy consumption and pollution converge [16]. Therefore, official turnover may deter enterprises with high energy consumption and high pollution by reducing their inefficient, high energy consumption and high pollution behaviors, which promots sustainable development. Taken together, we propose our second hypothesis:

Hypothesis 2 Official turnover may contribute to sustainable development through the deterrent effect.

Official Turnover and Government Construction Investment

Local government officials in China face the dual incentives of fiscal decentralization and vertical political concentration [36, 64]. The central government usually evaluates local officials according to economic growth; therefore, local government officials have strong incentives to compete around GDP growth and promote the transformation of the local system to a development-oriented government governance model [61, 62]. The public expenditure theory indicates that competition among governments can improve the supply of public goods and improve economic efficiency [81]. China's local government significantly influences local economic development [36]. In addition to the influence of local government officials on the delivery of important economic resources, such as land and credit, local government officials can also have certain discretion in the process of approval, taxation, and punishment [43, 56]. To promote the local economy's rapid growth in a short term of office, local government officials tend to allocate more financial expenditure on

capital construction investment with quick and obvious growth effects, especially concerning the construction of public goods with high visibility [26, 82].

Under the promotion championship system, local officials attach more importance to infrastructure construction investment than to environmental pollution control and technological progress, which also lead to overcapacity, low production efficiency, and environmental pollution [58]. The massive consumption of resources and the destruction of the ecological environment have also hindered the sustainable development of China's economy [78]. Official turnover could somewhat alleviate overcapacity, resource consumption, and environmental pollution caused by promotion to the championship. Some studies suggest that government construction and state-owned enterprise investments have significantly decreased during the official turnover period [17, 59]. Furthermore, the problems of inefficient production and overcapacity have also been appropriately alleviated [12]. Therefore, we develop the third hypothesis:

Hypothesis 3 Official turnover is likely to promote sustainable development by reducing government construction investment.

Measurements of GTFP

Redefine Production Feasible Set

A good Malmquist Productivity Index must satisfy four basic properties: transferability, uniqueness of productivity changes, technical regression, and the existence of no feasible solutions [19, 47]; however, some early methods struggle to meet these conditions simultaneously [4, 19, 52]. The first widely recognized and applied method was the global method constructed by Pastor and Lovell [50]. Based on Pastor and Lovell [50] and Afsharian and Ahn [1], we used the overall method to update the production feasible set of samples, making the production frontier composed of samples from different periods more accurate. Figure 1 shows the differences before and after the update of the production frontier.

Subscripts 1 and 2 represent two successive periods. The global method assumes that the production frontier of the sample is $O_1A_1A_2B_2B_1Z_1$, while the overall method assumes that the production frontier of the sample is $O_1A_1MA_2B_2NB_1Z_1$, eliminating the two production infeasible parts of A_1A_2M and B_1B_2N. In the global method, the sample point U maps to the production frontier B_1B_2, which is nonexistent. In the overall method, the sample point U is mapped to the real production frontier B_2N. Therefore, the overall method can provide more accurate calculations than the global method [1].

Calculation and Decomposition of GTFP

After using the overall method to determine the sample's production feasible set, we refer to Zhou et al. [79], Zhang and Choi [73], and Fan et al. [18] and calculate

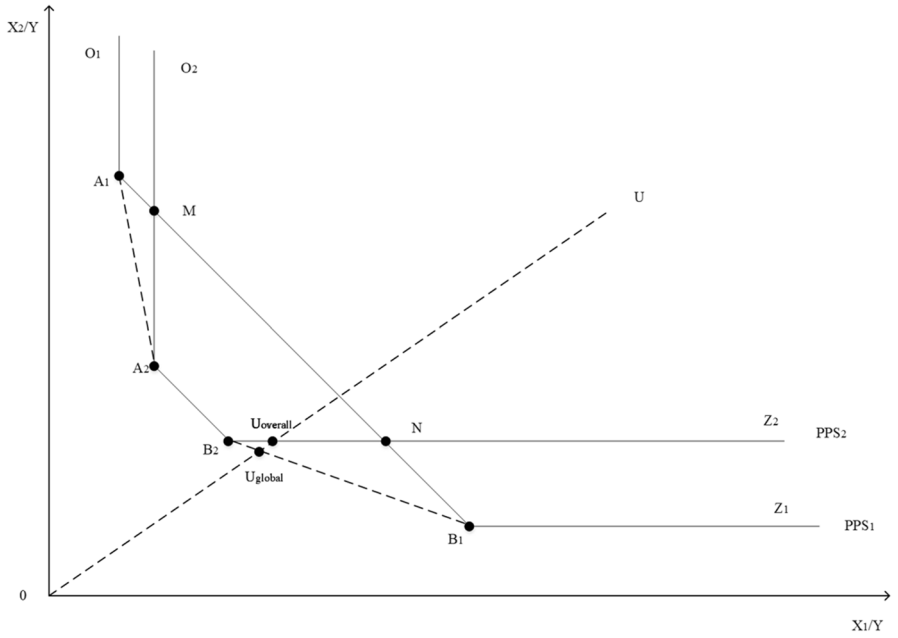


Fig. 1 Comparison of production frontiers calculated by Global and Overall methods

the GTFP of the sample using the non-angular and non-radial directional distance function (DDF) methods. According to the Malmquist–Luenberger productivity index [19, 47], and based on the overall method and non-angular, non-radial DDF, the GTFP growth between period t and period t + 1 can be defined as

$$GTFP = \theta(\pi)^t - \theta(\pi)^{t+1} \tag{1}$$

where θ represents non efficiency value. π represents a set of input-output relationships composed of (x, y, b; g), where x, y, and b represent the vector of input, desired output, and undesired output, respectively. g is the direction vector. GTFP improved if $GTFP > 0$ and deteriorated otherwise. According to Fan et al. [18], the growth of GTFP can be decomposed as follows:

$$EC = \theta^t(\pi)^t - \theta^{t+1}(\pi)^{t+1} \tag{2}$$

$$TC = \left[\theta(\pi)^t - \theta^t(\pi)^t\right] - \left[\theta(\pi)^{t+1} - \theta^{t+1}(\pi)^{t+1}\right] \tag{3}$$

where EC is an efficiency change index, meaning that each observation unit meets the best production frontier from t period to t + 1 period. EC > 0, EC < 0, and EC = 0 means that the distance between the observed unit and the best production frontier is closer, farther away, and remains unchanged, respectively.

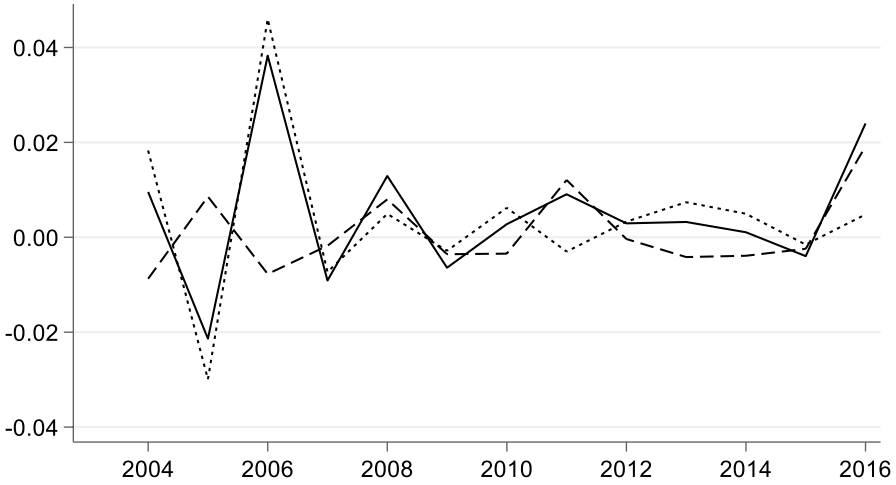


Fig. 2 General trends of GTFP, EC and TC. Note: The y-axis represents GTFP growth, and the x-axis represents the year. The solid line represents GTFP, while the long dashed, dotted line represents TC and EC, respectively

TC is an indicator of technological change, which means that the boundary of the PPS moves from period t to period t + 1, and measures the change in the technology of the decision-making unit. TC > 0, TC < 0, and TC = 0 represent the expansion, shrinkage, and unchanged of the best production frontier toward "less input, more desired output, and less undesired output."

Generally speaking, in calculating GTFP, the more input elements selected, the more accurate the calculation of GTFP can be [18]; however, due to data limitations in the actual calculation process, the existing literature cannot select all input elements, especially at the city level [71]. Therefore, most literature selects labor, capital, and energy to characterize the input factors [70]; the labor force represents the labor factor. Due to the serious lack of energy data at the city level, we select the most important secondary energy (electricity) to represent energy input. Output factors include GDP, and undesired output factors include industrial wastewater discharge, solid waste, $PM_{2.5}$ concentrations, and carbon dioxide concentrations. The $PM_{2.5}$ concentration data comes from the satellite light data measured by Columbia University using National Aeronautics and Space Administration (NASA) satellites, and the carbon dioxide concentration comes from China Emission Accounts and Datasets using satellite light data to calculate the carbon dioxide concentrations [39]. Other input–output data are from the China City Statistical Yearbook. Through the calculation of formula (1) and the decomposition of formulas (2) and (3), Fig. 2 shows the changing trend of GTFP, TC, and EC in China from 2004 to 2016.

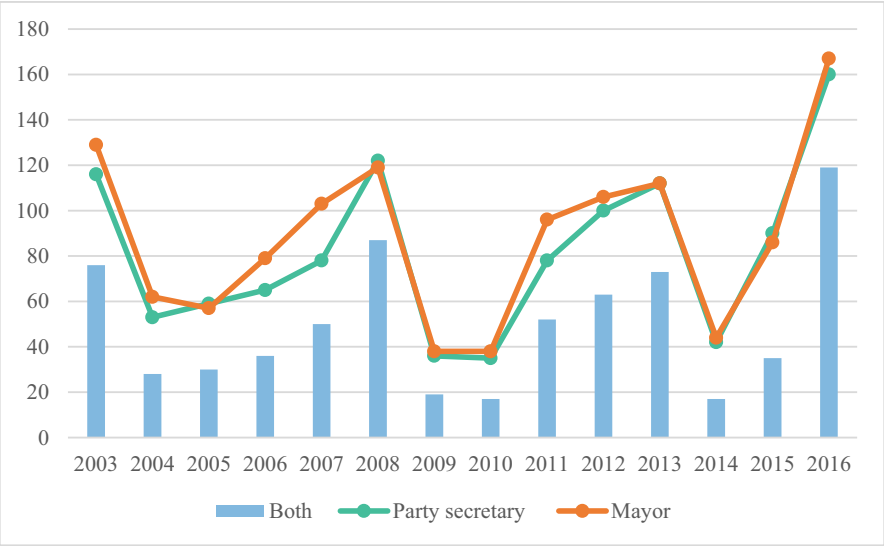


Fig. 3 Sample distribution of official turnover

Data and Methodology

Data and Variables

The data used in this paper are the annual data of 277 prefecture-level and above cities in China from 2004 to 2016. The explained variable is the GTFP calculated at the city level. The core explanatory variable is the official turnover (*Turnover*); we set *Turnover* = 1 if the mayor or municipal party secretary turnover, and 0 otherwise. Since we study at the annual level, and the impact of official turnover may have time lag effects, we treat turnover after June as turnover in the following year [35]. Information on official turnover comes mainly from Xinhua.com, People's Daily Online, and the government websites of each city. In addition to the dates of specific turnover in party secretaries and mayors, we compiled the individual characteristics of these officials, such as age, educational background, places of origin, and reasons for turnover, based on their CVs and official media messages.

Figure 3 shows the turnover of party secretaries and mayors. During the sample period, 1,146 party secretaries and 1,236 mayors changed at the city level in China. The number of party secrets and mayoral changes each year is roughly the same, and the trend of changes is at its peak every 5 years, which is also consistent with China's 5-year tenure system.

When selecting the control variables, we mainly controlled the weather variables to avoid the error caused by the "bad control problem" on the estimated results [3, 8] and to accurately study the impact of official turnover on GTFP. On the one hand, weather variables could affect the concentration of pollutants in the air [61, 62]. On the other hand, extreme weather and other conditions would also

Table 1 Descriptive statistics

Variables	Description	Obs	Mean	S.D.	Min	Max
GTFP		3,509	0.0046	0.0561	-0.4376	0.4282
Turnover	0/1	3,509	0.4209	0.4982	0	1
Party secretary turnover	0/1	3,509	0.2870	0.4405	0	1
Mayor turnover	0/1	3,509	0.2092	0.4064	0	1
AT	Celsius degree	3,509	15.070	5.1692	-1.2000	24.700
AH	%	3,509	68.566	9.6526	40.000	91.000
PR	Millimetre	3,509	1906.9	28910	41.800	2959.3
LI	Hour	3,509	1936.5	510.46	598.40	3376.1

affect the production efficiency of industrial enterprises [9, 35], thus affecting GTFP. More importantly, weather variables would not be affected by official turnover. The weather variables we selected include average temperature (*AT*), average humidity (*AH*), precipitation (*PR*), and light intensity (*LI*). The weather variables are from the National Weather Bureau monitoring stations. Table 1 presents the descriptive statistics of these variables.

Econometric Model

We used the idea of "event study" to explore the causal relationship between official turnover and GTFP:

$$GTFP_{it} = \alpha_0 + \alpha_1 Turnover_{it} + \alpha_2 X_{it} + \alpha_3 \delta_i + \alpha_4 \varphi_t + \alpha_5 \gamma_{pt} + \varepsilon_{it} \quad (4)$$

where subscript i represents the city and t is the year. $GTFP_{it}$ is the proxy of sustainable development of each city, calculated by the input, output, and undesired output. $Turnover_{it}$ denotes a dummy variable for official turnover, indicating whether the party secretary or the mayor has turnover. $\alpha_1$is our key coefficient, which is used to reflect the average impact of official turnover on sustainable development in China. X_{it} refers to control variables. δ_i is the city-fixed effect (FE), φ_t is the year-FE, and γ_{pt} is the interaction term of province-FE and year-FE. ε_{it} represents the random error term; we cluster standard errors to the city level.

Results

Baseline Results: The Effect of Official Turnover on Sustainable Development in China

Table 2 shows the impact of official turnover on sustainable development in China. Column (1) controls the city, year, and province-year-FEs. Column (2) further controls the *AT* and *AH* based on column (1), while column (3) controls

Table 2 Baseline results

	GTFP			*TC*	*EC*
	(1)	(2)	(3)	(4)	(5)
Turnover	0.0051***	0.0054***	0.0058***	0.0022	0.0036**
	(0.0018)	(0.0018)	(0.0018)	(0.0020)	(0.0017)
AT		-0.0105***	-0.1805	-0.2490	0.0685
		(0.0037)	(0.3997)	(0.1638)	(0.3917)
AH		0.0171***	0.0005	0.0007	-0.0002
		(0.0043)	(0.0016)	(0.0007)	(0.0016)
PR			0.0050**	-0.0010	0.0060***
			(0.0023)	(0.0011)	(0.0022)
LI			-0.0006	0.0073*	-0.0079*
			(0.0081)	(0.0042)	(0.0043)
City FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes
Constant	0.0026***	-0.0356*	-0.0356*	0.0016	-0.0372*
	(0.0007)	(0.0190)	(0.0190)	(0.0081)	(0.0190)
Observations	3509	3509	3509	3509	3509
Adjusted R-squared	0.0023	0.0028	0.0028	0.0010	0.0013

The values in parentheses are standard errors, and we cluster standard errors to the city level

*** p<0.01, ** p<0.05, * p<0.1

AT, *AH*, *PR,* and *LI*. The results in column (3) of Table 2 indicate that the turnover of officials could promote the growth of GTFP at the 1% significance level. More specifically, we find that offical turnover increases the GTTP by 0.0058 points. Translated into the standard deviation form, this effect is equivalent to 0.0515 standard deviations. Thus these estimates demonstrate the potential of official turnover in promoting sustainable development.

To study the impact of official turnover on sustainable development in more detail, we divide the growth sources of GTFP into TC and EC and report the results in columns (4) and (5) of Table 2, respectively. The promotion effect of official turnover on the growth of GTFP is more from efficiency change than technological change. In other words, the growth of GTFP during the official turnover period is not due to the improvement of production technology; however, it may be due to the reduction of enterprise production activities, the reduction of pollutant emissions, and the strengthening of enterprise management. In sum, the baseline results support Hypothesis 1, confirming the positive role of official turnover in promoting sustainable development in China.

Table 3 Robustness checks I

	Replace the GTFP indicator		Replace turnover indicator	Greening process	Greening process and regional differences
	(1)	(2)	(3)	(4)	(5)
Turnover	0.0115***	0.0223***	0.0339***	0.0158***	0.0110***
	(0.0016)	(0.0024)	(0.0058)	(0.061)	(0.0037)
$Year_{2007}$				0.0092***	
				(0.0007)	
Environmental					0.0070***
					(0.0004)
Control variables	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	No	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes
Constant	0.0019	0.0159	-0.0315***	-0.0183***	-0.0078***
	(0.0015)	(0.0111)	(0.0039)	(0.0021)	(0.0028)
Observations	3509	3509	3509	3509	3509
Adjusted R-squared	0.0028	0.0033	0.0026	0.0031	0.0019

The values in parentheses are standard errors, and we cluster standard errors to the city level

*** p<0.01, ** p<0.05, * p<0.1

Robustness Checks

Replacing Key Variables and Regression Models

First, we change the source of $PM_{2.5}$ in the calculation of GTFP. The $PM_{2.5}$ data used in estimating the growth of GTFP is the three-year average estimation of $PM_{2.5}$ by Columbia University through NASA satellites; therefore, the $PM_{2.5}$ data is relatively smooth, which may bias the calculation results of GTFP. We use the $PM_{2.5}$ data calculated by Dalhousie University in Canada to test the robustness of the results [7]; Dalhousie University also uses NASA satellites to estimate the average annual $PM_{2.5}$, which improves the accuracy of $PM_{2.5}$ concentration estimation and can more accurately estimate the growth of GTFP. Column (1) of Table 3 presents the estimated results of replacing GTFP growth measurement indicators. We find that official turnover can still promote sustainable development at the 1% significant level, which means the previous results are credible.

Second, we add SO_2 to GTFP as an undesired output. The central government's priorities for controlling different pollutants vary in different sample periods, such as focusing on SO_2 pollution before 2013 and $PM_{2.5}$ pollution after 2013. To comprehensively reflect the change in GTFP and more accurately identify the impact of official turnover on sustainable development, under the condition of data limitations, we treat SO_2 emissions in the Statistical Yearbook of Chinese Cities as undesired output and add them to the calculation of GTFP. The results in column (2) of

Table 3 show that when SO_2 is taken considered, official turnover can still promote GTFP growth to a greater extent.

Third, we convert the $Turnover_{it}$ into an ordinal variable for further robustness checks. In the baseline regressions, we use a dummy variable to indicate whether there is any official turnover. If the secretary of the municipal party committee and the mayor of city i turnover in the same year t, the turnover of officials may have a greater role in promoting the growth of the local GTFP. We regard $Turnover_{it}$ as an ordered variable used to indicate the number of officials who turnover in city i and year t. The results in column (3) of Table 3 show that the turnover of officials could significantly promote the growth of GTFP at the level of 1%, and the promotion effect is roughly 6 times the baseline regression result. This estimate indicates that the simultaneous turnover of officials could promote sustainable development to a greater extent.

Greener Appraisal for Officials

China has been working in recent years to make official appraisal standards green; thus, we eliminate the influence of the greening process in promoting and assessing officials. The Chinese government proposed the concept of ecological civilization around 2007 and strengthened the greening process of officials' assessments, which may lead to errors in the estimation results. We add the time trend term $Year_{2007}$ to control the impact of these policies [49]. We assigned a value of 1 after 2007 to indicate that the government began implementing the greening process of officials' promotion assessment, and a value of 0 before 2007 indicates that the greening process has not been implemented[1]. The results in column (4) of Table 3 indicate that after controlling the greening process that began around 2007, the baseline results remain unchanged that official turnover could still significantly promote the growth of GTFP.

Moreover, we rule out the influence of the greening process and regional heterogeneity; the differences in environmental protection intensity and officials' evaluation indicators in different regions may bias the estimated results. To eliminate such influence and get a purer effect of official turnover on sustainable development, we sorted the word frequencies related to green development and environmental protection in the "Government Work Report" of each city during the sample period [6] to control the importance that each city attached to green development each year. The results in column (5) of Table 3 show that official turnover could still significantly promote sustainable development.

Other Confounding Factors

We conduct robustness checks by eliminating some confounding factors in Table 4. First, we conduct a counterfactual test of changing the time by advancing the turnover time of party secretaries and mayors by one year. Column (1) shows

[1] In the case of controlling for the year-FE, the time trend term cannot be estimated due to the full collinearity. Therefore, here we do not add the year-FE.

Table 4 Robustness checks II

	GTFP			
	(1)	(2)	(3)	(4)
Turnover	-0.0048	0.0054***	0.0039***	0.0225***
	(0.0112)	(0.0012)	(0.0010)	(0.0017)
Control variables	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes
Constant	-0.0145***	0.0170***	0.0043***	-0.0016
	(0.0027)	(0.0017)	(0.0004)	(0.0033)
Observations	3218	3434	3361	3410
Adjusted R-squared	0.1436	0.2861	0.1710	0.2934

The values in parentheses are standard errors, and we cluster standard errors to the city level

*** p<0.01, ** p<0.05, * p<0.1

the regression results after the turnover time is advanced by one year. Clearly, after artificially advancing the turnover time by one year, the effect of official turnover on sustainable development has no continued promotion; the significance also becomes insignificant at the 10% significance level. This result proves the credibility of the above conclusion.

Second, we exclude the situation where the political system may overestimate the effects. The provincial capital city is the political center of a province where a large number of political resources are concentrated, and the governments at all levels prioritize the effect of policy improvement in capital cities. Therefore, we exclude the influence of provincial capital cities and municipalities directly under the central government. The estimated result in column (2) shows that the impact of official turnover on sustainable development is still robust, confirming the reliability of the baseline results above.

Third, we exclude the overestimation that industrial relocation may cause. China's Western Development Plan was fully implemented during the sample period when more polluting enterprises and enterprises with low production efficiency in the eastern region moved to the western region. Economic development was not substantially affected in the eastern region, but the pollution level in the eastern region was relatively reduced, causing an overestimation of the effect of official turnover. Column (3) shows the estimated results after excluding the industrial cities in the east and supports the reliability of our main results.

Finally, we exclude cities where official turnover occurred too frequently. For example, in Wenzhou and Handan, from 2004 to 2016, only 4 years were not replaced by officials, and the remaining 9 years experienced a turnover of a mayor or party secretary. Under such unstable conditions, it is difficult or even impossible for industrial enterprises to form a collusion with the government. Column (4) excludes cities with the highest frequency of official turnover, and the estimation is consistent with previous results.

Table 5 Mechanism tests

	Deterrent effect		Government construction investment	
	(1)	(2)	(3)	(4)
Panel A. The impact of official turnover on mechanism variables				
Turnover	-0.0151***	0.0093***	-0.0129***	-0.0183***
	(0.0009)	(0.0015)	(0.0004)	(0.0025)
Panel B. The impact of mechanism variables on GTFP				
enterprise pollution level	-0.2397***			
	(0.0018)			
enterprise production efficiency		0.5120***		
		(0.0017)		
investment in fixed assets			-0.6231***	
			(0.0050)	
expenditure within the general budget				-0.4025***
				(0.0010)

The values in parentheses are standard errors, and we cluster standard errors to the city level. We add control variables, city FE, year FE, and province-year FE in all regressions in Panels A and B

*** p<0.01, ** p<0.05, * p<0.1

Mechanism Tests

The above analysis indicates that official turnover has a significant role in promoting sustainable development. Based on the previous theoretical analysis, official turnover would promote sustainable development through a deterrent effect and government construction investment; therefore, this paper test the mechainsms by which official turnover affect the growth of GTFP from these two perspectives. Specifically, we use the enterprise pollution level and enterprise production efficiency to express the level of deterrent effect, following Ma et al. [45]. Government construction investment is expressed as the investment in fixed assets and the expenditure within the general budget of local finance. Table 5 reports the estimation results.

We use the classic mechanism analysis model in economics [14]. First, we test the regional technological progress. Columns (1) and (2) of Panel A in Table 5 suggest that official turnover could restrain the pollutant discharge level and promote enterprises' production efficiency. In general, official turnover could promote the increase of GTFP through the deterrent effect. Second, we tested the role of government construction investment, and the estimated results are reported in columns (3) and (4) of Panel A in Table 5. The results show that official turnover can restrain government construction investment, represented by fixed assets investment and local expenditure in the general budget, thus promoting GTFP growth. Official turnover could

Table 6 Heterogeneity analysis I

	Party secretary (1)	Mayor (2)	Normal (3)	Abnormal (4)
Turnover	0.0016	0.0198***	0.0144***	0.0020
	(0.0011)	(0.0029)	(0.0019)	(0.0021)
Control variables	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes
Constant	-0.0094***	-0.0056**	0.0026	-0.0019***
	(0.0016)	(0.0023)	(0.0047)	(0.0002)
Observations	3486	3486	3325	2610
Adjusted R-squared	0.2549	0.2668	0.1331	0.1248

The values in parentheses are standard errors, and we cluster standard errors to the city level

*** p<0.01, ** p<0.05, * p<0.1

promote sustainable development by the deterrent effect and reducing government construction investment.

We also show the influence of mechanism variables on dependent variables in the Panel B of Table 5. All the mechanism variables selected can significantly affect the dependent variable, and the influence direction is also in line with expectations. For example, the results of column (1) in Panel A in Table 5 show that official turnover inhibits the enterprise pollution level, while the enterprise pollution level reduces the GTFP growth in column (1) of Panel B in Table 5. Similar conclusions can be drawn for other mechanism variables in Panel B in Table 5, indicating that official turnover would accelerate sustainable development by inhibiting *enterprise pollution levels*. Overall, these estimates indicate the reliability of Hypotheses 2 and 3, implying that official turnover promotes sustainable development by the deterrent effect and reducing government construction investment.

Heterogeneity Analysis

First, we regress the GTFP growth on the samples of party secretary and mayor turnover—columns (1) and (2) of Table 6 report the estimated results. Through comparison, the turnover of a mayor has a greater promotion effect on sustainable development in China, promoting GTFP growth by an average of 1.98% at a 1% significance level. The reason is that the party secretary mainly engages in local party affairs. Meanwhile, the mayor is responsible for presiding over the work of the municipal people's government and is in charge of all matters related to economic growth and social development in the jurisdiction. Therefore, the mayor's turnover may have a greater impact on the production department.

We further analyzed the impact of whether the local officials were changed normally. Column (3) and (4) of Table 6 examines the impact of different types

Table 7 Heterogeneity analysis II

	Postgraduate		Local		Old	
	No	Yes	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)
Turnover	0.0022	0.0122***	0.0020	0.0164***	0.0064	0.0165***
	(0.0023)	(0.0001)	(0.0017)	(0.0006)	(0.0052)	(0.0025)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.0087***	0.0024	0.0015	-0.0074***	-0.0051***	-0.0021
	(0.0019)	(0.0052)	(0.0015)	(0.0011)	(0.0015)	(0.0023)
Observations	2650	3285	2629	3306	2677	3258
Adjusted R-squared	0.1763	0.1684	0.1664	0.1531	0.1418	0.1392

The values in parentheses are standard errors, and we cluster standard errors to the city level

*** p<0.01, ** p<0.05, * p<0.1

of official turnover on sustainable development. Column (3) examines the impact of normal (including promotion, lateral move, or retirement) official turnover on sustainable development. Column (4) examines the impact of abnormal (including being arrested due to corruption, bribery, and crime or being dismissed) official turnover on GTFP. On average, normal turnover can increase the GTFP by 1.44%, while the abnormal turnover of the mayor's role in promoting GTFP is not significant at the 10% significance level. The possible explanation is that officials dismissed due to corruption and bribery are more likely to develop the economy at any environmental cost.

Second, since the influence of the party secretary is not significant, we proceed from the perspective of the mayor's characteristics to study the impact of their turnover on sustainable development; the regression results are shown in Table 7. Columns (1) and (2) report the heterogeneity of a mayor's educational background, indicating that appointing a mayor with a high degree of education can improve GTFP by 1.22% at the 1% significance level. Highly educated officials have a longer-term vision; they tend to promote economic development and consider the coordinated development of the economy and the environment.

Columns (3) and (4) indicate that a mayor with a local household registration can accelerate sustainable development more effectively than a mayor with a household registration in other areas. The possible explanation is that local officials have a better understanding of local developments to promote sustainable development more efficiently.

Columns (5) and (6) explore the impact of official turnover at different ages on sustainable development, indicating that turnover in older officials is more likely

to promote the growth of GTFP. Older officials (over 53 years old, the mean) have less motivation for promotion and want to gain a good reputation after retirement; therefore, they pay more attention to sustainable development of the economy and improve residents' quality of life.

Conclusion

Based on panel data of 277 prefecture-level and above cities in China from 2004 to 2016, we use an updated method to accurately calculate the GTFP of cities to measure the level of sustainable development, and we use the idea of "event study" to identify the causal relationship between official turnover and sustainable development. The results show that official turnover significantly promotes sustainable development, supported by a series of robustness tests. Furthermore, mechanism analysis shows that official turnover contributes to sustainable development, primarily by the deterrent effect and reducing government construction investment. Furthermore, heterogeneity analysis shows that mayor turnover contributes more to GTFP growth than party secretary turnover, while official turnover with higher education, local *hukou*, and older age have a greater positive impact on sustainable development.

Our study has important theoretical implications. Most literature studied the separate impacts of official turnover on economic development, ecological environment, and enterprise investment and obtained different or even contrasting research conclusions [27, 36, 42, 56]. This study examines the impact of official change on sustainable development from a more macro and systematic perspective to get a more general conclusion. Our research has systematically identified the impact and mechanism of official turnover on sustainable development, enriching relevant research in environmental politics [22, 42, 49].

Moreover, this paper's empirical conclusion has strong practical significance. From the official governance perspective, official turnover provides new ideas for improving governance and environment improvement. In the phase of China's extensive development and pursuit of economic growth, official turnover could promote sustainable development through the deterrent effect and by reducing government investment in construction. The turnover of officials temporarily alleviates political connection and overcapacity, thus promoting sustainable economic development. Therefore, on the one hand, the central government should strengthen the construction of local government officials and reduce the occurrence of some local government officials exchanging ecological environments for economic benefits by strengthening learning. On the other hand, the central government should improve the supervision mechanism for local government officials, reduce the political connection between some local governments and high pollution and energy consuming enterprises, and reduce the occurrence of corruption and other negative events of local government officials. With the transformation of China's economy to high-quality development and the improvement of the central government's governance mechanism for local government officials, the sustainable development of China's

economy will increasingly depend on the governance and protection of resources and the environment during officials' tenure.

The coordinated, sustainable, and efficient development of the economy and environment is a common goal pursued by both developed and developing countries. Although we focus on China, the research perspective, mechanisms, and conclusions could also provide a reference for other countries. Future studies could compare the conclusions of our study with the relevant research in other countries and determine the advantages, disadvantages, and heterogeneity of the effect to supplement the literature in related fields.

A few limitations to our study are worth emphasizing. First, due to our research design, the GTFP calculated is an annual change, limiting our ability to investigate the dynamic relationship between official turnover and GTFP at the monthly or daily level. Second, multiple paths may exist for the effect of official turnover on sustainable development; however, we only examine the two most important paths. Further studies can identify other mechanisms by which official turnover affects GTFP growth. Third, with the transformation of the Chinese economy towards high-quality development from 2017, the Chinese government is striving to achieve sustainable development from various aspects, and the impact of official turnover on sustainable development will also be decreasing. We will further focus on the dynamic effects.

Acknowledgement This work was supported by the Fundamental Research Funds for the Central Universities (No. QCDC-2020-21).

Declarations

Conflict of Interest (COI) Statement We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

1. Afsharian, M., and H. Ahn. 2015. The overall Malmquist index: A new approach for measuring productivity changes over time. *Annals of Operations Research 226*: 1–27.
2. Aizenman, J., and N.P. Marion. 1993. *Policy Uncertainty. Persistence and Growth 1*: 145–163.
3. Angrist, J.D., and J.S. Pischke. 2009. *Mostly harmless econometrics: An empiricist's companion*. USA: Princeton University Press.
4. Berg, S.A., F.R. Førsund, and E.S. Jansen. 1992. Malmquist indices of productivity growth during the deregulation of Norwegian banking, 1980-89. *Scandinavian Journal of Economics 94*: S211–S228.
5. Chen, L., and M. Gao. 2020. The effects of three types of China's official turnover on air quality: A regression discontinuity study. *Growth and Change 51* (3): 1081–1101.
6. Chen, Z., M.E. Kahn, Y. Liu, and Z. Wang. 2018. The consequences of spatially differentiated water pollution regulation in China. *Journal of Environmental Economics and Management 88*: 468–485.
7. Cheng, Z., L. Luo, S. Wang, Y. Wang, S. Sharma, H. Shimadera, et al. 2016. Status and characteristics of ambient $PM_{2.5}$ pollution in global megacities. *Environment International 89*: 212–221.
8. Cinelli, C., A. Forney, and J. Pearl. 2022. A crash course in good and bad controls. *Sociological Methods & Research*. https://doi.org/10.1177/0049124122109955.
9. Colmer, J. 2021. Temperature, labor reallocation, and industrial production: Evidence from India. *American Economic Journal: Applied Economics 13* (4): 101–124.

10. Dahlström, C., and M. Holmgren. 2019. The political dynamics of bureaucratic turnover. *British Journal of Political Science 49* (3): 823–836.
11. Darby, J., C.W. Li, and V.A. Muscatelli. 2004. Political uncertainty, public expenditure and growth. *European Journal of Political Economy 20* (1): 153–179.
12. Deng, Y., Y. Wu, and H. Xu. 2019. Political turnover and firm pollution discharges: An empirical study. *China Economic Review 58*: 101363.
13. Deng, Y., Y. Wu, and H. Xu. 2020. Political connections and firm pollution behaviour: An empirical study. *Environmental and Resource Economics 75*: 867–898.
14. Deschenes, O., H. Wang, S. Wang, and P. Zhang. 2020. The effect of air pollution on body weight and obesity: Evidence from China. *Journal of Development Economics 145*: 102461.
15. Du, X. 2015. Is corporate philanthropy used as environmental misconduct dressing? Evidence from Chinese family-owned firms. *Journal of Business Ethics 129* (2): 341–361.
16. Earle, J.S., and S. Gehlbach. 2015. The productivity consequences of political turnover: Firm-level evidence from Ukraine's orange revolution. *American Journal of Political Science 59* (3): 708–723.
17. Ennser-Jedenastik, L. 2014. Political control and managerial survival in state-owned enterprises. *Governance 27* (1): 135–161.
18. Fan, M., S. Shao, and L. Yang. 2015. Combining global Malmquist–Luenberger index and generalized method of moments to investigate industrial total factor CO_2 emission performance: A case of Shanghai (China). *Energy Policy 79*: 189–201.
19. Färe, R., S. Grosskopf, B. Lindgren, and P. Roos. 1992. Productivity changes in Swedish pharamacies 1980–1989: A non-parametric Malmquist approach. *Journal of Productivity Analysis 3* (1-2): 85–101.
20. Feng, L., T. Fu, and A.M. Kutan. 2019. Can government intervention be both a curse and a blessing? Evidence from China's finance sector. *International Review of Financial Analysis 61*: 71–81.
21. Gao, N., and P. Liang. 2016. Fresh cadres bring fresh air? Personnel turnover, institutions, and China's water pollutions. *Review of Development Economics 20* (1): 48–61.
22. Hajer, M., and W. Versteeg. 2005. A decade of discourse analysis of environmental politics: Achievements, challenges, perspectives. *Journal of Environmental Policy & Planning 7* (3): 175–184.
23. Hu, X., and F. Kong. 2021. Policy Innovation of Local Officials in China: The administrative choice. *Journal of Chinese Political Science 26* (4): 695–672.
24. Hu, K., and D. Shi. 2021. The impact of government-enterprise collusion on environmental pollution in China. *Journal of Environmental Management 292*: 112744.
25. Huang, Y. 2002. Managing Chinese bureaucrats: an institutional economics perspective. *Political Studies 50* (1): 61–79.
26. Jiang, S.S., and J.M. Li. 2021. Do political promotion incentive and fiscal incentive of local governments matter for the marine environmental pollution? Evidence from China's coastal areas. *Marine Policy 128*: 104505.
27. Jiang, Z., Z. Wang, C. Feng, and L. Yi. 2022. Local political turnover, R&D investment leap and corporate innovation performance: Evidence from China. *Science and Public Policy 49* (2): 347–364.
28. Joseph, J. 2023. State, society and environmental security in international relations theory. *Fudan Journal of the Humanities and Social Sciences 16* (2): 171–190.
29. Julio, B., and Y. Yook. 2012. Political uncertainty and corporate investment cycles. *Journal of Finance 67*: 45–83.
30. Kim, D. 2011. Do local policy networks deter the race to the bottom in environmental regulation? The case of South Korea. *Environment and Planning C: Government and Policy 29*: 1037–1053.
31. Kostka, G., and C. Zhang. 2018. Tightening the grip: Environmental governance under Xi Jinping. *Environmental Politics 27* (5): 769–781.
32. Krause, G.A., and M. Zarit. 2020. Understanding political−economic development in the American South. *Economics & Politics 32*: 172–195.
33. Lægreid, O.M., and M. Povitkina. 2018. Do political institutions moderate the GDP-CO_2 relationship? *Ecological Economics 145*: 441–450.
34. Li, K., and B. Lin. 2016. Impact of energy conservation policies on the green productivity in China's manufacturing sector: Evidence from a three-stage DEA model. *Applied Energy 168*: 351–363.
35. Li, X., and D. Wang. 2022. Does transfer payments promote low-carbon development of resource-exhausted cities in China?. *Earth's. Future 10* (1): e2021EF002339.

36. Li, H., and L.A. Zhou. 2005. Political turnover and economic performance: The incentive role of personnel control in China. *Journal of Public Economics 89*: 1743–1762.
37. Li, Z., N. Meng, and X. Yao. 2017. Sustainability performance for China's transportation industry under the environmental regulation. *Journal of Cleaner Production 142*: 688–696.
38. Li, X., Z. Hu, J. Cao, and X. Xu. 2022. The impact of environmental accountability on air pollution: A public attention perspective. *Energy Policy 161*: 112733.
39. Liao, L., C. Zhao, X. Li, and J. Qin. 2021. Towards low carbon development: The role of forest city constructions in China. *Ecological Indicators 131*: 108199.
40. Lin, J.Y., and Z. Liu. 2000. Fiscal decentralization and economic growth in China. *Economic Development and Cultural Change 49* (1): 1–21.
41. Liu, Z., and L. Xin. 2019. Has China's Belt and Road Initiative promoted its green total factor productivity——Evidence from primary provinces along the route. *Energy Policy 129*: 360–369.
42. Liu, H., Z. Zhang, Y. Wang, C. Zhang, and D. Wang. 2022. Testing the correlation between eco-environmental performance and provincial official promotion in China. *Journal of Chinese Political Science*. https://doi.org/10.1007/s11366-022-09834-y.
43. Lu, S., Y. Yao, and H. Wang. 2022. Testing the relationship between land approval and promotion incentives of provincial top leaders in China. *Journal of Chinese Political Science 27*: 105–131.
44. Luo, W., and S. Qin. 2021. China's local political turnover in the Twenty-First Century. *Journal of Chinese Political Science 26*: 651–674.
45. Ma, G., J. Qin, and Y. Zhang. 2023. Does the carbon emissions trading system reduce carbon emissions by promoting two-way FDI in developing countries? Evidence from Chinese listed companies and cities. *Energy Economics 120*: 106581.
46. Mahlberg, B., M. Luptacik, and B.K. Sahoo. 2011. Examining the drivers of total factor productivity change with an illustrative example of 14 EU countries. *Ecological Economics 72*: 60–69.
47. Malmquist, S. 1953. Index numbers and indifference surfaces. *Trabajos de Estadística 4* (2): 209–242.
48. McBeath, J. 2010. Ligang Song and Wing Thye Woo, eds. China's dilemma: Economic growth, the environment, and climate change. *Journal of Chinese Political Science 15*: 215–216.
49. Moser, P., and A. Voena. 2012. Compulsory licensing: Evidence from the trading with the enemy act. *American Economic Review 102* (1): 396–427.
50. Pastor, J.T., and C.A.K. Lovell. 2005. A global Malmquist productivity index. *Economics Letters 88*: 266–271.
51. Piotroski, J.D., and T. Zhang. 2014. Politicians and the IPO decision: The impact of impending political promotions on IPO activity in China. *Journal of Financial Economics 111*: 111–136.
52. Shestalova, V. 2003. Sequential Malmquist indices of productivity growth: An application to OECD industrial activities. *Journal of Productivity Analysis 19*: 211–226.
53. Shi, Q., C. Shi, and F. Guo. 2020. Political blue sky: Evidence from the local annual "Two Sessions" in China. *Resource and Energy Economics 61*: 101165.
54. Shih, V., C. Adolph, and M. Liu. 2012. Getting ahead in the communist party: explaining the advancement of central committee members in China. *American Political Science Review 106* (1): 166–187.
55. Tang, P., Q. Jiang, and L. Mi. 2021. One-vote veto: the threshold effect of environmental pollution in China's economic promotion tournament. *Ecological Economics 185*: 107069.
56. Tian, Z., A. Hu, Y. Chen, and S. Shao. 2023. Local officials' tenure and CO_2 emissions in China. *Energy Policy 173*: 113394.
57. Wang, Z. 2021. The elusive pursuit of incentive systems: Research on the cadre management regime in post-Mao China. *Journal of Chinese Political Science 26* (3): 573–592.
58. Wang, X. 2022. Managing the problem of too much capital: Coal power overcapacity, migrant labor sacrifice, and structural problems in contemporary Chinese political economy. *Environment and Planning E: Nature and Space 5* (4): 1895–1918.
59. Wang, X., and W. Shan. 2013. Chairman's government background, excess employment and government subsidies: Evidence from Chinese local state-owned enterprises. *China Journal of Accounting Research 6* (1): 51–74.
60. Wang, C., J. Hong, M. Kafouros, and M. Wright. 2012. Exploring the role of government involvement in outward FDI from emerging economies. *Journal of International Business Studies 43* (7): 655–676.

61. Wang, Y., Z. Guo, and J. Han. 2021a. The relationship between urban heat island and air pollutants and them with influencing factors in the Yangtze River Delta. *China. Ecological Indicators 129*: 107976.
62. Wang, K.L., B. Zhao, L.L. Ding, and Z. Miao. 2021b. Government intervention, market development, and pollution emission efficiency: Evidence from China. *Science of the Total Environment 757*: 143738.
63. Wang, S., L. Yuan, and B. Gong. 2023. China's agricultural green transition and high-quality development toward carbon neutrality. *Chinese Political Science Review 8* (2): 240–272.
64. Wong, M.Y. 2022. Performance, factions, and promotion in China: the role of provincial transfers. *Journal of Chinese Political Science 27* (1): 41–75.
65. Wu, J., Y. Deng, J. Huang, R. Morck, and B. Yeung. 2014. Incentives and outcomes: China's environmental policy. *Capitalism and Society 9* (1): 1–41.
66. Xia, F., and J. Xu. 2020. Green total factor productivity: A re-examination of quality of growth for provinces in China. *China Economic Review 62*: 101454.
67. Xie, R., Y. Yuan, and J. Huang. 2017. Different types of environmental regulations and heterogeneous influence on "green" productivity: Evidence from China. *Ecological Economics 132*: 104–112.
68. Xu, N., X. Xu, and Q. Yuan. 2013. Political connections, financing friction, and corporate investment: Evidence from Chinese listed family firms. *European Financial Management 19* (4): 675–702.
69. Yan, Y., X. Zhang, J. Zhang, and K. Li. 2020. Emissions trading system (ETS) implementation and its collaborative governance effects on air pollution: The China story. *Energy Policy 138*: 111282.
70. Yang, Z., M. Fan, S. Shao, and L. Yang. 2017. Does carbon intensity constraint policy improve industrial green production performance in China? A quasi-DID analysis. *Energy Economics 68*: 271–282.
71. Yang, Z., M. Lu, S. Shao, M. Fan, and L. Yang. 2023. Carbon regulation and economic growth: City-level evidence from China. *Environmental Impact Assessment Review 99*: 107020.
72. Yin, L., and C. Wu. 2022. Promotion incentives and air pollution: From the political promotion tournament to the environment tournament. *Journal of Environmental Management 317*: 115491.
73. Zhang, N., and Y. Choi. 2014. A note on the evolution of directional distance function and its development in energy and environmental studies 1997–2013. *Renewable and Sustainable Energy Reviews 33*: 50–59.
74. Zhang, X., W. Luo, and D. Xiang. 2022. Political Turnover and Innovation: Evidence from China. *Journal of Chinese Political Science*. https://doi.org/10.1007/s11366-022-09800-8.
75. Zhao, C., M. Deng, and X. Cao. 2021. Does haze pollution damage urban innovation? Empirical evidence from China. *Environmental Science and Pollution Research 28* (13): 16334–16349.
76. Zhao, X., M. Jia, and Z. Zhang. 2023a. Promotion vs. pollution: City political status and firm pollution. *Technological Forecasting and Social Change 187*: 122209.
77. Zhao, M., R. Shi, R. Du, and L. Yao. 2023b. The pathway to China's carbon–neutral agriculture: Measures, potential and future strategies. *Chinese Political Science Review 8* (2): 304–324.
78. Zheng, W., and P. Chen. 2020. The political economy of air pollution: Local development, sustainability, and political incentives in China. *Energy Research & Social Science 69*: 101707.
79. Zhou, P., B.W. Ang, and D.Q. Zhou. 2012. Measuring economy-wide energy efficiency performance: A parametric frontier approach. *Applied Energy 90* (1): 196–200.
80. Zhu, X., and Y. Yang. 2023. Did the turnover of officials reduce the pollution emission of enterprises? *Science of The Total Environment 868*: 161459.
81. Zodrow, G.R., and P. Mieszkowski. 1986. Pigou, Tiebout, property taxation, and the underprovision of local public goods. *Journal of Urban Economics 19*(3): 356–370.
82. Zuo, C. 2022. Political selection institutions and policy performance: Evidence from China. *Chinese Political Science Review*:1-23.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Xing Li PhD in Economics, is a lecturer in the College of Economics and Management, Southwest University. The research fields of interest are mainly in environmental economics and political economics. His research has been published in such journals as Earth's Future, Energy Policy, Journal of Business Research, Ecological Indicators, and Cities.

Chunkai Zhao PhD in Economics, is an associate professor in the College of Economics and Management, South China Agricultural University. His areas of interest are demographic and labor economics, rural development, and environmental economics. His research has been published in such journals as Journal of Happiness Studies, Journal of Business Research, China Economic Review, Cities, Transport Policy, and Environment, Development, and Sustainability.

Jianhua Cao PhD in Economics, is the professor and doctoral supervisor of the Institute of Finance and Economics Research, Shanghai University of Finance and Economics. His main research area is energy and environmental economics. He has published more than 70 papers, and these journals include Renewable and Sustainable Energy Reviews, Business Strategy and the Environment, Journal of Environmental Management, Tourism Economics, and Economic Research Journal (in Chinese).

China Economic Review 87 (2024) 102247

Contents lists available at ScienceDirect

China Economic Review

journal homepage: www.elsevier.com/locate/chieco

Online markets, offline happiness: E-commerce development and subjective well-being in rural China

Binhui Wei [a], Chunkai Zhao [a,b,*], Mingzhong Luo [a]

[a] College of Economics and Management, South China Agricultural University, No.483, Wushan Street, Tianhe District, Guangzhou, Guangdong Province, China
[b] Agricultural and Rural Policy and Reform Innovation Laboratory, South China Agricultural University, No.483, Wushan Street, Tianhe District, Guangzhou, Guangdong Province, China

ARTICLE INFO

JEL codes:
I31 R28 O38

Keywords:
E-commerce development
Subjective well-being
National Rural E-commerce Comprehensive Demonstration County
Digital divide
Rural China

ABSTRACT

Employing a quasi-natural experiment of the National Rural E-commerce Comprehensive Demonstration County (NRECDC) policy, our research aimed to explore the causal effect of e-commerce development and residents' subjective well-being (SWB) in rural China. By adopting the data of China Family Panel Studies (CFPS) from 2014 to 2020 and using a staggered difference-in-differences (DID) method, we found that rural e-commerce development resulted in an average increase of 2.40% in happiness scores. This positive effect can be explained by absolute material welfare enhancement, relative disparity reduction, and social capital accumulation. Further evidence suggests that e-commerce users are more likely to benefit from the NRECDC policy and that the synergistic effect of digital finance and e-commerce contributes to rural residents' SWB. However, we revealed that the potential digital divide effect in e-commerce development, due to technological barriers and self-digital exclusion, has not enhanced the SWB of some relatively disadvantaged groups in rural China. Our findings enrich the research field of digital well-being and shed light on the potential digital divide effects of digital tools such as e-commerce in rural China.

1. Introduction

With the increasing abundance of material life, the public, government, and academics have shifted their focus to subjective well-being (SWB).[1] How to improve people's SWB has long been a concern of scholars. Previous studies showed that people's happiness is affected by a considerable number of factors, such as personal characteristics (e.g., Bartolini & Sarracino, 2015; Hartog & Oosterbeek, 1998; Ram, 2010; Zhao & Guo, 2022), family characteristics (Myrskylä & Margolis, 2014; Pina & Bengtson, 1993; Schnittker, 2008), economic characteristics (e.g., Caporale et al., 2009; Clark et al., 2008; Di Tella et al., 2001; Knight & Gunatilaka, 2022; Kumar et al., 2021), and social and institutional factors (e.g., Ahuvia, 2002; Knight & Gunatilaka, 2010; Zhang et al., 2017; Zhou et al., 2023). Recently, with the arrival of the digital era, the impact of digital economic sectors on people's SWB has inevitably attracted academic interest (e.g., Lei et al., 2023; Lu & Kandilov, 2021; Zheng & Ma, 2021). However, the effect of e-commerce development, particularly in rural areas, on residents' SWB has not been thoroughly explored (Jin et al., 2020). Our study aims to fill this deficiency by

* Corresponding author.
E-mail address: ckz_zhao@163.com (C. Zhao).

[1] In our study, SWB and happiness are considered consistent and interchangeable.

https://doi.org/10.1016/j.chieco.2024.102247
Received 8 November 2023; Received in revised form 27 May 2024; Accepted 27 July 2024
Available online 3 August 2024
1043-951X/© 2024 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

confirming the causal relationship between e-commerce development and residents' SWB in rural China.

E-commerce is a business operation mode in which buyers and sellers rely on the Internet platform and digital financial tools to realize online shopping for consumers and online transactions for merchants. China is currently one of the most successful countries in terms of e-commerce development and has the largest e-commerce market in the world (Guo et al., 2022; Zang et al., 2023). In recent years, with the popularization of the Internet and the construction of logistics infrastructure in rural areas, more and more Chinese rural residents are becoming e-commerce users. This transition is quantitatively evidenced by the meteoric rise in rural online retail sales, which soared from 180 billion yuan in 2014 to an impressive 2.49 trillion yuan by 2023, marking a thirteenfold increase and an average annual growth rate of 33.90%. More tellingly, the proportion of rural online retail sales within the national total has expanded from 6% to 16.17% over the same period.[2] These trends show that rural markets are becoming powerful drivers of e-commerce expansion. E-commerce is positively changing the daily lives of rural households in China and bringing them happiness.

There are two branches of literature closely related to our study. The first strand of research looked closely at the benefits of e-commerce development but did not develop a unified view. Scholars analyzed the positive role of e-commerce development at both macro and micro levels. At the macro level, previous literature recognized that e-commerce contributed to narrowing the urban-rural gap, alleviating poverty, and reducing inequality (Chen et al., 2021; Liu et al., 2020; Liu & Zhou, 2023; Qin et al., 2023; Tang & Zhu, 2020; Wang, Miao, et al., 2021). At the micro level, existing studies found that e-commerce could help promote rural household entrepreneurship, consumption upgrading, and income growth (Fan et al., 2018; Li & Qin, 2022; Martinsons, 2008; Sebora et al., 2009; Yang et al., 2017; Zang et al., 2023). Thus, these studies emphasized the inclusive character of e-commerce development. However, some other scholars held an opposing view, because e-commerce is inaccessible to some residents as a digital tool due to weak digital infrastructure and logistics (Couture et al., 2021; Yin & Choi, 2022). Moreover, digital technology thresholds and digital self-exclusion effects may also hinder e-commerce use (Liu et al., 2022), especially in rural areas (Niu et al., 2022; Xia & Lu, 2008; Yeung et al., 2017).

The second strand of literature deals with digital well-being, a topic on which scholars are similarly divided. First, there is a considerable body of research confirming the happiness effects of Internet use, stemming from the advantages of the Internet in terms of information transfer, improved livelihoods, expanded social networks, and increased productivity (e.g., DiMaggio & Bonikowski, 2008; Lu & Kandilov, 2021; Mitchell et al., 2011; Nie et al., 2017). However, some scholars argued that the Internet might reduce individual SWB due to overuse, comparison, and information cocoon effects (Bao et al., 2021; Zhan & Zhou, 2018). Second, with the proliferation of digital technologies, researchers found that while the adoption of digital tools—like mobile internet, digital payments, and e-commerce—can enhance SWB (Lei et al., 2023; Lohmann, 2015; Zhao et al., 2024; Zheng & Ma, 2021), there are concerns about the digital divide. This divide could restrict the benefits of such tools for underprivileged segments of the population, such as those in rural areas (Lu et al., 2024; Nie et al., 2017). Additionally, another literature, exemplified by Jin et al. (2020), is closely related to our study about the happiness effect of e-commerce. Based on the case of Tonglu County in Zhejiang Province, they found that rural e-commerce service centers enhance local farmers' SWB by providing several convenient services.

Unlike the above studies, we employed the nationally representative data from the China Family Panel Studies (CFPS) and used a quasi-natural experiment with the National Rural E-commerce Comprehensive Demonstration County (NRECDC) policy to assess the causal effect of rural e-commerce development on residents' SWB (Chen et al., 2021; Qin et al., 2023). This policy shock is more exogenous compared to digital use in prior studies. By using a staggered difference-in-differences (DID) empirical design, we were able to compare changes in people's SWB in NRECDCs and non-NRECDCs, which helps us clarify the controversy about the impact of e-commerce development on the welfare of rural households in the previous literature mentioned above. Moreover, our analysis extends to uncovering the potential mechanisms and examines the heterogeneity of these effects at different dimensions.

We had some interesting findings that complement existing research in the following ways. First, our estimates suggest that e-commerce development promotes the happiness scores of rural residents, which enriches a class of studies on the benefits of e-commerce development. Prior literature involved in e-commerce benefits mainly focuses on material life (e.g., Li & Qin, 2022; Yang et al., 2017; Zang et al., 2023), and only little addressed the potential effect on SWB. In particular, unlike Jin et al. (2020), this study aims to identify causal effects between e-commerce development and SWB enhancement in rural households by using a nationally representative dataset and a staggered DID approach, and by delving deeper into the underlying mechanisms and heterogeneous effects. Thus, our study contributes to a further understanding of the impact of e-commerce development on people's subjective welfare.

Second, this study further added to the literature in the recently emerging field of digital well-being based on the case of e-commerce (e.g., Lei et al., 2023; Nie et al., 2017; Zhao et al., 2024). Also, in terms of the happiness effect of e-commerce, existing studies focused on e-commerce use (Zhan & Zhou, 2018; Zheng & Ma, 2021). We provided new evidence from a particular perspective of the e-commerce development policy in rural China. Moreover, our estimates show the synergistic effects of digital finance and e-commerce, which help to further enhance the SWB of rural residents in China. As more and more countries enter the digital economy era, this research raises a new topic about how to utilize multiple digital tools to achieve a happy life for a broader population.

Third, this work supplements the theoretical mechanisms by which rural e-commerce development translates into happiness improvement and tests them empirically. We enriched the mechanisms from three dimensions: relative material welfare enhancement, relative disparities reduction, and social capital accumulation. In addition, we found that the entrepreneurial effect of e-commerce, previously emphasized in the literature (Sebora et al., 2009; Tang & Zhu, 2020; Zang et al., 2023), does not effectively translate into increased SWB for rural residents in China. Thus, this study has, to a certain extent, bridged the research gap in this area.

[2] Data source: http://www.mofcom.gov.cn/.

Last, we offered new insights into the theoretical divergence between the digital dividend and the digital divide. We found that in rural China, e-commerce development has not boosted the SWB of some disadvantaged groups, such as the elderly, the low-educated, and those from poor households, which hints at a potential digital divide effect. Importantly, our estimates imply that the digital divide effect of e-commerce may stem from digital technology thresholds and digital exclusion. However, further enhancement of e-commerce penetration is expected to alleviate this challenge. Thus, in the debate between the digital dividend and the digital divide, our results tend to support the latter (Xia & Lu, 2008; Yeung et al., 2017), emphasizing that at present, rural e-commerce development needs to be further made more inclusive to reach more disadvantaged groups in the digital age. In summary, our results show a mixed effect of e-commerce development on SWB in rural China, which may provide direction for subsequent research on mitigating the digital divide in e-commerce.

2. Background

The Ministry of Finance and the Ministry of Commerce of China introduced the NRECDC policy in 2014, aiming to bridge the "last kilometer" of rural e-commerce. The first pilots were launched in 56 counties in eight provinces and these NRECDCs are provided with a financial support package amounting to 20 million yuan per year, allocated to construct a rural e-commerce system. Key areas of focus lie in the integration of existing distribution resources, the promotion of rural logistics and distribution systems, and the strengthening of rural e-commerce application capacity (Chen et al., 2021). Moreover, this policy also actively broadens the scope of e-commerce applications in rural areas, strengthens the two-way circulation of industrial goods to the countryside and agricultural products to the city, and realizes the diversification of sales channels for agricultural products (Qin et al., 2023). From 2014 to 2020, the NRECDC policy has been applied cumulatively in 1456 counties. The temporal and geographical distribution of NRECDCs is illustrated in Figs. 1 and 2.

Rural e-commerce transactions have grown significantly with the construction of NRECDCs. As shown in Fig. 3, rural online retail sales reached 1.79 trillion yuan in 2020. More importantly, rural e-commerce sales saw a consistent rise, rising from 1100 billion yuan in 2013 to 8945.40 billion yuan by 2016, marking an increase of more than eightfold. From 2016 onwards, while the growth rate slowed compared to the previous surge, there was still a steady rise in sales each year. By the end of 2020, sales reached approximately double the figure from 2016.[3] These statistics show a substantial expansion or possibly broader adoption of e-commerce platforms in rural China after NRECDC policy implementations.

Moreover, based on the NRECDC policy, a total of 2120 county-level e-commerce public service centers and logistics distribution centers were established nationwide, complemented by 137,000 village-level e-commerce service sites. These sites provide nearly full coverage of townships and courier outlets, raising the proportion of direct courier deliveries to rural regions to over 50%. These efforts culminated in the collection and delivery of over 30 billion express parcels in rural areas, accounting for 36% of the national total.[4] Conclusively, the robust implementation of the NRECDC policy has played a pivotal role in accelerating rural e-commerce development in China (Chen et al., 2021).

It is important to mention that the policy goal of NRECDC constructions is twofold. In addition to promoting rural revitalization and reducing urban-rural disparities, it also seeks to alleviate poverty and promote the development of revolutionary areas through e-commerce development. Thus, the NRECDC policy prioritizes pilot projects in some less developed areas, like poor counties and old revolutionary areas (Chen et al., 2021; Qin et al., 2023). After considering the pilot selection criteria, the steady progression of NRECDCs in batches provides good conditions for quasi-natural experiment analysis to identify potential policy effects (Liu et al., 2023; Niu et al., 2022).

3. Theoretical mechanisms

In this section, we discussed the theoretical mechanisms by which e-commerce affects the SWB of rural residents, including absolute material welfare enhancement, relative disparity reduction, and social capital accumulation. Fig. 4 illustrates the theoretical framework of this study.

3.1. Absolute material welfare enhancement

E-commerce may improve rural residents' happiness by enhancing their absolute material welfare. Income and consumption are key objective indicators for measuring the utility and well-being, and income growth and consumption upgrading are important ways to improve residents' material welfare. However, due to relatively closed market environments and high transaction costs, rural households in China have long suffered from the challenges of low consumption and liquidity constraints (Curtis et al., 2015; Démurger & Xu, 2011; Zhao et al., 2022). Fortunately, e-commerce has brought new transaction models that may help address this challenge to improve rural residents' material welfare. For instance, from the supply side, e-commerce is likely to promote online market sales of rural households (Tang & Zhu, 2020; Zang et al., 2023), such as opening up online channels for the sale of agricultural products (Wang, Luo, et al., 2021). From the demand side, it reduces the transaction threshold for rural consumers with lower prices (Couture et al.,

[3] The data come from the China Rural E-Commerce Development Report (2021−2022) and the website is https://ciecc.ec. com.cn/upload/article/20221024/20221024113120420.pdf.

[4] Data source: https://www.gov.cn/xinwen/2021-07/03/content_5622668.htm.

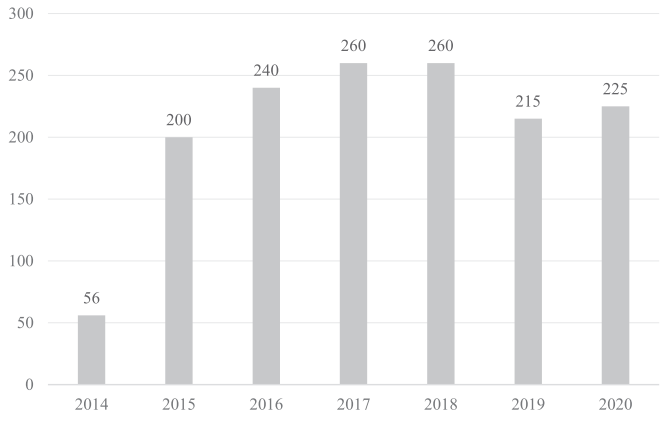


Fig. 1. The temporal distribution of the NRECDCs in batches.

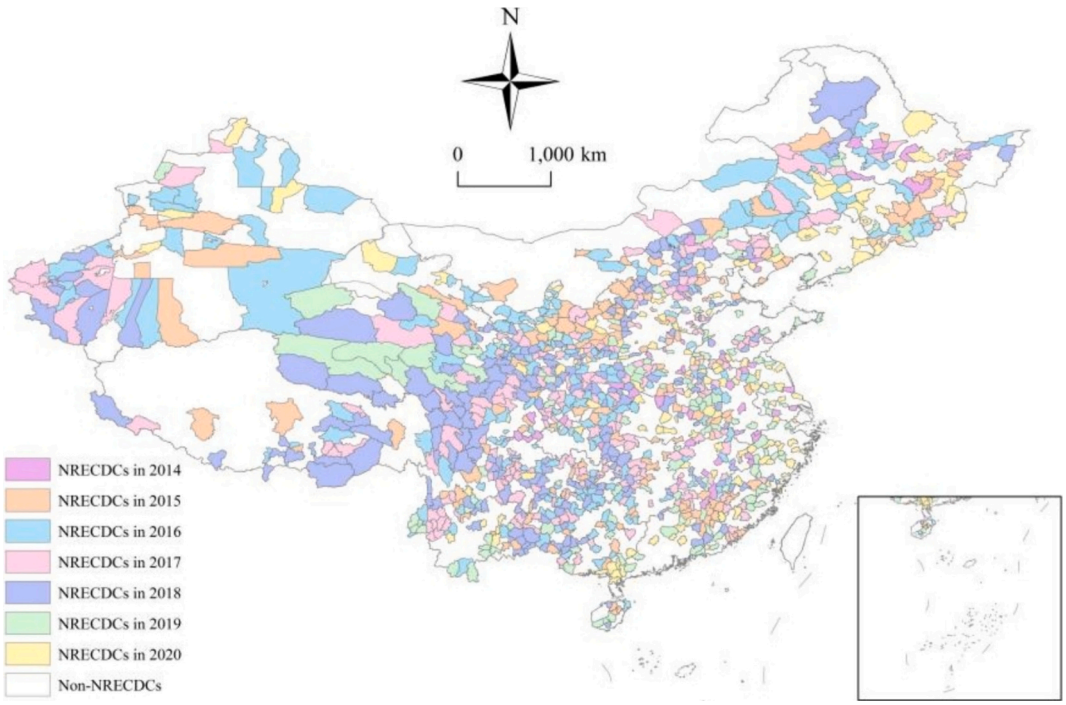


Fig. 2. The spatial distribution of the NRECDCs in batches.

2021; Zhao et al., 2022).

With the help of electronic information technology, digital financial tools, and modern logistics systems, e-commerce bridges areas with smaller local markets and lower market accessibility (Li & Qin, 2022; Tang & Zhu, 2020). This helps reduce transaction costs due to geographic and information factors, thereby facilitating market participation by rural households (Batjargal, 2007; Sebora et al., 2009; Zang et al., 2023). On one hand, e-commerce directly facilitates online sales of agricultural products, which in turn increases rural households' absolute incomes through entrepreneurial effects (Liu et al., 2020; Qi et al., 2019). Moreover, e-commerce development can stimulate rural business entities such as leading enterprises and family workshops to participate in online sales (Wang, Miao, et al., 2021) and take advantage of scale effects to expand output (Liu et al., 2022; Yang et al., 2017). This scale effect may have positive spillover effects on household incomes, such as increasing non-farm jobs and extending supply chains of local agricultural products. On the other hand, e-commerce facilitates the seamless distribution of industrial products and consumer goods across urban and rural areas, overcoming temporal and spatial constraints. This integration significantly lowers transaction costs (Fan et al., 2018; Zhou et al., 2023). Furthermore, e-commerce helps to narrow the differences in price, variety, and quality of commodities across regions, reducing information asymmetry and resolving the conflict between supply and demand (Li et al., 2019; Lin et al., 2022; Tran, 2021). Through channels such as online shopping, the increasingly diverse consumption needs of rural consumers can be satisfied (Jo

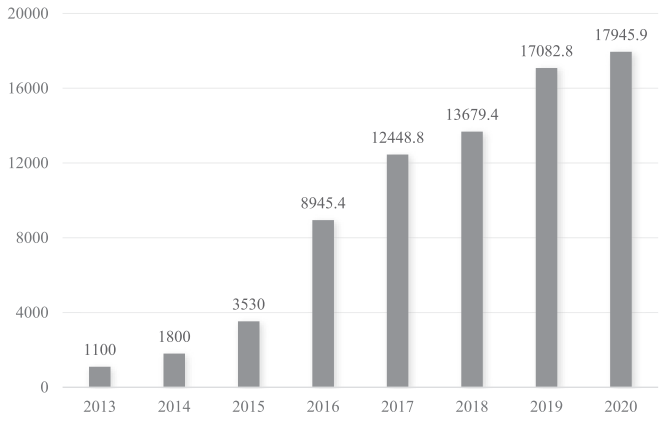


Fig. 3. Rural e-commerce sales in China.
Data source: The China Rural *E*-Commerce Development Report (2021−2022).

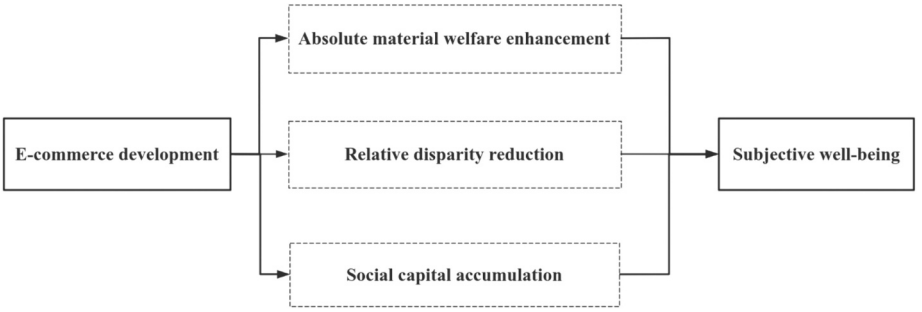


Fig. 4. Theoretical framework.

et al., 2019; Zhao et al., 2022).

The amplification of happiness often hinges on advancements in life quality and material welfare. A considerable body of literature confirmed the positive correlation between material welfare and happiness, suggesting that both absolute income and consumption contribute to people's SWB (e.g., Ahuvia, 2002; Clark et al., 2008; Dehejia et al., 2007; Zheng & Ma, 2021; Zhou et al., 2023). Thus, we considered that e-commerce may enhance rural residents' SWB by improving their absolute material welfare.

3.2. Relative disparity reduction

In addition to absolute material welfare, e-commerce development has the potential to reduce relative disparities, such as income and consumption disparities and relative social status, which further affects the SWB of rural residents. Individuals frequently engage in social comparison by selecting benchmarks with similar social attributes and evaluating their welfare against these references. A notable consequence of perceived discrepancies in entitlements is the emergence of relative deprivation, adversely affecting SWB (Wang, Wang, et al., 2021). This phenomenon underscores the significance of equitable access to resources and opportunities. Particularly for rural groups, they are often marginalized due to logistical, financial, and information barriers and rarely enjoy the benefits of a modern market economy (Niu et al., 2022; Zhao et al., 2022). However, e-commerce has the potential to bridge this gap (Li & Qin, 2022; Yin & Choi, 2022). By providing equitable access to markets, information, and opportunities, rural e-commerce platforms diminish this relative disparity—a key detractor from SWB (Asher & Novosad, 2020; Dohman et al., 2010; Lei et al., 2023).

First, e-commerce development has transformed traditional commerce models by enabling more agricultural products to access broader markets. Previously, the geographical isolation of rural areas limited the sales channels available to local producers, confining them to low-income levels. E-commerce platforms, however, offer an alternative by connecting rural producers directly with consumers nationwide, thereby eliminating many of the logistical and financial barriers associated with traditional market access (Guo et al., 2022; Yang et al., 2017). These facilities are likely to further reduce the relative income gap. Previous evidence confirmed the pivotal role of e-commerce development in narrowing the urban-rural income divide and fostering poverty alleviation (Chen et al., 2021; Li et al., 2019; Li et al., 2021). The positive income distributional effects of e-commerce have the potential to further ameliorate income disparities by trickling down to higher levels of income for disadvantaged rural groups (Aghion & Bolton, 1997; Li & Qin, 2022; Yin & Choi, 2022). The decline in relative income disparity brought about by the e-commerce development may boost the SWB of rural residents.

Next, e-commerce also plays a critical role in narrowing the consumption disparity. Historically, rural consumers in China have faced limited access to goods and services, both in terms of variety and quality. This has been primarily due to logistical challenges and the high cost of distribution in less accessible areas. With the rise of e-commerce, however, rural residents can now easily buy a variety of goods that were previously difficult to purchase offline (Falk & Hagsten, 2015; Fan et al., 2018; Li et al., 2021). This change not only enhances the quality of life for rural residents but also promotes consumer rights and interests in these regions (Zhao et al., 2022). Furthermore, the increased competition among online retailers has the added benefit of driving down prices, making a wider variety of goods more affordable to low-income rural residents (Chen et al., 2021; Liu & Zhou, 2023). This narrowing of the consumption disparity may also contribute to improving the SWB of rural residents.

Last, the role of e-commerce development in reducing the relative disparity is not only at the economic level but also potentially in improving the relative social status of rural residents. Through e-commerce participation in online markets, rural residents gain visibility and recognition, challenging traditional perceptions of rural inferiority and marginalization. This visibility translates into enhanced social status and a stronger sense of belonging (Kumar et al., 2021; Zhao et al., 2024). As rural consumers engage in e-commerce, they also acquire valuable digital literacy and entrepreneurial skills, further elevating their social standing and self-esteem (Couture et al., 2021). These skills enable them to engage in the digital economy, a domain that was previously dominated by urban dwellers or high-income groups (Liu & Zhou, 2023; Yin & Choi, 2022). All those shifts in social dynamics contribute to reducing the social status disparity, thereby enhancing their SWB (Chen et al., 2021; Li et al., 2021; Lohmann, 2015).

3.3. Social capital accumulation

E-commerce is a digital platform that amalgamates social, commercial, and economic activities, which may further promote social capital accumulation among rural households. This facilitation can be understood in terms of social networks, social interactions, and social trust. First, in terms of social networks, as e-commerce penetrates the daily lives of rural residents, it further removes spatial constraints and strengthens connections with relatives and friends through commerce activities (Lin et al., 2022; Zhou et al., 2021). Moreover, e-commerce technological advancement supplements traditional communication means and reinforces offline social networks (Liu et al., 2020; Qi et al., 2019), aiding individuals in maintaining and consolidating existing social connections.

Second, e-commerce is likely to facilitate social interactions. E-commerce is an effective vehicle for carrying and transmitting business information, where different users can communicate with each other (Liu et al., 2020; Tran, 2021). More importantly, relying on the low cost, high efficiency, and popularity of the Internet, e-commerce consumers or entrepreneurs may become partners (Teo & Liu, 2007; Yoon, 2009). Based on their identity and similar interests, they may exchange business ideas effectively, promoting the accumulation of social capital. For rural residents, e-commerce training programs also help to promote social interactions, such as participation in training courses and learning organizations.

Third, e-commerce may foster a greater level of social trust among rural residents. The virtual social networks formed through e-commerce, with mechanisms like "post-sale credit evaluation," help reduce mistrust between rural residents and outsiders, addressing the information asymmetry and uncertainty typical of the anonymized online market (Kwon & Arenius, 2010; Li & Ku, 2018). These regulations engender greater willingness among rural residents to trust and complete transactions or engage in social interactions with diverse individuals, thereby fostering a trust attitude based on transactions (Doney et al., 1998; Lin et al., 2019).

Obviously, the richer the social capital accumulation, the more powerful the resources an individual can mobilize. These resources may provide effective physical support and emotional support and counteract risky events. Therefore, there is a strong link between social capital accumulation and an individual's SWB (e.g., Bartolini & Sarracino, 2015; Helliwell, 2006; Ram, 2010). To sum up, we believe that by promoting social capital accumulation, e-commerce development is likely to enhance rural residents' SWB.

4. Research design

4.1. Data resource

The data for our study are sourced from the publicly accessible CFPS, conducted by the Institute of Social Science Survey of Peking University. This multidisciplinary social tracking survey covers a broad array of topics including economic activities, educational attainment, family relationships, and family dynamics, with a focus on the economic and non-economic well-being of the Chinese population. The survey is expansive, covering 25 provinces, municipalities, and autonomous regions, which comprise 94.5% of the total population of China. Initiated in 2010, the survey has been carried out biennially. Given concerns about the setting of variables and the timeliness of data,[5] the empirical analyses in this study primarily utilize data from the four surveys carried out in 2014, 2016, 2018, and 2020. Relevant data about the list of NRECDCs was obtained from the website of the Ministry of Commerce.[6]

The data filtering rules are strictly enforced in the following manner. First, the initial household head sample was carefully scrutinized,[7] and we elected to omit responses from individuals below the age of 18 and those above 80, to reduce potential bias and

[5] Such as inconsistency in SWB statistics in the 2010 wave and too much missing data on SWB in the 2012 wave.

[6] Data source: http://www.mofcom.gov.cn/.

[7] According to the CFPS questionnaire, the household head is here the primary decision-maker in daily economic activities in a family and not necessarily the person in charge as registered in the household register. Household head's SWB can reasonably represent the overall happiness level of the household (Zhou et al., 2023).

variability within the sample. Second, we excluded non-rural households according to the *hukou* and place of residence, with the final rural households coming from 140 different counties. Among them, 63 counties have been classified as NRECDCs and there were 10, 15, 8, 11, 14, and 5 NRECDCs in each batch from 2015 to 2020, respectively. Notably, these counties were first exposed to the NRECDC policy in 2015, meaning that our sample from CFPS data did not address the first pilots in 2014.[8]

4.2. Empirical strategy

Given that the approval of the NRECDC policy did not occur simultaneously across all regions, traditional DID models are limited to evaluating the policy effect at a single time point (Callaway & Sant'Anna, 2021; Zhou et al., 2023). Consequently, in our model specification, we employed a staggered DID approach to examine the influence of rural e-commerce development on individuals' SWB by accommodating the staggered implementation of the policy across diverse regions (Goodman-Bacon, 2021; Liu et al., 2023; Niu et al., 2022; Wei et al., 2024). The happiness effect of e-commerce can be determined by testing whether the increase in SWB scores of rural residents in pilot counties after the implementation of the NRECDC policy is higher than those from non-pilot counties. The baseline DID specification is as follows:

$$SWB_{i,c,t} = \alpha_0 + \alpha_1 NRECDC_{c,t} + \alpha_2 X_{i,t} + \mu_c + \delta_t + \gamma_t H_c + \varepsilon_{i,c,t} \tag{1}$$

where $SWB_{i,c,t}$ is the explained variable which indicates SWB scores of the respondent i in county c and year t. In the CFPS questionnaire, respondents' SWB is evaluated based on their answers to the question: "How happy do you think you are?" Options are integers from 0 to 10, with larger values indicating greater happiness (Knight & Gunatilaka, 2010; Sun et al., 2016). It is important to note that *SWB* is an ordered variable. When the explained variable has numerous categories (five or more), the coefficients and significance of econometric regression results derived from the linear OLS model do not exhibit substantial discrepancies from those obtained through the nonlinear Oprobit model, provided that the model is correctly specified (Ferrer-i-Carbonell & Frijters, 2004). The estimates garnered from OLS methods are more intuitive and straightforward to interpret, and this approach has been extensively utilized in established happiness studies (Brockmann et al., 2009; Zhou et al., 2023). Therefore, we employed OLS estimates in the main specification and the Oprobit estimates for robustness checks.

The core explanatory variable $NRECDC_{c,t}$ indicates whether county c is selected as an NRECDC in year t. In the DID framework, this variable is assigned according to the following rules. If a county is NRECDC, set to 1, and 0 otherwise. The coefficient α_1 is the core DID estimate we were interested in, representing the causal effect of e-commerce development on the SWB of rural residents in China. The definitions of these two key variables are reported in the Panel A of Table A1 in Appendix A.

The basic control variable set is denoted by $X_{i,t}$, including household head and household characteristics. At the household head level, control variables include *Age*, *Age2*, *Gender*, *Marriage*, *Education*, and *Self-rated health*. These variables are introduced based on evidence in the literature that suggests significant impacts on an individual's SWB. For instance, education level has been linked to higher SWB due to better economic opportunities and awareness (Lohmann, 2015). Similarly, self-rated health is a strong predictor of happiness, reflecting the critical role of physical well-being in overall life satisfaction (Lu & Kandilov, 2021). At the household level, variables such as *Family structure*, *Major event*, *Additional properties*, *Credit availability*, and *Household deposit* are considered. These are chosen for their potential influence on a household's economic stability and capacity to absorb shocks, which directly correlates with the SWB of its members (Clark et al., 2019; Zhao & Guo, 2022; Zhou et al., 2023). Specific definitions for these control variables are shown in the Panel B of Table A1 in Appendix A.

The terms μ_c and δ_t denote county fixed effect (FE) and year FE, respectively, which helps us control for invariant selection bias at the county level and absorb common time-varying shocks to rural households. $\varepsilon_{i,c,t}$ is a random disturbance term. Moreover, H_c refers to county-level characteristics in 2014, and γ_t is a vector of year-specific coefficients. Controlling for this set of interaction terms means that we allowed the time trend in rural residents' SWB to vary with county characteristics in the initial period. The main concern here is that the choice of NRECDCs was not randomly assigned but was the result of government selection based on county characteristics (Chen et al., 2018; Gentzkow, 2006; Liu et al., 2023). For instance, as mentioned above, revolutionary areas may be more likely to become NRECDCs.

Specifically, we selected some county characteristics, including *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*.[9] The screening process for these variables is described below. First, we looked at differences in a range of socioeconomic county characteristics between the treated and control groups before the policy was implemented, as shown in Table A2 of Appendix A. The *t*-test results suggest that the means of these variables are almost significantly different between the two groups, implying that they may be potential selection criteria for a county to become an NRECDC. Second, we regressed these variables on NRECDC in column (4), finding that only the coefficients for these five variables are significant. That is, these five variables are pivotal in affecting the NRECDC treatment status and are essential for inclusion in our primary specification (Niu et al., 2022). Third, we constructed separate cross-sectional data by NRECDC batches and the results are presented in Table A3 of Appendix A. These characteristic variables are statistically significant in most columns. Finally, in addition to the economic factors and the characteristics of revolutionary areas that are focused on at the policy level, the other three variables are closely related to e-commerce development, especially the high dependence of e-commerce development on geographic conditions and transportation and logistics (Lin et al., 2022; Liu et al., 2020; Qi et al.,

[8] The potential reason for this is that the first batch included only 56 counties in 2014 (as shown in Fig. 1), far fewer than the later batches.
[9] Specific definitions are shown in Panel C of Table A1 in the Appendix A.

2019). Therefore, it is theoretically feasible to select these variables.

We further tested the effect of controlling for these county-level characteristics, as Table 3 shows that some individual and household characteristics remain significantly different between the NRECDC and Non-NRECDC groups. To alleviate potential concerns about pilot selectivity, we observed whether the difference between these two groups were reduced with the inclusion of these particular control variables (Ma & Mu, 2020). Table A4 in Appendix A presents the regression results. In columns (1) and (2), without controlling for the selected county characteristics, the coefficients for *Age*, *Gender*, *Education*, *Credit availability*, and *Household deposit* are significant. However, in column (3), after adding these county-level predetermined characteristics, all household head and household characteristic variables except for *Gender* lose their significance. This implies that the observed differences are primarily driven by county-level factors rather than intrinsic household characteristics. In other words, by rigorously controlling for the interaction effects of these county characteristics and the year FE, we effectively mitigate potential endogeneity issues associated with the pilot selection (Chen et al., 2018; Liu et al., 2023). Moreover, we attempted to use county characteristic variables as predictors for propensity score matching (PSM) in our robustness tests (Zhou et al., 2023), which further increased comparability between NRECDCs and non-NRECDCs.

Another concern comes from the heterogeneous treatment effects and potential spillover effects in staggered DID designs. Because the NRECDC is a year-by-year pilot policy, treatment effects vary by group or time, leading to potential bias in our DID estimates (Callaway & Sant'Anna, 2021; De Chaisemartin & d'Haultfoeuille, 2020). We employed a Bacon decomposition to detect the effect of this bias on our main estimates (Goodman-Bacon, 2021). Moreover, in the DID framework, if policy spillover effects exist, the stable unit treatment value assumption (SUTVA) is then violated, implying that rural residents in non-NRECDCs are also subject to policy shocks. As a result, the control group is contaminated, obtaining biased DID estimates (Lu et al., 2019). In the robustness checks, we excluded potential spillover effects through empirical analysis.

4.3. Descriptive statistics

Table 1 presents descriptive statistics for our main variables mentioned above, offering insightful trends and averages. Among our core explanatory variables, it is revealed that about 29.54% of rural residents lived in NRECDCs. The SWB of rural residents, evaluated on a scale from 0 to 10, averages at 7.22. This figure suggests that rural residents generally perceive their well-being positively, aligning with findings from other recent studies on rural residents' SWB in China. For example, Luo et al. (2021) observed an average SWB of 7.41 in 2016, while Zheng and Ma (2021) documented an average SWB of 7.27 in 2019. In addition, we found that rural residents in NRECDCs had an average SWB score of 7.20 compared to 7.22 scores for those in non-NRECDCs. As mentioned earlier, some less developed areas have priority in the NRECDCs , which may result in lower SWB for rural residents in the treated group than those in the control group without considering any control variables and FEs, although this difference was not statistically significant.

A comprehensive understanding of the effects of this shock forms the focal point of our study. The control variables elucidate further demographic and socioeconomic characteristics. A significant portion of the respondents are married, and the median level of educational attainment aligns roughly with primary school completion. The sample exhibits a nearly balanced gender distribution, with approximately equal proportions of males and females. Additional housing property ownership is reported among 13.58% of the respondents. Furthermore, differences for most of the control variables across these two groups were significant, suggesting that controlling for these individual and household characteristics may help to further reduce the interference of these heterogeneous factors. These statistics, when juxtaposed with existing literature, corroborate our understanding of rural China's socioeconomic landscape as depicted in prior studies (Hu et al., 2021; Li & Qin, 2022; Zhao et al., 2022).

Table 1
Descriptive statistics.

	Full sample			NRECDCs		Non-NRECDCs		Diff.
	N	Mean	SD	Mean	SD	Mean	SD	Mean
Panel A. Key variables								
SWB	15,321	7.2164	2.3376	7.2000	0.0335	7.2239	0.0229	0.0239
NRECDC	21,620	0.2954	0.4562					
Panel B. Household head and household characteristics								
Age	21,620	51.1985	13.4287	50.9580	0.1678	51.2993	0.1089	0.3412**
Age2	21,620	28.0161	13.6890	27.7649	0.1706	28.1213	0.1111	0.3564**
Gender	21,620	0.5733	0.4946	0.6212	0.0061	0.5532	0.0040	−0.0680***
Marriage	21,516	0.8659	0.3408	0.8580	0.0044	0.8692	0.0027	0.0112**
Education	21,263	5.2350	4.4169	5.2147	0.0572	5.2434	0.0357	0.0288
Self-rated health	21,573	3.0027	1.3472	3.1050	0.0172	2.9600	0.0108	−0.1450***
Family structure	21,620	0.5839	0.3238	0.5588	0.0040	0.5944	0.0026	0.0357***
Major event	21,607	0.1539	0.3609	0.1375	0.0043	0.1608	0.0030	0.0234***
Additional properties	21,606	0.1358	0.3426	0.1361	0.0043	0.1357	0.0028	−0.0004
Credit availability	21,602	0.1568	0.3636	0.2162	0.0052	0.1319	0.0027	−0.0843***
Household deposit	21,620	2.2892	7.4311	2.1540	0.0840	2.3459	0.0625	0.1919**

Notes: The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

5. Results

5.1. Baseline results

Table 2 presents the baseline results, probing the causal association between e-commerce development and the SWB of rural residents in China. In the first columns, we only controlled for year and county FEs. We found that the estimated coefficients on *NRECDC* are statistically significantly positive, suggesting that e-commerce development is likely to contribute to the SWB improvement of Chinese rural residents. In column (2), we further controlled for household head characteristics, household characteristics, and interactions between the 2014 county characteristics mentioned above and year dummies to mitigate NRECDC pilot selectivity bias. Not surprisingly, our DID estimate remains significantly positive at the 1% level.

According to our estimates from our preferred specification in column (2), NRECDC policy implementations increased the happiness score of rural residents by 0.18 points. Correspondingly, rural e-commerce development led to an increase in residents' SWB by about 2.40%, as the average happiness score is 7.22 in the full sample. The magnitude of the impact of e-commerce development on the happiness of rural residents estimated in this research is comparable to the estimates from existing studies on digital well-being. For example, estimates from Lohmann (2015) suggested that Internet use increased residents' SWB by about 4.10% using multi-country data. In Chinese cases, results in Lu and Kandilov (2021) suggested that the effect of mobile internet use in improving the happiness score is about 2.04%; Lei et al. (2023) found that digital financial inclusion increased residents' life satisfaction by 3.04%. Thus, we found a statistically significant and economically large effect of e-commerce development on rural people's SWB.

In column (3), the regression model is adjusted to an Oprobit model for robustness checks (e.g., Ferrer-i-Carbonell & Frijters, 2004; Zhao & Guo, 2022; Zheng & Ma, 2021). Despite this alteration, the core inference remains unscathed, reiterating the reliability of the key findings. In Table A5 of Appendix A, results show when the outcomes of *SWB* are equal to other integers. These results still support our conclusion that e-commerce development is an effective way to enhance rural happiness in China.

5.2. Applicability tests

Applying the DID method necessitates adherence to certain assumptions. Non-compliance can cause significant deviation in estimates from the actual causal effect. This section discusses the key assumptions of the DID approach, specifically the Parallel Trends Assumption (PTA) and potential spillover effects.

5.2.1. PTA tests

The baseline DID results confirmed that NRECDC policy implementations effectively enhance people's happiness in rural China. However, one of the prerequisites for DID design is to satisfy the PTA, meaning that before policy implementations, the treated group and control group would have followed the same trend over time. We adopted an event study analysis method for the staggered DID and the model is set up as follows:

$$Happiness_{i,c,t} = \beta_0 + \sum_{j\geq -4, j\neq 0}^{j\leq 5} \beta_j NRECDC_{c,j} + \beta_n X_{i,c,t} + \mu_c + \delta_t + \gamma_t H_c + \varepsilon_{i,c,t} \quad (2)$$

In Eq. (2), $NRECDC_{c,j}$ is assigned to 1 if the county is in the jth year after the NRECDC policy implementations and 0 otherwise. To mitigate the risk of multicollinearity and ensure robustness in our estimations, the current year of policy enactment (period 0) serves as the referential baseline in conducting parallel trend tests. This approach is integral for the precision of PTA tests within our model, as underscored by the pertinent literature (Leng, 2022; Yu et al., 2021). Moreover, recognizing the limited sample size associated with the treatment group in the fifth and sixth years before the NRECDC policy implementations, we have opted to aggregate these observations into a consolidated period, labeled as −4+ (Niu et al., 2022).

Fig. B1 in Appendix B provides a representation of the coefficients along with their 95% confidence intervals for each period. We found that none of the estimated coefficients were statistically significant before the NRECDC policy implementation. This observation aligns with the PTA, suggesting that the SWB of rural residents living in NRECDCs and non-NERCDCs followed similar trends before policy implementation. That is, using a staggered DID method is appropriate in this empirical design. Moreover, after the introduction of the NRECDC policy, the coefficients became notably significant and showed an increase in magnitude. This indicates a growing disparity in the happiness levels of rural residents between areas with and without NRECDCs following policy implementations.

5.2.2. Spillover effect tests

The DID method relies on another key assumption known as the SUTVA. This assumption posits that the treatment status of one unit is independent of another, meaning that the exposure of the treated group to the policy shock does not impact any units in the control group (Callaway & Sant'Anna, 2021). In simpler terms, the pilot policy should ideally have no or negligible spillover effects (Lu et al., 2019). In this study, two methods are employed to check the potential spillover effects of the NRECDC policy, following the approaches used in previous studies (Lu et al., 2019; Marein, 2023). First, a dummy variable named *NRECDC_Near* is created, which is assigned a value of 1 if a county shares a border with an NRECDC, and 0 otherwise. Second, the distance between non-NRECDCs and the nearest NRECDC is calculated, and we construct a set of dummy variables *NRECDC_Dis* according to different distance windows in each batch. With the help of the idea of placebo tests, it is possible to determine whether the NRECDC policy has any spillover effects (Lu et al., 2019).

Table 2
Baseline results: DID methods.

	OLS estimates		Oprobit estimates
	(1)	(2)	(3)
NRECDC	0.2036***	0.1800***	0.0236**
	(0.0648)	(0.0694)	(0.0096)
Age		−0.0900***	−0.0119***
		(0.0107)	(0.0015)
Age2		0.1034***	0.0138***
		(0.0107)	(0.0015)
Gender		−0.1735***	−0.0241***
		(0.0451)	(0.0062)
Marriage		0.8991***	0.1184***
		(0.0693)	(0.0090)
Education		0.0232***	0.0018**
		(0.0059)	(0.0008)
Self-rated health		0.3292***	0.0458***
		(0.0162)	(0.0022)
Family structure		−0.0303	−0.0032
		(0.0697)	(0.0096)
Major event		0.0130	0.0031
		(0.0497)	(0.0068)
Additional properties		0.1362**	0.0176**
		(0.0555)	(0.0078)
Credit availability		−0.1014*	−0.0139*
		(0.0562)	(0.0076)
Household deposit		0.0072**	0.0009**
		(0.0029)	(0.0004)
2014 county characteristics×year FE	No	Yes	Yes
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
R^2	0.0471	0.1068	0.0278
N	15,314	15,073	15,073

Notes: The explained variable is *SWB* in all columns. County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses. Column (3) reports the marginal effect when *SWB* is equal to 10.

The results in Table B1 in Appendix B show that the coefficient on *NRECDC_Near* is insignificant in column (1), suggesting that the NRECDC policy did not appear to have significant spillover effects on neighboring counties. Moreover, in columns (2)–(6), the coefficients of *NRECDC_Dis* are also insignificant after considering non-NRECDCs at different distances, which further reinforces the conclusion that policy spillovers are negligible in our empirical design. In sum, these results provide supportive evidence that the DID method used in this study satisfies the SUTVA, indicating that our DID design is quite reasonable.

5.3. Robustness tests

5.3.1. Bacon decomposition

In the causal interpretation of the Two-Way Fixed-Effect (TWFE) DID estimates, it is essential to ensure the control group remains unchanged after a policy shock. However, meeting this condition can be challenging in a framework of staggered DID, as the staggered DID estimates may be subject to potential estimation bias due to heterogeneity of treatment effects or negative weighting issues (De Chaisemartin & d'Haultfoeuille, 2020; Goodman-Bacon, 2021). To address this concern, our research employed the Bacon decomposition to decompose the multi-period DID coefficients into a series of 2 × 2 DID coefficients. As shown in Fig. B2 and Table B2 in Appendix B, this decomposition revealed a negative weighting problem in the estimation results, with approximately 2.10% of the treatment effects being inappropriate. However, the majority of the treatment effects (over 70%) were deemed appropriate. Thus, although treatment effect heterogeneity exists in our staggered DID estimates, it is not severe, suggesting that our baseline results are reliable.

5.3.2. Eliminating disturbing policies

During the sample period, some other related policies were introduced, which may result in our DID estimates not capturing the policy effects of NRECDC on rural residents' SWB (Chen et al., 2020). We aimed to consider and exclude several types of policies that are directly related to the development of rural e-commerce. The first is the National E-commerce Demonstration City (NEDC), which is a policy centered on the development of e-commerce established at the prefecture-level city level (Ma et al., 2018). From 2009 to 2017, 70 cities in China became NEDCs in three batches. In column (1) in Table B3 of Appendix B, we found that the DID estimated coefficient of *NRECDC* remains almost unchanged after controlling for the effect of this important disruptive policy. Thus, the positive effect of this rural e-commerce development-centered policy on people's SWB is not shocked by the NEDC policy implementations.

Second, we focused on several policies related to rural digitization, as previous research highlighted the dual impact of digital

transformation on e-commerce use and residents' happiness (e.g., Lei et al., 2023; Tran, 2021; Zhan & Zhou, 2018; Zhao et al., 2020). These policies include the Rural Information to Villages and Households Project (RIVHP) implemented by the Ministry of Rural Affairs of China in 2016, and the Broadband China Strategy (BCS) and Smart City Construction (SCC) policy announced by the Ministry of Industry and Information Technology of China. In columns (2)–(4) of Table B3, we found that the positive effect of e-commerce development on rural residents' SWB was not challenged after full consideration of these digitization-related policies.

Last, we were concerned about the potential impact of the Returning Entrepreneurship Demonstration County (BEDC) policy due to the link between entrepreneurial behaviors and happiness (Batjargal, 2007; Naudé et al., 2014; Tan & Li, 2022). This policy has been implemented by the NDRC since 2016 in three batches in over 340 counties. The results in column (5) of Table B3 show that our baseline results remain unchanged. Further, we controlled for all these policies in column (6) and the DID estimate for *NRECDC* is still significantly positive. Taken together, these results tell us that the influence of these disruptive policies on our baseline results is limited.

5.3.3. Placebo tests

To ascertain if unseen random factors could skew our DID results, we executed placebo tests by constructing an artificial treatment group, as documented in studies by Niu et al. (2022) and Zhou et al. (2023). In this approach, we randomly chose a fraction of policy pilot districts from the complete sample, designating them as the "simulated" treated group. The rest of the sample served as the "simulated" control group. This ensured a consistent number of treatment group observations at the micro-household level, aligning with our foundational regression analysis. Fig. B3 in Appendix B shows that after repeating this randomization process 1000 times, the mean of the regression coefficients of these dummy policy variables is close to 0, roughly following a normal distribution. Simultaneously, it is far from the estimate in the baseline regression. These outcomes rule out the potential impact of unobservable chance factors, indirectly suggesting that the increase in rural residents' SWB is genuinely due to the implementation of the NRECDC policy.

5.3.4. PSM-DID method

As mentioned earlier, to further overcome the pilot selectivity bias, we conducted robustness testing through the PSM-DID method by selecting five aforementioned county-level characteristic variables as predictor variables (Abadie & Imbens, 2006; Zhou et al., 2023). The outcomes, delineated across various matching techniques such as Kernel matching, Radius matching, Nearest neighbor matching, and Local linear regression matching, are delineated in columns (1)–(4) within Panel A of Table B4 of Appendix B. We can see that the estimate for *NRECDC* is still significantly positive, indicating that the causal relationship between rural e-commerce development and residents' SWB is not affected by the further exclusion of policy selectivity for mitigation. Moreover, extending the scope of matching variables, we incorporated all significant county-level variables from column (3) of Table A2 based on *t*-tests.[10] The resulting estimates, as portrayed in Panel B of Table B4, further affirm the robustness of our prior findings.

It is imperative to note that the assumption of common support is one of the prerequisites for employing the PSM method (Caliendo & Kopeinig, 2008). Taking the Kernel matching in Panel B as an illustrative example, our study presents a propensity score common support domain bar graph (as depicted in Fig. B4 of Appendix B), visually demonstrating that the majority of observational values fall within the common support range, with only a negligible portion of observations outside this range.[11] Furthermore, Fig. B5 of Appendix B indicates a significant convergence trend in the kernel density distributions of the experimental and control groups post-matching, suggesting a close alignment in the probability distributions of propensity score values between the two groups. Additionally, Fig. B5 underscores a substantial area of common support below the overlapping range of the two kernel density lines, signifying ample data for matching and passing the common support test, indicative of the high quality of data matching.

Another crucial assumption of the PSM approach is the balance assumption, necessitating balance tests. The results of specific covariate matching quality tests are presented in Table B5 of Appendix B. Analysis reveals that prior to matching, significant differences exist in the means of most covariates between the treatment and control groups. However, post propensity score matching, the distribution of all covariates between the treatment and control groups becomes balanced, with the majority exhibiting a reduction in bias of over 90%, and t-test *p*-values markedly exceeding 10%. Furthermore, Fig. B6 of Appendix B demonstrates that the majority of covariates pass the balance test post-matching, with standard deviations approaching zero, indicative of achieving the requisite balance.

5.3.5. Alternative specifications

We now performed robustness checks by using alternative specifications in Table B6 of Appendix B. First, we considered the potential impact of expected effects. We recognized that there are certain selection criteria for the NRECDCs, so government departments in some counties that meet these criteria may have different expectations and preparations in response. These actions are likely to affect local rural residents' SWB; thus, such expectations may bias our DID estimates. Thus, we included in the regressions a dummy term *Pre1* for the year before being an NRECDC to exclude expected effects. Not surprisingly, the coefficient on *NRECDC* is significantly positive while the coefficient on *Pre1* is not significant in line (1). These estimation results by adopting staggered DID approaches imply that the happiness effect of rural e-commerce development is unlikely to be affected by the expectation effect.

[10] This expanded set of matching variables now additionally comprises *Financial expenditures*, *Internet infrastructure*, *Social consumption*, and *Planted area*.

[11] In accordance with other alternative matching methods, we rigorously tested both the common support and the balance assumptions, all of which successfully passed the tests.

Second, we incorporated additional control variables for robustness checks. On one hand, we controlled for other county-level variables, as the NRECDC is implemented at the county level. These variables cover *GDP* per capita, *Fixed-asset Investment*, *Financial Development*, *Government,* and *Human capital*.[12] As shown in line (2), we found that the DID estimate is similar to the baseline results in Table 4. On the other hand, we attempted to control for more FEs (Chen et al., 2020; Liu et al., 2023). In line (3), we replaced the county FE with the household FE, controlling for characteristics that do not change over time at the household level. In line (4), we further included a high-dimensional FE that captures the characterization of the provincial level over the years. These DID estimates tell us that our baseline results are satisfactory for tighter controls.

Third, we excluded some special conditions. We first considered the potential impact of COVID-19 on SWB (Wang, Wang, et al., 2021; Zhou et al., 2021), even though we have controlled for year FE. However, as a public health crisis event with huge implications, we cannot completely ignore the interference it causes in our estimates. Therefore, in line (5), we excluded the data from the 2020 wave, finding that the results are also reliable. We then excluded the 2021 NRECDCs because the NRECDC list is still being updated. However, due to our data availability, we were unable to observe the happiness effect of the NRECDC policy after 2020. Thus, these prospective NRECDCs may not be qualified control groups. In line (6), after excluding the rural sample from these counties, we discovered that this change has little impact on the main specification.

Last, our data spans from 2014 to 2020, determining the policy effects for the latest three batches of pilot counties (2018–2020) can be challenging, as they were implemented for a relatively short period. Consequently, we focused on the initial three batches of NRECDC from 2015 to 2017. The DID estimate shown in line (7) aligns closely with our primary results, albeit with a slightly increased magnitude. Moreover, in line (8), we redefined *SWB* as a dummy variable (Zhao et al., 2020; Zhao et al., 2024). Specifically, it was assigned to 1 if the initial SWB score is >5, and 0 otherwise. We can clearly see that the DID estimate remains unchanged. Additionally, in line (9), we engaged with the work of Matsuura and Ma (2022) by adopting subjective life satisfaction as a surrogate measure for SWB, offering an alternative lens through which the efficacy of rural e-commerce initiatives can be assessed. The DID results show that the effect of rural e-commerce development in promoting happiness is rather insensitive to the way the explained variable is defined.

5.4. Mechanism tests

Given that our baseline results and robustness tests suggest that rural e-commerce significantly enhances people's SWB, we further explored multiple potential channels for this positive effect. As analyzed in Section 3, these mechanisms include absolute material welfare enhancement, relative disparity reduction, and social capital accumulation.

5.4.1. Absolute material welfare enhancement

We theorized that rural e-commerce development may enhance residents' SWB by increasing their absolute material welfare. Based on the practice of previous studies, we measured material welfare at both consumption and income levels (e.g., Luo et al., 2022; Zhao et al., 2024; Zhou et al., 2023). According to the CFPS questionnaire, absolute material well-being consists of five variables, namely *Total income*, *Operational income*, *Wage income*, *Total consumption*, and *Enjoyable consumption*.[13] The definitions of these mechanism variables are available in Panel A in Table C1 of Appendix C. The results for mechanism tests related to material welfare are reported in Table 3. In terms of household absolute income in columns (1)–(3), NRECDC policy implementations have led to an increase in the absolute income of rural residents, especially operational income.

These estimates imply that rural household income growth is likely to come mainly from a boost in operational income. Consistent with previous literature, our findings emphasized how rural households as entrepreneurs or sellers benefit from the development of e-commerce (Li & Qin, 2022; Sebora et al., 2009; Wang, Miao, et al., 2021). Moreover, from a consumption perspective in columns (4)–(5), we could see that the NRECDC policy has not only raised the total consumption level of rural households but has also improved their consumption structure. This change is manifested in the increase in enjoyment consumption. Thus, here we were concerned with the absolute material welfare-enhancing effect for rural households as consumers, with the development of e-commerce. This aligns with findings from earlier research that observed comparable outcomes (Fan et al., 2018; Jo et al., 2019; Tran, 2021).

5.4.2. Relative disparity reduction

As emphasized in Section 3, e-commerce development may also contribute to relative disparity reduction, thereby enhancing rural residents' SWB. This reduction of relative disparities may be reflected in social status, income, and consumption. Based on previous research and the CFPS questionnaire (Wang et al., 2024; Zou et al., 2020), we selected four variables to measure relative material welfare, including *Relative income*, *Relative social status*, *Income inequality*, and *Consumption inequality*.[14] Among them, the first two are self-reported by respondents, reflecting relative disparities in individual subjective perceptions; the last two are based on the Kakwani index measurements, which are relative disparities at the objective level. These mechanism variables related to relative material

[12] *Fixed-asset Investment* is represented as the ratio of total investment in fixed assets to GDP. *Financial Development* is depicted as the proportion of loan balances from financial institutions to GDP. *Government* is derived from the general budget expenditure as a percentage of GDP. *Human capital* is quantified as the number of college students for every 10,000 persons.

[13] Enjoyment consumption encompasses expenditures on cultural and entertainment activities, travel, healthcare, beauty treatments, dining out, and durable goods.

[14] *Income inequality* and *Consumption inequality* are measured by the Kakwani index (Zhou et al., 2023), which ranges from 0 to 1. A value closer to 1 indicates a more pronounced deprivation in income or consumption.

Table 3
Absolute welfare enhancement.

	Total income	Operational income	Wage income	Total consumption	Enjoyable consumption
	(1)	(2)	(3)	(4)	(5)
NRECDC	0.0578***	0.1904***	0.0154	0.1242***	0.0075**
	(0.0212)	(0.0768)	(0.0816)	(0.0324)	(0.0035)
Control variables	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
2014 county characteristics×year FE	Yes	Yes	Yes	Yes	Yes
N	20,778	21,072	21,101	10,032	11,670

Notes: County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

welfare are defined as shown in Panel B in Table C1 of Appendix C.

Table 4 presents the mechanism test results for relative disparity reduction. The DID estimates for *NRECDC* are all statistically significant and positive in columns (1) and (4) and negative in columns (2) and (3), indicating the potential that e-commerce development exhibits in improving relative social status and mitigating rural income and consumption inequality. Taking column (1) as an example, NRECDC policy implementations increased by about 0.44 points in the relative income of rural residents in China. These results are similar to the role of e-commerce development in mitigating inequality as highlighted in the previous literature (Li et al., 2021; Liu & Zhou, 2023). From a micro perspective, we confirmed that e-commerce development has indeed improved the relative welfare of rural households in China. Given the close link between relative disparity and SWB, the above results strongly suggest that reducing relative disparity may be one of the key pathways through which rural e-commerce development enhances residents' happiness.

5.4.3. Social capital accumulation

We further validated the social capital mechanism. Combining theoretical analysis, existing research, and data availability, we selected three types of mechanism variables. Social networks are measured by *Relatives' expenses* and *Interpersonal relationships* (Batjargal, 2007; Chen et al., 2023; Kwon & Arenius, 2010). Two variables act as proxies for social interactions, and they are *Communication expenditures* and *Organizational membership*. Communication expenditures are effective proxies for measuring social interactions in rural China because they quantify the extent of individuals' connectivity and engagement with digital platforms, which are integral to both social interaction and economic participation in the context of e-commerce development (Chen et al., 2024; Li et al., 2024; Meng et al., 2023). Meanwhile, organizational membership underscores the formal engagement in village groups, clubs, or associations, which facilitates shared interests, collective action, and social support, further embedding individuals in the social fabric of their rural setting (Appau et al., 2020; Li et al., 2024; Yang & Zhang, 2024). *Social trust* is reflected by how much respondents trust other people (Guiso et al., 2004). The definitions of these mechanism variables are detailed in Panel C in Table C1 of Appendix C.

The estimation results are shown in Table 5. We could see that although the coefficient of *NRECDC* is insignificant in column (1), it is significantly positive in column (2), partially supporting our view that rural e-commerce development contributes to the expansion of household social networks in China. In the last three columns, we found that NRECDC policy implementations improved social interactions and increased social trust among Chinese rural households. These findings are similar to existing literature, emphasizing how e-commerce as a digital tool is effective in promoting rural household social capital accumulation (e.g., Martinsons, 2008; Qi et al., 2019; Wang, Miao, et al., 2021). In sum, we confirmed the validity of the mechanism of social capital accumulation, given the positive role of social capital in promoting people's SWB (Bartolini & Sarracino, 2015; Helliwell, 2006; Ram, 2010).

5.5. Digital dividend or digital divide in e-commerce

As discussed earlier, there is no consensus in the existing literature on whether e-commerce can deliver digital dividends to more vulnerable groups. Potential dilemmas stem from technological barriers to e-commerce use and self-digital exclusion (Xia & Lu, 2008; Yoon, 2009). Therefore, from the perspective of SWB, we sought to explore whether NRECDC policy implementations have a higher marginal effect on the happiness enhancement of rural disadvantaged groups in China.

We attempted to determine the digital dividend and digital divide effects in e-commerce in terms of individual characteristics, including age, education, and health.[15] In the previous literature, the elderly, less educated, and relatively unhealthy individuals were

[15] Synthesizing the sample and previous literature (e.g., Hu et al., 2021; Nie et al., 2017; Zhao et al., 2022), the grouping criteria were as follows. For age categorization, the terms "young," "middle," and "old" signify that the age is below 50, between 50 and 60, and over 60 years old, respectively. In terms of education, "relatively low," "indifferent," and "relatively high" correspond to respondents having <6 years, between 6 and 8 years, and 9 years or more of education, respectively. Regarding self-rated health classifications, "worse," "general," and "better" represent the health scores being 5 and 4, 3, and 2 and 1, respectively.

Table 4
Relative disparity reduction.

	Relative income	Income inequality	Consumption inequality	Relative social status
	(1)	(2)	(3)	(4)
NRECDC	0.4403**	−0.0216***	−0.0233***	0.0605**
	(0.2066)	(0.0057)	(0.0083)	(0.0275)
Control variables	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
2014 county characteristics×year FE	Yes	Yes	Yes	Yes
N	19,456	20,354	10,166	20,157

Notes: County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

Table 5
Social capital accumulation.

	Relatives' expenses	Interpersonal relationship	Organizational membership	Communications expenses	Social trust
	(1)	(2)	(3)	(4)	(5)
NRECDC	−0.0261	0.1732***	−0.0026	0.2141**	0.0410**
	(0.0482)	(0.0653)	(0.0093)	(0.0888)	(0.0171)
Control variables	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
2014 county characteristics×year FE	Yes	Yes	Yes	Yes	Yes
N	19,531	13,958	19,789	20,874	13,025

Notes: County characteristics include GDP, Revolutionary area, Topographical relief, Capital distance, and Port distance. The control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

viewed as relatively disadvantaged groups with a greater likelihood of facing the digital divide (Chen et al., 2021; Lu et al., 2024; Zhao et al., 2022). In Table 6, we found that these DID estimates for *NRECDC* are insignificant for all these relatively vulnerable groups, but are instead significantly positive for young and middle-aged, more educated, and healthier groups. In conclusion, in terms of happiness, our estimates do not support the inclusive effect of e-commerce as a digital tool in rural China, which is at variance with findings in the previous literature (e.g., Chen et al., 2021; Li & Qin, 2022). Instead, our results recognize the potential digital divide effect that prevents some disadvantaged groups from enhancing their SWB through rural e-commerce development.

Furthermore, to explore whether technological barriers or self-digital exclusion play a dominant role in the digital divide effect of e-commerce development in rural China, we selected two variables according to the CFPS questionnaire, namely *Internet use* and the *Importance of the Internet*.[16] We considered that Internet users have a lower technological barrier to using e-commerce, at least as online consumers (Jo et al., 2019). Also, those respondents who viewed the Internet as more important were likely to have less rejection of the Internet and Internet-based digital tools. In Table 7, we restricted the sample to relatively disadvantaged groups who are shown to face the digital divide in Table 6 and then constructed the interaction terms for *NRECDC* and these two variables, separately. The positive coefficients for these interaction terms across all columns suggest that, although certain groups may be considered relatively vulnerable, their SWB may improve with e-commerce development, provided they surmount technological obstacles and digital exclusion. Among them, the coefficients are statistically significant in columns (1), (4), and (6). These findings indicate that for the rural elderly, the primary obstacle appears to be the digital technology barrier. Meanwhile, for those who are educationally and health-wise disadvantaged in rural areas, the predominant issue is self-digital exclusion.

Taken together, our estimates favored the digital divide effect, indicating that the happiness-enhancing effect of e-commerce on rural disadvantaged groups is likely to be limited by technological thresholds and self-digital exclusion. Although prior literature considered the positive effects of e-commerce development on raising farmers' incomes and alleviating poverty in rural China (Li et al., 2019; Li & Qin, 2022), we were unable to provide supportive evidence for this inclusive feature, at least from the perspective of SWB. More importantly, our findings emphasized the benefits of enhancing digital outreaches and lowering digital technology barriers in enhancing the welfare of rural disadvantaged groups.

[16] *Internet use* is a binary variable which is set to 1 if the household head used the Internet and 0 otherwise. The variable *Importance of the Internet* ranges from 1 to 5, where a higher value signifies greater significance placed on the Internet for accessing information (self-reported).

Table 6
Digital dividend or digital divide: individual heterogeneity.

Panel A. Age	Young	Middle	Old
	(1)	(2)	(3)
NRECDC	0.1750*	0.4116***	−0.0283
	(0.1064)	(0.1375)	(0.1416)
N	6635	4349	4085
Panel B. Education	Relatively low	Indifferent	Relatively high
	(4)	(5)	(6)
NRECDC	0.1678	0.1304	0.2192**
	(0.1628)	(0.1346)	(0.0956)
N	4355	4243	6469
Panel C. Self-rated health	Worse	General	Better
	(7)	(8)	(9)
NRECDC	0.0792	0.1408	0.4584***
	(0.1079)	(0.1630)	(0.1340)
N	6647	3163	5253

Notes: The explained variable is *SWB* for all regressions in four Panels. Baseline controls, county FE, year FE, and 2014 county characteristics×year FE are added in all regressions. Among them, county characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*; the baseline control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

Table 7
Digital thresholds and self-digital exclusion effects in e-commerce.

	Age		Education		Self-rated health	
	Old		Relatively low or indifferent		Worse or general	
	(1)	(2)	(3)	(4)	(5)	(6)
NRECDC	−0.0899	−0.1294	0.1242	0.0144	0.0389	−0.1069
	(0.1472)	(0.1598)	(0.1105)	(0.1182)	(0.1004)	(0.1086)
NRECDC×Internet usage	0.5382**		0.0275		0.0097	
	(0.2709)		(0.1492)		(0.1013)	
NRECDC×Importance of the Internet		0.2151		0.2781**		0.2141**
		(0.2044)		(0.1325)		(0.1047)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
2014 county characteristics×year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	3947	3695	8372	8016	9612	9354

Notes: The explained variable is *SWB* in all columns. *Internet usage* is controlled in columns (1), (3), and (5); *Importance of the Internet* is controlled in columns (2), (4), and (6). County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

6. Further studies

6.1. Further evidence for e-commerce use

In the above explorations, our identification strategy relies on the exposure of rural households to the e-commerce development policy, but exactly whether rural households use e-commerce and the extent to which e-commerce tools are used has not yet been addressed. After all, we do not guarantee that all rural households in the pilot counties become e-commerce users.

We measured e-commerce use in two ways. At the macro level, the Taobao village, a particular form of e-commerce organization, was widely used in studies of the welfare effect of e-commerce development in China (e.g., Lin et al., 2022; Liu & Zhou, 2023; Wang, Miao, et al., 2021; Zhou et al., 2021). Living in counties with Taobao villages, rural residents are more likely to be e-commerce users due to spillover effects (Liu et al., 2020; Zang et al., 2023). At the micro level, consistent with what we discussed earlier, we explored both consumers and operators in the e-commerce system. Based on the CFPS questionnaire, two variables, namely *Online shopping*

expenses and *Online business activities*,[17] were constructed. In Table 8, we set three interaction terms between *NRECDC* and e-commerce use variables, finding that these estimated coefficients all remain significantly positive in Panel A. The implementation of the NRECDC policy confers advantages to rural e-commerce users, subsequently enhancing their SWB.

Moreover, in the context of NRECDC policy implementations, we were also interested in whether relatively disadvantaged groups in rural areas improved their SWB through e-commerce use. Therefore, we set three dummy variables to represent these vulnerable groups in terms of age, education, and health.[18] In Panel B of Table 8, we found that the estimated coefficient on *NRECDC*×*Taobao village*×*Relatively low or indifferent education* is significantly positive in column (2), which suggests that rural residents with relatively low and moderate levels of education, may derive more happiness from NRECDC policy implementations if they live in counties with Taobao villages. Similarly, in Panels C and D of Table 8, we observed that compared to other groups, the elderly who spend more on online purchases and have online business behaviors are likely to be happier after the NRECDC policy is implemented. These findings indicate that despite the digital divide impacting some disadvantaged rural groups within e-commerce, the expansion of digital access and the advancement of e-commerce in rural China present opportunities for these groups to overcome such divides and derive increased SWB from e-commerce development policies.

6.2. Distinguishing levels of SWB

Given the distribution of SWB's inherent skewness, e-commerce development may affect the rural residents with different SWB scores, thereby engendering varied implications. To delve into this phenomenon, we drew inspiration from Binder and Coad (2011) and employed a quantile regression model for analysis. Specifically, our study focused on estimating SWB at the 10th, 30th, 50th, 70th, and 90th percentiles among rural residents, thereby encompassing a wide range of happiness levels from the lowest to the highest. According to the estimation results presented in columns (1) through (5) of Table 9, the core explanatory variable, *NRECDC*, exhibits significant positive coefficients only at the 10th and 30th quantiles. This indicates that the development of rural e-commerce is particularly beneficial in enhancing the SWB of individuals who initially have lower levels of happiness.

Rural e-commerce development improves access to goods and services that were previously unavailable or hard to come by. This enhanced access can lead to an improvement in the quality of life for individuals with lower levels of happiness, as it directly addresses some of the material shortages that can contribute to the decline in SWB (Couture et al., 2021; Tu & Bulte, 2010). Moreover, *E*-commerce platforms often foster village engagement and connectivity, offering social support networks that were previously lacking in these rural settings. For individuals in the lower happiness group, this increase in social interaction and the feeling of being connected to a larger village can significantly enhance one's sense of belonging, which are crucial components of SWB (Bartolini & Sarracino, 2015; Ram, 2010). These insights dovetail with the study's mechanism analyses, elucidating pathways for ameliorating relative disparities and fostering social capital accumulation. However, the findings also suggest that this significant positive effect of rural e-commerce development is not as pronounced in the groups with medium and high levels of happiness, which means that the marginal utility they gain from improving access through e-commerce may be lower.

6.3. The synergistic effect of digital finance and e-commerce development

In addition to the Internet base, digital financial development is also key to the success of e-commerce in China. On the one hand, funds in e-commerce transactions are mainly transacted through digital financial tools, such as mobile payment (Guo et al., 2022; Ma et al., 2018). Meanwhile, digital financial development has further accelerated the spread of e-commerce in China. On the other hand, digital finance also provides more financial support to e-commerce merchants and consumers, such as start-up capital and consumer credit (Sebora et al., 2009; Zhao et al., 2022). Especially for rural areas in China, the positive role of digital finance in promoting residents' welfare has been supported by a considerable body of literature (e.g., Lei et al., 2023; Liu et al., 2022; Zhao et al., 2024). Hence, it becomes intriguing to investigate if the combined influence of digital finance and e-commerce growth can amplify the SWB of rural residents in China.

Similar to previous literature (Lei et al., 2023; Niu et al., 2022), we measured digital finance using the county-level Digital Financial Inclusion Index compiled by Peking University and Ant Group, including four variables *Digital finance*, *Coverage breadth*, *Usage depth*, and *Digitization level*.[19] Consistent with prior research (Fan & Liang, 2023; Sun et al., 2023), we formulated interaction terms to quantify the synergistic effects. The results in Table 10 show that coefficients on all interaction terms are significantly positive, suggesting that the happiness effect on rural residents is more pronounced when the NRECDC policy is pursued in regions with higher levels of digital finance. Our findings are similar to those highlighted in the prior studies, confirming that digital finance is an important foundation for e-commerce development (Guo et al., 2022; Sebora et al., 2009). These estimates tell us that the synergistic effect of digital finance and e-commerce development further contributes to the SWB of rural residents in China.

[17] Online shopping expense refers to the total amount spent on online purchases over the past year. Meanwhile, online business activities denote how frequently the Internet is used for various commercial purposes, such as internet banking or making online purchases.

[18] This is consistent with Tables 6 and 7. *Old* is assigned a value of 1 if the respondent is elderly and 0 otherwise; For *Relatively low or indifferent education*, it is defined as 1 if the respondent has a relatively low or moderate level of education and 0 otherwise. *Worse or general health* is equal to 1 if respondent's self-rated health is low or general and 0 otherwise.

[19] More details can be found in the website: https://idf.pku.edu.cn/.

Table 8
Further evidence for e-commerce use.

	(1)	(2)	(3)
Panel A. E-commerce use and happiness			
NRECDC	0.1468**	−0.8795	0.1402
	(0.0739)	(0.6083)	(0.2270)
NRECDC×Taobao village	0.0750*		
	(0.0400)		
NRECDC×Online shopping expense		0.1190*	
		(0.0679)	
NRECDC×Online business activities			0.4489**
			(0.2410)
N	15,073	1156	2332
Panel B. Taobao village and digital divide			
NRECDC	0.1280	0.1749**	0.2908***
	(0.0788)	(0.0826)	(0.0955)
NRECDC×Taobao village×Old	−0.0344		
	(0.0853)		
NRECDC×Taobao village×Relatively low or indifferent education		0.1706**	
		(0.0789)	
NRECDC×Taobao village×Worse or general health			−0.0614
			(0.0864)
N	15,073	15,073	15,073
Panel C. Online shopping and digital divide			
NRECDC	−0.8319	−0.6367	−1.4295
	(0.6287)	(0.6292)	(1.3048)
NRECDC×Online shopping expense×Old	0.3490**		
	(0.1662)		
NRECDC×Online shopping expense×Relatively low or indifferent education		0.0998	
		(0.1694)	
NRECDC×Online shopping expense×Worse or general health			−0.1101
			(0.1882)
N	1156	1156	1156
Panel D. Online business and digital divide			
NRECDC	0.1282	0.1529	0.3575
	(0.2294)	(0.2346)	(0.2956)
NRECDC×Online business activities×Old	1.9407**		
	(0.7736)		
NRECDC×Online business activities×Relatively low or indifferent education		−0.8948	
		(0.7749)	
NRECDC×Online business activities×Worse or general health			0.6574
			(1.0265)
N	2332	2332	2332

Notes: The explained variable is *SWB* for all regressions in four Panels. Baseline controls, county FE, year FE, and 2014 county characteristics×year FE are added in all regressions. Among them, county characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*; the baseline control variables are consistent with column (2) of Table 2. *Taobao village*, *Online shopping expense*, and *Online business activities* are controlled for in columns (1)–(2) of Panel A, respectively. For all columns in Panels B—D, three dual interaction terms are also added. For example, in column (1) of Panel B, *Taobao village*, *Old*, *NRECDC×Taobao village*, *NRECDC×Old*, and *Taobao village×Old* are all controlled for. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

6.4. The role of entrepreneurial activities

A significant body of research contends that e-commerce fosters entrepreneurship (e.g., Li & Qin, 2022; Lin et al., 2022; Qi et al., 2019; Sebora et al., 2009), as it affords rural households a wealth of entrepreneurial opportunities and access to pertinent information. Moreover, entrepreneurial activities have the potential to enhance people's SWB, given the benefits that entrepreneurship may increase household income, expand social networks, and improve social status (Liu et al., 2022; Zhao et al., 2020). However, entrepreneurship is a high-risk business activity. The relatively weaker physical and human capital foundation, coupled with limited social support, subjects rural households to heightened uncertainty in their entrepreneurial endeavors (Batjargal, 2007; Démurger & Xu,

Table 9
Distinguishing levels of SWB.

	Q10	Q30	Q50	Q70	Q90
	(1)	(2)	(3)	(4)	(5)
NRECDC	0.1030***	0.1945*	0.0589	0.2332	0.0554
	(0.0315)	(0.1102)	(0.0647)	(0.1807)	(0.0503)
Control variables	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
2014 county characteristics×year FE	Yes	Yes	Yes	Yes	Yes
N	15,081	15,081	15,081	15,081	15,081

Notes: The explained variable is *SWB* for all regressions. County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

Table 10
The synergistic effect of digital finance.

	(1)	(2)	(3)	(4)
NRECDC	−1.1024*	−2.1589**	−0.4889	−1.0237
	(0.6224)	(0.9271)	(0.3619)	(0.6959)
NRECDC×Digital finance	0.0124**			
	(0.0060)			
NRECDC×Coverage breadth		0.0259**		
		(0.0102)		
NRECDC×Usage depth			0.0053*	
			(0.0029)	
NRECDC×Digitization level				0.0108*
				(0.0062)
Control variables	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
2014 county characteristics×year FE	Yes	Yes	Yes	Yes
N	14,561	14,561	14,561	14,561

Notes: The explained variable is *SWB* for all regressions. County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

2011; Tan & Li, 2022), potentially diminishing their happiness (Naudé et al., 2014). Therefore, it is uncertain whether rural e-commerce development can enhance residents' happiness through entrepreneurial activities.

Based on the CFPS data, we selected some variables related to household entrepreneurial activities in three dimensions: entrepreneurial behavior, business assets, and business profits.[20] In columns (1)–(3) of Table 11, we found that all DID estimated coefficients on *NRECDC* are insignificant, suggesting that this e-commerce development policy did not stimulate entrepreneurial passion among rural households. Unlike previous literature that focused on the entrepreneurial effects of Taobao villages, the NRECDC policy did not create an effective entrepreneurial ecology and scale organization (Liu et al., 2020; Qi et al., 2019; Zang et al., 2023), which may lead to less pronounced entrepreneurial effects. Nonetheless, as demonstrated in Table 4, we observed that e-commerce development boosts the operational income of rural households. These findings are not contradictory; they suggest that for rural households in China, e-commerce primarily promotes home-based business activities, such as small-scale online sales (Guo et al., 2022; Tang & Zhu, 2020), even though these ventures may not yet meet the scale associated with private enterprise establishments.

Furthermore, to further test the role of entrepreneurial activities in this framework of e-commerce happiness, we constructed interaction terms between *NRECDC* and these entrepreneurial activity variables in columns (4)–(6). Not surprisingly, none of the coefficients for these interaction terms are significant. Thus, given the policy orientation of NRECDC and the mixed relationship between entrepreneurial activities and SWB, these results pointed out that entrepreneurial activity is not an effective channel in which NRECDC policy enhances the SWB of Chinese rural residents.

[20] We assigned a value of 1 to *Entrepreneurial behavior* if a household operates a private business, and 0 otherwise. *Business assets* denotes the total assets owned by household businesses and *Business profits* refers to the net after-tax profits.

Table 11
The role of entrepreneurial activities.

	Entrepreneurial behavior	Business asset	Business profit	SWB		
	(1)	(2)	(3)	(4)	(5)	(6)
NRECDC	0.0041 (0.0055)	0.0117 (0.0139)	0.0134 (0.0535)	0.1880*** (0.0704)	0.1864*** (0.0701)	0.1893*** (0.0705)
NRECDC×Entrepreneurial behavior				−0.1257 (0.1526)		
NRECDC×Business asset					−0.0641 (0.0557)	
NRECDC×Business profit						−0.0179 (0.0154)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
2014 county characteristics×year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	21,103	21,103	21,062	15,073	15,073	15,046

Notes: County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

7. Conclusions and discussion

7.1. Main findings

With the advent of the digital economy, digital well-being has received increasing attention from scholars (e.g., Lei et al., 2023; Lu & Kandilov, 2021; Zhan & Zhou, 2018). However, as one of the most widely used digital tools, the happiness effect of e-commerce development is rarely mentioned. Especially in Chinese rural areas, e-commerce has been a considerable success and is deeply embedded in the daily production and lives of rural households (Chen et al., 2021; Li & Qin, 2022). Therefore, based on a quasi-natural experiment of NRECDC policy implementations, we attempted to explore the causal effect of e-commerce development and rural residents' SWB by using a staggered DID approach.

Based on nationally representative panel data, we found that the implementation of NRECDC policy increases the happiness scores of rural households by 2.40% on average, confirming the happiness effect of rural e-commerce development in rural China. A series of robustness tests provide supportive evidence for our staggered DID estimation results, such as PTA tests, spillover effect tests, Bacon decomposition, placebo tests, and alternative specifications. In mechanism tests, we revealed that the positive effect of e-commerce development on rural residents' SWB may be explained by absolute welfare enhancement, relative disparity reduction, and social capital accumulation, while definitively excluding the impact of entrepreneurial activities as a significant factor.

Moreover, our analysis reveals that the evolution of rural e-commerce significantly boosts the SWB of individuals who initially reported lower happiness levels. We further explored whether e-commerce development brought a digital dividend or a digital divide to Chinese rural households. The results show that e-commerce development fails to promote SWB for rural residents who are relatively disadvantaged in age, education, and health, but instead effectively enhances non-disadvantaged groups' happiness, suggesting a potential digital divide effect in e-commerce. Furthermore, we identified technological barriers and self-digital exclusion as the main causes of the digital divide in e-commerce. Fortunately, actively promoting e-commerce use in rural households may overcome this difficulty. Last, we confirmed that the synergistic effect of digital finance and e-commerce development contributes to rural residents' SWB in China.

7.2. Implications

7.2.1. Theoretical implications

There are several theoretical insights from our study. First, we enriched the field of research on digital well-being of recent interest (Lei et al., 2023; Lohmann, 2015; Zheng & Ma, 2021). Based on the case of e-commerce, we further demonstrated how this digital tool, which is widely used in China, is effective in improving the SWB of rural residents. Also, in contrast to previous literature focusing on e-commerce use (Zhan & Zhou, 2018; Zheng & Ma, 2021), we centered on a rural e-commerce development policy in China. Based on a staggered DID empirical framework, we provided more evidence for the happiness effects of e-commerce.

Second, in the ongoing discussion regarding whether e-commerce and similar digital tools contribute to a digital divide or enhance the digital dividend in rural areas, our findings lean towards the presence of a digital divide effect (Couture et al., 2021; Xia & Lu, 2008; Yin & Choi, 2022). This underscores that the advent of e-commerce may not uniformly boost SWB among some of the more vulnerable groups in rural China. Due to the technology bias effect (Akerman et al., 2015; Niu et al., 2022), those with digital skills are instead more likely to derive SWB from this e-commerce development policy. Moreover, we further confirmed the reasons for the current digital divide phenomenon in rural e-commerce in China, which provides a new direction for subsequent research. Thus, the phenomena of digital refugees and the digital divide in this digital age still deserve more academic attention, particularly in developing countries and backward regions.

Third, in the process of e-commerce development, the fundamental role of the Internet, has been mainly emphasized in the previous studies (Couture et al., 2021; Falk & Hagsten, 2015; Li & Ku, 2018), but they ignored the synergistic effects of digital finance. Our study adds to the body of research by affirming that the e-commerce development policy's positive impact on the SWB of rural residents is more significant in regions with a stronger digital financial foundation. E-commerce has quite complex operational mechanisms and internal operating systems. With the emergence of more and more digital tools, the synergy between e-commerce and other digital tools is equally worthy of scholarly attention.

7.2.2. Policy implications

Our study not only underscores the transformative potential of digital tools, exemplified by e-commerce, in enhancing the welfare of rural households but also highlights the critical need to address emerging digital divides. First, to bolster the effectiveness of the NRECDC policy, policymakers must initiate targeted measures aimed at enhancing digital inclusivity. Specifically, the formulation and implementation of comprehensive digital literacy programs tailored to the unique needs of rural areas are essential. These programs may focus on imparting essential skills required for navigating e-commerce platforms, understanding online payment systems, and leveraging digital tools for entrepreneurial endeavors. Concurrently, further increase the construction of digital infrastructure in rural areas and accelerate the process of full penetration of rural informatization.

Second, addressing technical barriers and self-exclusion requires a multifaceted approach. To address the challenge of technical barriers, the initiative should focus on simplifying e-commerce platforms. This can be achieved by developing new software solutions with intuitive user interfaces designed specifically for rural users who may have limited digital literacy. These solutions should prioritize ease of navigation, straightforward transaction processes, and local language support to ensure a seamless online shopping experience. Furthermore, possible measures also include active sensitization and training. These should not only inform residents about the opportunities e-commerce presents, such as access to broader markets and the ability to purchase goods and services at lower prices but also provide hands-on training sessions.

Third, the synergy between digital finance and e-commerce represents a significant opportunity to boost rural residents' SWB. By integrating digital finance platforms into the NRECDC policy framework, rural residents can benefit from lower interest rates, faster loan approvals, and more flexible repayment options compared to traditional banking systems. Additionally, there is a need to improve the digital financial literacy of rural residents. This includes understanding how to safely engage in online transactions, manage digital wallets, and conduct digital financing.

7.3. Limitations

Several limitations of this study deserve attention. First, similar to the previous literature (e.g., Brockmann et al., 2009; Ferrer-i-Carbonell & Frijters, 2004; Hu et al., 2021; Knight & Gunatilaka, 2010; Lu & Kandilov, 2021; Luo et al., 2021; Sun et al., 2016; Zhao et al., 2024; Zhou et al., 2023), our measure for SWB comes from a simple question in the CFPS, which inevitably introduces the subjectivity and potential selectivity of self-reporting. This challenge may not fully capture the multifaceted nature of happiness. Acknowledging this limitation, future research could explore additional, more objective measures of SWB to complement our findings. Second, despite our attempts to mitigate the self-selectivity of the NRECDC pilot as much as possible, such as controlling for potential pilot selection criteria characteristics and using the PSM method (Liu et al., 2023; Niu et al., 2022), we were still unable to completely rule out the potential for the selectivity of pilots to interfere with the estimation results (Angrist & Pischke, 2008). Third, although we tried to provide more direct evidence on the happiness effects of e-commerce development at the level of online shopping and online business activities, this information is not detailed and comprehensive, and there are even many missing observations. In subsequent studies, a special questionnaire designed for rural e-commerce development for further identification is expected to address this shortcoming. Last, some mechanism tests are incomplete due to CFPS limitations. For example, for social capital accumulation, we cannot fully capture the expansion of digital networks and the formation of digital trust brought about by rural e-commerce development. This limitation is expected to be addressed through designing and developing a set of questionnaires and data related to digital social networks for e-commerce.

Declaration of competing interest

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Data availability

The authors do not have permission to share data.

Acknowledgments

The authors acknowledge the support from the National Natural Science Foundation of China (No. 72304102; No.72373043), the National Social Science Fund of China (No. 23&ZD112), the Soft Science Program for Ministry of Agriculture and Rural Affairs of China (No. RKX2024030), and Philosophy and Social Science Program Discipline Co-construction Project of Guangdong Province (No. GD23XYJ68).

Appendix A

Table A1
Variable definitions.

Variables	Definitions
Panel A. Key variables	
SWB	Integers from 0 to 10, with higher numbers representing greater happiness
NRECDC	=1 if the county implemented the NRECDC policy, 0 otherwise
Panel B. Household head and household characteristics	
Age	household head age
Age2	Square of the age/100
Gender	Male = 1; female = 0
Marriage	Married =1; otherwise =0
Education	Years of education
Self-rated health	Ranging from 1 to 5, with higher numbers representing less health.
Family structure	Labor force (16–64 years old) as a proportion of total household population
Major event	=1 if the household has occurred some important events (such as weddings, funerals, and birth of children) within the past year, 0 otherwise
Additional properties	=1 if the household owns a second house, 0 otherwise
Credit availability	=1 if the household has borrowed from banks, 0 otherwise
Household deposit	Household cash deposits, 10,000, yuan
Panel C. County characteristics	
GDP	100-million-yuan, logarithm
Revolutionary area	=1 if the county is located in a revolutionary area, 0 otherwise
Topographical relief	The degree of topographic relief in the county
Capital distance	Distance between the county and the capital city of the province to which it belongs (km)
Port distance	Distance between the county and the nearest port (km)

Table A2
County characteristics between the control and treated groups.

	Control group	Treated group	T-test	OLS
	(1)	(2)	(3)	(4)
GDP	4.5110 [0.0081]	4.5466 [0.0137]	−0.0356**	0.0427*** (0.0073)
Revolutionary area	0.0729 [0.0019]	0.1647 [0.0062]	−0.0917***	0.0410*** (0.0042)
Topographical relief	0.8767 [0.0072]	1.2409 [0.0191]	−0.3642***	0.0132*** (0.0014)
Capital distance	132.093 [0.5327]	153.7267 [1.2584]	−21.6337***	0.0060** (0.0026)
Port distance	609.8086 [4.3578]	780.3655 [11.0650]	−170.557***	0.0527** (0.0240)
Financial expenditures	0.2957 [0.0026]	0.3532 [0.0029]	−0.0575***	−0.0100 (0.0095)
Internet infrastructure	9.3563 [0.2079]	10.4497 [0.4017]	−1.0934**	−0.0001 (0.0002)
Social consumption	1.1409 [0.0105]	0.8786 [0.0134]	0.2623***	−0.0036 (0.0029)
Loan balances	1.0924 [0.0168]	1.0832 [0.0196]	1.1754	−0.0032 (0.0020)
Planted area	62.2221 [0.3356]	68.4744 [0.8689]	−6.2523***	0.0006 (0.0005)

Notes: *Financial expenditures* refers to the general budget expenses. *Internet infrastructure* can be defined as the number of Internet users for every 10,000 individuals. *Social consumption* represents the per capita total retail sales of consumer goods. *Loan balances* indicates the year-end balance of loans obtained from financial institutions. *Planted area* denotes the total cultivated area for crops. Columns (1) and (2) report the means and standard deviation. Standard deviations are reported in the brackets. Column (3) shows the results of the T-test between control and treated groups. In column (4), the OLS model is used and the explained variable is *NRECDC*. The standard errors clustered at the county level are reported in parentheses. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Table A3
County characteristics and NRECDCs by batches.

	2015	2016	2017	2018	2019	2020
	(1)	(2)	(3)	(4)	(5)	(6)
GDP in2014	0.0022	−0.0072	−0.0217**	−0.0603***	−0.0566***	−0.0381***
	(0.0063)	(0.0083)	(0.0095)	(0.0098)	(0.0102)	(0.0105)
Revolutionary area	−0.0084	0.0002	0.0429***	0.0822***	0.0948***	0.0944***
	(0.0072)	(0.0095)	(0.0108)	(0.0112)	(0.0116)	(0.0120)
Topographical relief	0.0850***	0.1840***	0.2567***	0.3276***	0.3067***	0.2834***
	(0.0217)	(0.0286)	(0.0327)	(0.0339)	(0.0351)	(0.0363)
Capital distance	3.2088***	3.6795***	6.9953***	7.6936***	7.2704***	8.3213***
	(0.8297)	(1.0961)	(1.2496)	(1.2970)	(1.3425)	(1.3896)
Port distance	0.0001	0.5827***	0.5411***	0.5056***	0.2585	0.1292
	(0.1146)	(0.1514)	(0.1726)	(0.1791)	(0.1854)	(0.1919)

Notes: The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Robust standard errors are reported in parentheses.

Table A4
Tests for NRECDC pilot selectivity.

	(1)	(2)	(3)
Age	−0.0017***	−0.0016***	0.0001
	(0.0003)	(0.0003)	(0.0001)
Gender	0.0367***	0.0360***	−0.0060*
	(0.0059)	(0.0059)	(0.0033)
Marriage	−0.0023	−0.0033	−0.0002
	(0.0091)	(0.0091)	(0.0040)
Education	−0.0066***	−0.0062***	−9.85e-06
	(0.0008)	(0.0008)	(0.0004)
Self-rated health	0.0017	0.0022	0.0015
	(0.0021)	(0.0021)	(0.0012)
Family structure		0.0036	0.0049
		(0.0095)	(0.0051)
Major event		0.0019	0.0002
		(0.0066)	(0.0049)
Additional properties		−0.0030	0.0040
		(0.0080)	(0.0048)
Credit availability		0.0589***	0.0041
		(0.0083)	(0.0047)
Household deposit		−0.0023***	−0.0004
		(0.0006)	(0.0003)
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
2014 county characteristics×year FE	No	No	Yes

Notes: The explained variable is *NRECDC* in all columns. County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

Table A5
Marginal effects of the Oprobit estimates.

	$SWB = 0$	$SWB = 1$	$SWB = 2$	$SWB = 3$	$SWB = 4$
NRECDC	−0.0025**	−0.0014**	−0.0021**	−0.0035**	−0.0026**
	(0.0010)	(0.0006)	(0.0008)	(0.0014)	(0.0011)
	$SWB = 5$	$SWB = 6$	$SWB = 7$	$SWB = 8$	$SWB = 9$
	−0.0132**	−0.0033**	−0.0015**	0.0033**	0.0032**
	(0.0054)	(0.0014)	(0.0006)	(0.0013)	(0.0013)

Notes: All regressions add the baseline control variables as shown in column (2) of Table 2, county FE, year FE, and 2014 county characteristics×year FE. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

Appendix B

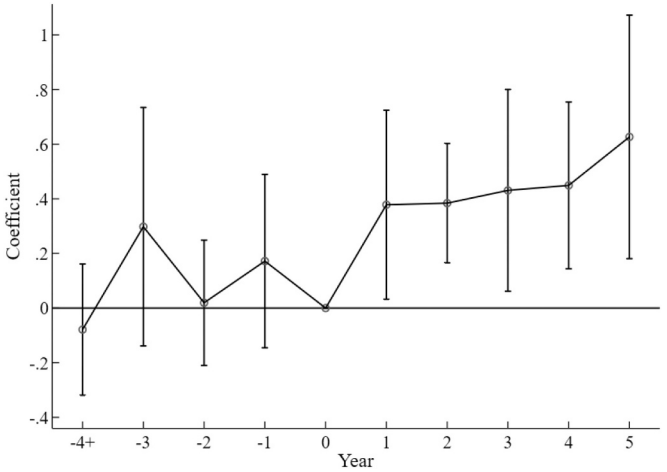


Fig. B1. Parallel trend tests

Table B1
Spillover effect tests.

	Full sample	≤25 km	≤50 km	≤75 km	≤100 km	≤125 km
	(1)	(2)	(3)	(4)	(5)	(6)
NRECDC_Near	0.1851					
	(0.1586)					
NRECDC_Dis		0.0574	0.0019	−0.0552	−0.0089	0.0547
		(0.1388)	(0.1147)	(0.0900)	(0.0839)	(0.0778)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
2014 county characteristics×year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	15,073	15,073	15,073	15,073	15,073	15,073

Notes: The explained variable is *SWB* in all columns. County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

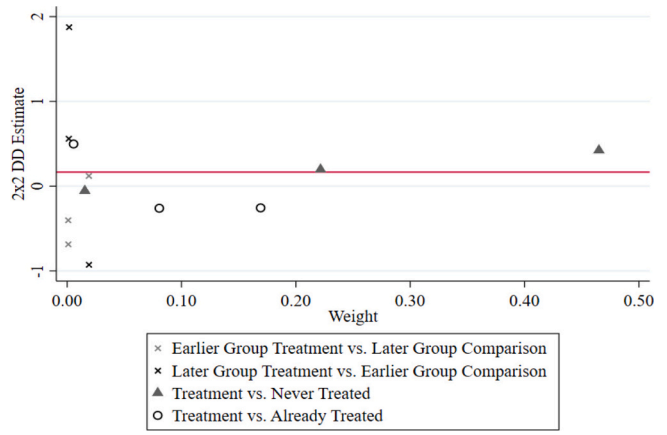


Fig. B2. Distribution of TWFE estimation bias weights.

Table B2
Bacon decomposition results.

	Estimates	Weight
Early Treatment vs. Later Control	0.067	0.021
Later Treatment vs. Early Control	−0.633	0.022

(*continued on next page*)

Table B2 (*continued*)

	Estimates	Weight
Treatment vs. Never treated	0.341	0.702
Treatment vs. Already treated	−0.241	0.255

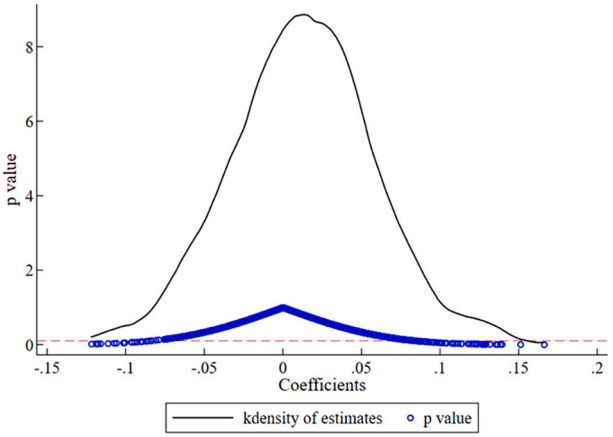


Fig. B3. Placebo tests.

Table B3
Disturbing policies.

	NEDC	RIVHP	BCS	SCC	BEDC	All policies
	(1)	(2)	(3)	(4)	(5)	(6)
NRECDC	0.1816***	0.1819***	0.1805***	0.1800***	0.1790***	0.1829***
	(0.0694)	(0.0695)	(0.0694)	(0.0694)	(0.0694)	(0.0694)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
2014 county characteristics×year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	15,073	15,073	15,073	15,073	15,073	15,073

Notes: The explained variable is *SWB* in all columns. County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

Table B4
PSM-DID Method.

	Kernel matching	Radius matching	Nearest neighbor matching	Local linear regression matching
	(1)	(2)	(3)	(4)
Panel A. Five county characteristics variables as predictors				
NRECDC	0.1729**	0.1591**	0.1678**	0.1765**
	(0.0717)	(0.0807)	(0.0702)	(0.0746)
Control variables	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
2014 county characteristics×year FE	Yes	Yes	Yes	Yes
N	13,554	11,229	14,353	12,680
Panel B. All county characteristics variables as predictors				
NRECDC	0.1495**	0.1476**	0.1716**	0.1453*
	(0.0747)	(0.0711)	(0.0706)	(0.0716)
Control variables	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
2014 county characteristics×year FE	Yes	Yes	Yes	Yes
N	12,680	14,236	14,458	14,064

Notes: The explained variable is *SWB* in all columns. County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

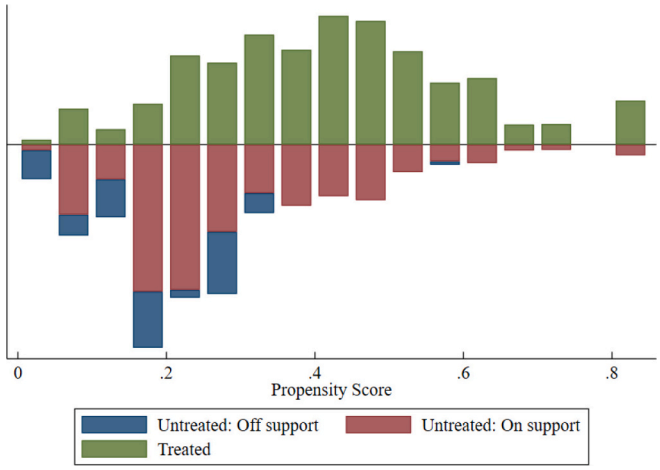


Fig. B4. Common range of values for the propensity score.

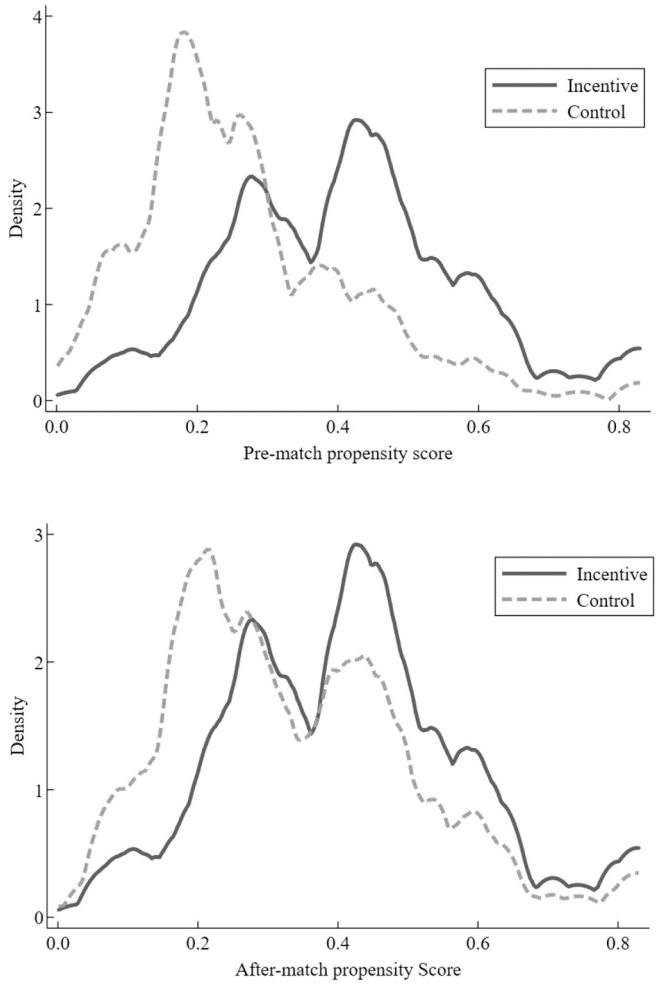


Fig. B5. Kernel density curves before and after matching.

Table B5
Results of balance test of control variables before and after matching.

Variable	Unmatched	Mean		%bias	%reduct |bias|	t-test	
	Matched	Treated	Control			t	p > |t|
GDP	U	101.06	179.51	−38.4	97.5	−21.49	0.000
	M	101.06	102.98	−0.9		−0.51	0.613
Revolutionary area	U	0.1418	0.0513	31.0	90.3	19.43	0.000
	M	0.1418	0.1330	3.0		1.25	0.211
Topographical relief	U	1.3745	0.7394	62.4	99.6	37.31	0.000
	M	1.3745	1.3721	0.2		0.11	0.916
Capital distance	U	156.73	130.68	38.2	98.7	21.74	0.000
	M	156.73	156.39	0.5		0.24	0.810
Port distance	U	674.64	487.72	47.8	95.6	27.47	0.000
	M	674.64	666.35	2.1		1.02	0.310
Financial expenditures	U	0.3716	0.2937	22.6	90.3	12.03	0.000
	M	0.3716	0.364	2.2		1.39	0.164
Internet infrastructure	U	13.399	8.4709	16.6	91.5	10.50	0.000
	M	13.399	13.837	−1.5		−0.59	0.553
Social consumption	U	0.6648	1.0255	−55.3	96.3	−29.91	0.000
	M	0.6648	0.6515	2.0		1.24	0.214
Planted area	U	73.118	74.56	−2.6	89.7	−1.51	0.132
	M	73.118	73.267	−0.3		−0.13	0.898

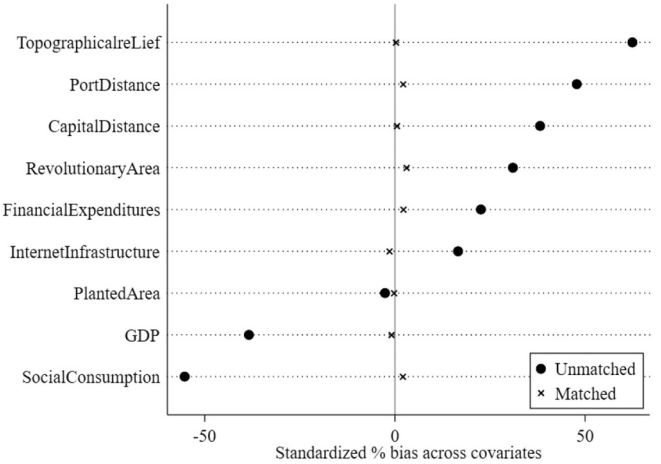


Fig. B6. Standardized bias of covariates before and after matching.

Table B6
Alternative specifications.

Line	Test methods	Coefficients on *NRECDC*	Standard errors	N
(1)	Expectation effect	0.2059***	0.0767	15,073
(2)	Other county variables	0.1601**	0.0712	15,073
(3)	Household FE	0.2153***	0.0681	13,100
(4)	Province-year FE	0.1587**	0.0803	15,071
(5)	Deleting the 2020 wave	0.2169**	0.0941	11,315
(6)	Deleting NRECDCs in 2021	0.1853***	0.0700	14,653
(7)	Considering the first three pilots	0.2871***	0.1015	10,807
(8)	Redefining *SWB* (dummy variable)	0.0285**	0.0137	15,073
(9)	Redefining *SWB* (life satisfaction)	0.0539**	0.0274	18,720

Notes: The explained variable is *SWB* in all columns. County characteristics include *GDP*, *Revolutionary area*, *Topographical relief*, *Capital distance*, and *Port distance*. The control variables are consistent with column (2) of Table 2. The significance levels of 1%, 5%, and 10% are denoted by ***, **, and *, respectively. Standard errors clustered at the household level are reported in parentheses.

Appendix C

Table C1
Definitions and descriptive statistics for mechanism variables.

Variables	Definitions	N	Mean	SD
Panel A. Absolute welfare				
Total income	Total household income (yuan, logarithm)	21,288	8.9795	1.0248
Operational income	Household business income per capita (yuan, logarithm)	21,588	5.6696	3.5357
Wage income	Household wage income per capita (yuan, logarithm)	21,616	2.4077	4.0814
Total consumption	Total household consumption (yuan, logarithm)	10,174	9.2243	0.8726
Enjoyable consumption	Enjoyable consumption/total household consumption	11,859	0.0652	0.0899
Panel B. Relative material disparity				
Relative income	Integers from 1 to 5 for self-assessment, with larger indicating a higher income level	19,790	3.4948	7.7830
Relative social status	Integers from 1 to 5 for self-assessment, with larger indicating a higher social status	20,513	3.1055	1.0947
Income inequality	Calculated using the Kakwani index based on per capita household income	20,835	0.4579	0.2616
Consumption inequality	Calculated using the Kakwani index based on per capita household consumption	10,306	0.4482	0.2304
Panel C. Social capital				
Relatives' expenses	Gift expenditure per capita (yuan, logarithm)	20,014	5.9419	1.9672
Interpersonal relationship	Integers from 0 to 10, with larger indicating better interpersonal relationship (self-related)	14,170	7.1955	1.9777
Communications expenses	Yearly household communications costs (yuan, logarithm)	21,378	3.6151	3.5431
Organizational membership	=1 if the respondent is a member of organizations such as the Communist Party of China, trade unions, religious groups, or individual labor associations, 0 otherwise	20,145	0.1520	0.3591
Social trust	=1 if respondents believed that most people could be trusted, 0 otherwise	13,204	0.5285	0.4992

References

Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica, 74*(1), 235–267.
Aghion, P., & Bolton, P. (1997). A theory of trickle-down growth and development. *Review of Economic Studies, 64*(2), 151–172.
Ahuvia, A. C. (2002). Individualism/collectivism and cultures of happiness: A theoretical conjecture on the relationship between consumption, culture and subjective well-being at the national level. *Journal of Happiness Studies, 3*, 23–36.
Akerman, A., Gaarder, I., & Mogstad, M. (2015). The skill complementarity of broadband internet. *Quarterly Journal of Economics, 130*(4), 1781–1824.
Angrist, J. D., & Pischke, J. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
Appau, S., Awaworyi Churchill, S., Smyth, R., & Zhang, Q. (2020). Social capital inequality and subjective wellbeing of older Chinese. *Social Indicators Research*, 1–23.
Asher, S., & Novosad, P. (2020). Rural roads and local economic development. *American Economic Review, 110*(3), 797–823.
Bao, T., Liang, B., & Riyanto, Y. E. (2021). Unpacking the negative welfare effect of social media: Evidence from a large scale nationally representative time-use survey in China. *China Economic Review, 69*, Article 101650.
Bartolini, S., & Sarracino, F. (2015). The dark side of Chinese growth: Declining social capital and well-being in times of economic boom. *World Development, 74*, 333–351.
Batjargal, B. (2007). Internet entrepreneurship: Social capital, human capital, and performance of internet ventures in China. *Research Policy, 36*(5), 605–618.
Binder, M., & Coad, A. (2011). From average Joe's happiness to miserable Jane and cheerful John: Using quantile regression to analyze the full subjective well-being distribution. *Journal of Economic Behavior & Organization, 79*(3), 275–290.
Brockmann, H., Delhey, J., Welzel, C., & Yuan, H. (2009). The China puzzle: Falling happiness in a rising economy. *Journal of Happiness Studies, 10*, 387–405.
Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys, 22*(1), 31–72.
Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics, 225*(2), 200–230.
Caporale, G. M., Georgellis, Y., Tsitsianis, N., & Yin, Y. P. (2009). Income and happiness across Europe: Do reference values matter? *Journal of Economic Psychology, 30*(1), 42–51.
Chen, B., Zeng, N., & Tam, K. P. (2024). Do social networks affect household financial vulnerability? Evidence from China. *Finance Research Letters, 59*, Article 104710.
Chen, P., Biao, M. A., & Zhang, C. (2021). Poverty alleviation through e-commerce: Village involvement and demonstration policies in rural China. *Journal of Integrative Agriculture, 20*(4), 998–1011.
Chen, S., Liu, C., Huang, J., & Zhi, H. (2023). Can social capital facilitate households' donations in rural China? *China Economic Review, 81*, Article 102034.
Chen, Y., Fan, Z., Gu, X., & Zhou, L. A. (2020). Arrival of young talent: The send-down movement and rural education in China. *American Economic Review, 110*(11), 3393–3430.
Chen, Y. J., Li, P., & Lu, Y. (2018). Career concerns and multitasking local bureaucrats: Evidence of a target-based performance evaluation system in China. *Journal of Development Economics, 133*, 84–101.
Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature, 46*(1), 95–144.
Clark, W. A., Yi, D., & Huang, Y. (2019). Subjective well-being in China's changing society. *Proceedings of the National Academy of Sciences, 116*(34), 16799–16804.
Couture, V., Faber, B., Gu, Y., & Liu, L. (2021). Connecting the countryside via e-commerce: Evidence from China. *American Economic Review: Insights, 3*(1), 35–50.
Curtis, C. C., Lugauer, S., & Mark, N. C. (2015). Demographic patterns and household saving in China. *American Economic Journal: Macroeconomics, 7*(2), 58–94.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review, 110*(9), 2964–2996.
Dehejia, R., DeLeire, T., & Luttmer, E. F. (2007). Insuring consumption and happiness through religious organizations. *Journal of Public Economics, 91*(1–2), 259–279.
Démurger, S., & Xu, H. (2011). Return migrants: The rise of new entrepreneurs in rural China. *World Development, 39*(10), 1847–1861.
Di Tella, R., MacCulloch, R. J., & Oswald, A. J. (2001). Preferences over inflation and unemployment: Evidence from surveys of happiness. *American Economic Review, 91*(1), 335–341.
DiMaggio, P., & Bonikowski, B. (2008). Make money surfing the web? The impact of internet use on the earnings of US workers. *American Sociological Review, 73*(2), 227–250.
Dohman, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability. *American Economic Review, 100*(3), 1238–1260.
Doney, P. M., Cannon, J. P., & Mullen, M. R. (1998). Understanding the influence of national culture on the development of trust. *Academy of Management Review, 23*(3), 601–620.
Falk, M., & Hagsten, E. (2015). E-commerce trends and impacts across Europe. *International Journal of Production Economics, 170*, 357–369.
Fan, H., & Liang, C. (2023). The pollutant and carbon emissions reduction synergistic effect of green fiscal policy: Evidence from China. *Finance Research Letters, 58*, Article 104446.
Fan, J., Tang, L., Zhu, W., & Zou, B. (2018). The Alibaba effect: Spatial consumption inequality and the welfare gains from e-commerce. *Journal of International Economics, 114*, 203–220.
Ferrer-i-Carbonell, A., & Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? *Economic Journal, 114*(497), 641–659.
Gentzkow, M. (2006). Television and voter turnout. *Quarterly Journal of Economics, 121*(3), 931–972.
Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics, 225*(2), 254–277.
Guiso, L., Sapienza, P., & Zingales, L. (2004). The role of social capital in financial development. *American Economic Review, 94*(3), 526–556.
Guo, J., Jin, S., Zhao, J., Wang, H., & Zhao, F. (2022). Has COVID-19 accelerated the E-commerce of agricultural products? Evidence from sales data of E-stores in China. *Food Policy, 112*, Article 102377.
Hartog, J., & Oosterbeek, H. (1998). Health, wealth and happiness: Why pursue a higher education? *Economics of Education Review, 17*(3), 245–256.
Helliwell, J. F. (2006). Well-being, social capital and public policy: What's new? *Economic Journal, 116*(510), C34–C45.
Hu, M., Xiang, G., & Zhong, S. (2021). The burden of social connectedness: Do escalating gift expenditures make you happy? *Journal of Happiness Studies, 22*(8), 3479–3497.
Jin, H., Li, L., Qian, X., & Zeng, Y. (2020). Can rural e-commerce service centers improve farmers' subject well-being? A new practice of internet plus rural public services' from China. *International Food and Agribusiness Management Review, 23*(5), 681–695.
Jo, Y. J., Matsumura, M., & Weinstein, D. E. (2019). The impact of e-commerce on relative prices and consumer welfare. In *National Bureau of Economic Research, No. w26506*.
Knight, J., & Gunatilaka, R. (2010). Great expectations? The subjective well-being of rural–urban migrants in China. *World Development, 38*(1), 113–124.
Knight, J., & Gunatilaka, R. (2022). Income inequality and happiness: Which inequalities matter in China? *China Economic Review, 72*, Article 101765.
Kumar, P., Kumar, P., & Garg, R. K. (2021). A study on farmers' satisfaction and happiness after the land sale for urban expansion in India. *Land Use Policy, 109*, Article 105603.
Kwon, S. W., & Arenius, P. (2010). Nations of entrepreneurs: A social capital perspective. *Journal of Business Venturing, 25*(3), 315–330.
Lei, X., Shen, Y., & Yang, L. (2023). Digital financial inclusion and subjective well-being: Evidence from China health and retirement longitudinal study. *China Economic Review, 81*, Article 102013.
Leng, X. (2022). Digital revolution and rural family income: Evidence from China. *Journal of Rural Studies, 94*, 336–343.
Li, C. Y., & Ku, Y. C. (2018). The power of a thumbs-up: Will e-commerce switch to social commerce? *Information & Management, 55*(3), 340–357.
Li, G., & Qin, J. (2022). Income effect of rural E-commerce: Empirical evidence from Taobao villages in China. *Journal of Rural Studies, 96*, 129–140.
Li, H., Mu, W., Chen, T., & Wu, J. (2024). A social network perspective on household cooking fuel transition: Evidence from China. *Energy Economics, 131*, Article 107314.
Li, L., Du, K., Zhang, W., & Mao, J. Y. (2019). Poverty alleviation through government-led e-commerce development in rural China: An activity theory perspective. *Information Systems Journal, 29*(4), 914–952.
Li, L., Zeng, Y., Ye, Z., & Guo, H. (2021). E-commerce development and urban-rural income gap: Evidence from Zhejiang Province, China. *Papers in Regional Science, 100*(2), 475–494.
Lin, J., Li, H., Lin, M., & Li, C. (2022). Rural e-commerce in China: Spatial dynamics of Taobao villages development in Zhejiang Province. *Growth and Change, 53*(3), 1082–1101.
Lin, X., Wang, X., & Hajli, N. (2019). Building e-commerce satisfaction and boosting sales: The role of social commerce trust and its antecedents. *International Journal of Electronic Commerce, 23*(3), 328–363.
Liu, H., Ma, J., & Zhao, L. (2023). Public long-term care insurance and consumption of elderly households: Evidence from China. *Journal of Health Economics, 90*, Article 102759.
Liu, M., Zhang, Q., Gao, S., & Huang, J. (2020). The spatial aggregation of rural e-commerce in China: An empirical investigation into Taobao villages. *Journal of Rural Studies, 80*, 403–417.
Liu, S., Koster, S., & Chen, X. (2022). Digital divide or dividend? The impact of digital finance on the migrants' entrepreneurship in less developed regions of China. *Cities, 131*, Article 103896.
Liu, Y., & Zhou, M. (2023). Can rural e-commerce narrow the urban-rural income gap? Evidence from coverage of Taobao villages in China. *China Agricultural Economic Review, 15*(3), 580–603.
Lohmann, S. (2015). Information technologies and subjective well-being: Does the internet raise material aspirations? *Oxford Economic Papers, 67*(3), 740–759.
Lu, H., & Kandilov, I. T. (2021). Does mobile internet use affect the subjective well-being of older Chinese adults? An instrumental variable quantile analysis. *Journal of Happiness Studies, 22*(7), 3137–3156.
Lu, Y., Wang, J., & Zhu, L. (2019). Place-based policies, creation, and agglomeration economies: Evidence from China's economic zone program. *American Economic Journal: Economic Policy, 11*(3), 325–360.
Lu, Z., Li, P., Li, Q., & Zhang, H. (2024). To be, or not to be, happy? That's the question: A study on three levels of the digital divide and individual happiness in China. *Information Technology & People, 37*(4), 1802–1836.
Luo, R., Lyu, Q., Rozelle, S., & Wang, S. (2021). Early child development and caregiver subjective well-being in rural China. *China Agricultural Economic Review, 13*(2), 302–318.
Luo, S., Sun, Y., & Zhou, R. (2022). Can fintech innovation promote household consumption? Evidence from China family panel studies. *International Review of Financial Analysis, 82*, Article 102137.
Ma, S., Chai, Y., & Zhang, H. (2018). Rise of cross-border e-commerce exports in China. *China & World Economy, 26*(3), 63–87.
Ma, S., & Mu, R. (2020). Forced off the farm? Farmers' labor allocation response to land requisition in China. *World Development, 132*, Article 104980.
Marein, B. (2023). Public health departments and the mortality transition in Latin America: Evidence from Puerto Rico. *Journal of Development Economics, 160*, Article 102980.
Martinsons, M. G. (2008). Relationship-based e-commerce: Theory and evidence from China. *Information Systems Journal, 18*(4), 331–356.
Matsuura, T., & Ma, X. (2022). Living arrangements and subjective well-being of the elderly in China and Japan. *Journal of Happiness Studies, 23*(3), 903–948.
Meng, W., Yuan, G., & Sun, Y. (2023). Expansion of social networks and household carbon emissions: Evidence from household survey in China. *Energy Policy, 174*, Article 113460.
Mitchell, M. E., Lebow, J. R., Uribe, R., Grathouse, H., & Shoger, W. (2011). Internet use, happiness, social support and introversion: A more fine grained analysis of person variables and internet activity. *Computers in Human Behavior, 27*(5), 1857–1861.

Myrskylä, M., & Margolis, R. (2014). Happiness: Before and after the kids. *Demography, 51*(5), 1843–1866.
Naudé, W., Amorós, J. E., & Cristi, O. (2014). "Surfeiting, the appetite may sicken": Entrepreneurship and happiness. *Small Business Economics, 42*, 523–540.
Nie, P., Sousa-Poza, A., & Nimrod, G. (2017). Internet use and subjective well-being in China. *Social Indicators Research, 132*, 489–516.
Niu, G., Jin, X., Wang, Q., & Zhou, Y. (2022). Broadband infrastructure and digital financial inclusion in rural China. *China Economic Review, 76*, Article 101853.
Pina, D. L., & Bengtson, V. L. (1993). The division of household labor and wives' happiness: Ideology, employment, and perceptions of support. *Journal of Marriage and the Family, 55*(4), 901–912.
Qi, J., Zheng, X., & Guo, H. (2019). The formation of Taobao villages in China. *China Economic Review, 53*, 106–127.
Qin, Q., Guo, H., Shi, X., & Chen, K. (2023). Rural e-commerce and county economic development in China. *China & World Economy, 31*(5), 26–60.
Ram, R. (2010). Social capital and happiness: Additional cross-country evidence. *Journal of Happiness Studies, 11*, 409–418.
Schnittker, J. (2008). Happiness and success: Genes, families, and the psychological effects of socioeconomic position and social support. *American Journal of Sociology, 114*(S1), S233–S259.
Sebora, T. C., Lee, S. M., & Sukasame, N. (2009). Critical success factors for e-commerce entrepreneurship: An empirical study of Thailand. *Small Business Economics, 32*, 303–316.
Sun, S., Chen, J., Johannesson, M., Kind, P., & Burström, K. (2016). Subjective well-being and its association with subjective health status, age, sex, region, and socio-economic characteristics in a Chinese population study. *Journal of Happiness Studies, 17*, 833–873.
Sun, X., Yan, S., Liu, T., & Wang, J. (2023). The impact of high-speed rail on urban economy: Synergy with urban agglomeration policy. *Transport Policy, 130*, 141–154.
Tan, Y., & Li, X. (2022). The impact of internet on entrepreneurship. *International Review of Economics and Finance, 77*, 135–142.
Tang, W., & Zhu, J. (2020). Informality and rural industry: Rethinking the impacts of E-commerce on rural development in China. *Journal of Rural Studies, 75*, 20–29.
Teo, T. S., & Liu, J. (2007). Consumer trust in e-commerce in the United States, Singapore and China. *Omega, 35*(1), 22–38.
Tran, L. T. T. (2021). Managing the effectiveness of e-commerce platforms in a pandemic. *Journal of Retailing and Consumer Services, 58*, Article 102287.
Tu, Q., & Bulte, E. (2010). Trust, market participation and economic outcomes: Evidence from rural China. *World Development, 38*(8), 1179–1190.
Wang, C. C., Miao, J. T., Phelps, N. A., & Zhang, J. (2021). E-commerce and the transformation of the rural: The Taobao village phenomenon in Zhejiang Province, China. *Journal of Rural Studies, 81*, 159–169.
Wang, H., Liu, H., & Fu, Q. (2024). Effects of common prosperity on China's education expenditure—Empirical analysis based on CFPS quasi-micro panel data. *International Review of Economics and Finance, 91*, 440–455.
Wang, J., Wang, C., Li, S., & Luo, Z. (2021). Measurement of relative welfare poverty and its impact on happiness in China: Evidence from CGSS. *China Economic Review, 69*, Article 101687.
Wang, Z., Luo, S., Xu, J., Wang, Y., Yun, H., Zhao, Z., … Wang, Y. (2021). Well-being reduces COVID-19 anxiety: A three-wave longitudinal study in China. *Journal of Happiness Studies, 22*, 3593–3610.
Wei, B., Zhao, C., Luo, M., Jia, Z., & Xia, H. (2024). Beyond the scenery: Understanding the impact of rural tourism development on household consumption in China. *Review of Development Economics*. https://doi.org/10.1111/rode.13096
Xia, J., & Lu, T. J. (2008). Bridging the digital divide for rural communities: The case of China. *Telecommunications Policy, 32*(9–10), 686–696.
Yang, T., & Zhang, T. (2024). Social capital meets guanxi: Social networks and income inequality in China. *China Economic Review, 83*, Article 102094.
Yang, Z., Shi, Y., & Yan, H. (2017). Analysis on pure e-commerce congestion effect, productivity effect and profitability in China. *Socio-Economic Planning Sciences, 57*, 35–49.
Yeung, G., He, C., & Zhang, P. (2017). Rural banking in China: Geographically accessible but still financially excluded? *Regional Studies, 51*(2), 297–312.
Yin, Z. H., & Choi, C. H. (2022). Does e-commerce narrow the urban-rural income gap? Evidence from Chinese provinces. *Internet Research, 32*(4), 1427–1452.
Yoon, C. (2009). The effects of national culture values on consumer acceptance of e-commerce: Online shoppers in China. *Information & Management, 46*(5), 294–301.
Yu, F., Xiao, D., & Chang, M. S. (2021). The impact of carbon emission trading schemes on urban-rural income inequality in China: A multi-period difference-in-differences method. *Energy Policy, 159*, Article 112652.
Zang, Y., Hu, S., Zhou, B. B., Lv, L., & Sui, X. (2023). Entrepreneurship and the formation mechanism of Taobao villages: Implications for sustainable development in rural areas. *Journal of Rural Studies, 100*, Article 103030.
Zhan, G., & Zhou, Z. (2018). Mobile internet and consumer happiness: The role of risk. *Internet Research, 28*(3), 785–803.
Zhang, X., Zhang, X., & Chen, X. (2017). Happiness in the air: How does a dirty sky affect mental health and subjective well-being? *Journal of Environmental Economics and Management, 85*, 81–94.
Zhao, C., & Guo, J. (2022). Are veterans happy? Long-term military service and the life satisfaction of elderly individuals in China. *Journal of Happiness Studies, 23*(2), 477–508.
Zhao, C., Li, X., & Yan, J. (2024). The effect of digital finance on residents' happiness: The case of mobile payments in China. *Electronic Commerce Research, 24*(1), 69–104.
Zhao, C., Wu, Y., & Guo, J. (2022). Mobile payment and Chinese rural household consumption. *China Economic Review, 71*, Article 101719.
Zhao, D., Jiang, J., & Yin, Z. (2020). Can entrepreneurship bring happiness? Evidence from China. *Economic Modelling, 91*, 679–686.
Zheng, H., & Ma, W. (2021). Click it and buy happiness: Does online shopping improve subjective well-being of rural residents in China? *Applied Economics, 53*(36), 4192–4206.
Zhou, J., Yu, L., & Choguill, C. L. (2021). Co-evolution of technology and rural society: The blossoming of taobao villages in the information era, China. *Journal of Rural Studies, 83*, 81–87.
Zhou, Y., Huang, X., Shen, Y., & Tian, L. (2023). Does targeted poverty alleviation policy lead to happy life? Evidence from rural China. *China Economic Review, 81*, Article 102037.
Zou, H., Xiong, Q., & Xu, H. (2020). Does subjective social status predict self-rated health in Chinese adults and why? *Social Indicators Research, 152*, 443–471.

Teaching and Teacher Education 119 (2022) 103857

Contents lists available at ScienceDirect

Teaching and Teacher Education

journal homepage: www.elsevier.com/locate/tate

Research paper

More is less: Homeroom teachers' administrative duties and students' achievements in China

Boou Chen [a], Chunkai Zhao [b, *]

[a] School of Economics, Nanjing University of Finance & Economics, No.3 Wenyuan Road, Qixia District, Nanjing, Jiangsu Province, China
[b] College of Economics and Management, South China Agricultural University, No.483, Wushan, Tianhe District, Guangzhou, Guangdong Province, China

HIGHLIGHTS

- Using a quasi-natural experiment for random student-classroom assignment.
- HTs' administrative duties are detrimental to students' academic achievements.
- Mechanisms include less teaching time and communication with students and class peer effects.
- This negative effect is more effective in the higher title, male, and middle-aged HTs.
- Reducing HTs' workload and stress from administrative duties helps improve student performance.

ARTICLE INFO

Article history:
Received 27 September 2021
Received in revised form
4 August 2022
Accepted 6 August 2022
Available online xxx

Keywords:
Administrative duties
Students' achievements
Homeroom teachers (HTs)
Random student-classroom assignment
Chinese secondary schools

ABSTRACT

This paper investigated causal effects of homeroom teachers' (HTs') administrative duties on students' achievements in Chinese secondary schools. Using data from China Education Panel Survey, we employed random student-classroom assignment as a quasi-natural experiment, finding that HTs' administrative duties curbed students' achievements, which is supported by some robustness checks like instrumental variable methods. Mechanism analysis indicated that this adverse impact can be explained by less teaching time, reduction of communication with students, and negative peer effects. Furthermore, negative effects of administrative duties were more pronounced among grade team leaders and school leaders, HTs with higher titles, male, and middle-aged HTs.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

Although it is very common in most countries for university teachers to hold administrative positions and be responsible for academic administrations (e.g., García-Gallego et al., 2015; Gigante & Firestone, 2008; Gutman, 2018; Tsang & Kwong, 2017), administrative duties for compulsory education[1] teachers are mainly found in some Asian countries, such as China, Korea, and Japan (Kim, 2019; Okumura, 2017; Park & Sung, 2013; Yang et al., 2009). In China, teachers are allowed and encouraged to take on administrative duties for purposes of easing staff management, compressing authorized strength, implementing educational reform, and saving education finance (Tsang & Kwong, 2017; Zheng & Sun, 2020), which is known as the "double-shouldered" (Li, 2018). These administrative duties involve school-related tasks such as school management, teacher management, training, meetings, and appraisals (Droogenbroeck & Spruyt, 2014; Zhong & Liu, 2021).[2] Recent reports indicated that approximately 16% of teachers in Chinese junior high schools were required to complete additional administrative tasks (Zheng & Sun, 2020); in addition, they spent less time correcting students' work, but worked longer hours and had lower job satisfaction (Li, 2018). Surprisingly, few studies

* Corresponding author.
E-mail addresses: drchen2015@163.com (B. Chen), ckz_zhao@163.com (C. Zhao).

[1] Compulsory education in China includes primary and junior high school.

[2] In China, teachers may also deal with social administrative tasks not related to school, such as poverty alleviation, knowledge promotion, community work, and cleaning (source: https://m.thepaper.cn/baijiahao_15782960). The administrative duties in this paper are school-related. Subject to the availability of data, these social administrative tasks are not included.

https://doi.org/10.1016/j.tate.2022.103857
0742-051X/© 2022 Elsevier Ltd. All rights reserved.

addressed the impact of teachers' administrative duties on students' academic achievements in Chinese compulsory education (Sun & Zhou, 2022).

There are two different effects of teachers' administrative duties on students' performance. Clearly, on the negative side, administrative tasks may consume teachers' energy, squeeze teaching time (Kim, 2019; Park & Sung, 2013; Sun & Zhou, 2022), and even create professional stress (Lasky, 2005; Yang et al., 2009), which leads to a decline in teaching quality (Ballet et al., 2006; Kim & Park, 2014). Additionally, there are different work norms and assessment systems between teaching and administrative work, and administrative duties require teachers to switch between different roles and orientations (Kim, 2019; Tsang & Kwong, 2017), accelerating their burnout and sense of job exhaustion (Kim & Park, 2014; Zheng & Sun, 2020). Especially for homeroom teachers (HTs) in China, it is difficult for them to have enough energy to coordinate teaching, classroom management, and administrative tasks.

On the positive side, under the influence of traditional Confucianism, Chinese schools always give preference to outstanding, famous, or senior teachers for administrative positions. These teachers have more teaching experience and better teaching performance, and holding administrative duties enhances their exemplary responsibility in teaching (Zheng & Sun, 2020), which may improve their students' achievements. Moreover, for students in compulsory education, teachers in administrative positions have more authority and positive influence than other teachers (Gerrard & Farrell, 2014).

These conflicts from previous studies may be influenced by potential endogenous problems. In theory, whether a teacher holds administrative duties is not random and may instead be selective (Sun & Zhou, 2022). For example, in China, some teachers with higher credentials and better teaching quality are more likely to obtain administrative positions (Zheng & Sun, 2020). Additionally, identifying the causal relationship between teachers' administrative duties and students' achievements is influenced by many confounding factors, such as school and teacher characteristics, teaching quality, subjective willingness, job satisfaction, and many other unobserved features that correlate with the entry of administrative duties. However, these confounders may also potentially affect students' achievements but are difficult to fully control (Caprara et al., 2006; Gong et al., 2018; Hanushek & Rivkin, 2010). To overcome the endogeneity problems caused by self-selection and omitted variable bias, according to previous studies (e.g., Gong et al., 2018; Hu, 2018; Huang et al., 2021; Luo et al., 2021; Sun & Zhou, 2022; Wang et al., 2018; Xu & Li, 2018), we used a quasi-natural experiment with the random student-classroom assignment to identify the causal effect of HTs' administrative duties on students' achievements.

The reason why we focused on HTs' administrative duties is multifaceted. First, before students are assigned to classes in Chinese junior high schools, the HT of a particular class is determined before the start of the school year as opposed to other teachers. Thus, although whether HTs hold administrative duties are not randomized, the quasi-natural experiment of random student−classroom assignment can be regarded as a random match of students and HTs, which facilitates causal identification. Second, at the compulsory education stage in China, the important impact of the HT on students is not only on academic achievements, but also involves the physical and mental health as well as the development of cognitive and non-cognitive skills of students (e.g., Chen et al., 2014; Liu & Barnhart, 1999; Shi & Leuwerke, 2010; Wang & Rao, 2019). However, after balancing teaching and class management, if HTs are still required to take on administrative tasks, it can consume a lot of their energy and may even disrupt the teaching schedule and lead to a decrease in teaching quality (Frenzel et al., 2019; Keller et al., 2016). In fact, HTs with administrative duties are very common in Chinese primary and secondary schools (Li, 2018). Therefore, the purpose of this study is twofold. We not only emphasized on the importance of HTs in Chinese junior high school students, but also focused on the consequence of increasing teachers' workload on students, which provides insight into the development of educational policies related to reducing the burdens of teachers.

We identify the impact of HTs' administrative duties on students' academic achievement. We exploited the baseline wave of the China Education Panel Survey (CEPS), a large-scale and nationally representative survey of Chinese middle school students, teachers, parents, and school administrators (Hu, 2018; Huang et al., 2021). Our contributions are summarized below. First, our empirical results show that HTs' administrative duties are detrimental to student achievement, therefore, we added to the literature on the adverse effects of teachers' administrative duties. Previous related studies on teachers holding administrative duties mainly focused on teachers' workload and efficiencies, psychological states, and opportunity costs of instructional activities (Kim, 2019; Lasky, 2005; Tsang & Kwong, 2017; Yang et al., 2009). In addition, a recent study was also interested in the negative effects of teacher administrative duties on student achievement (Sun & Zhou, 2022), but the starting point is not consistent with our paper. As in the literature we have mentioned (Kim, 2019; Lasky, 2005), we placed more emphasis on the role of HT and the burden and stress associated with HT administrative duties in China, which in turn affects student achievements.

Second, our research supplemented the broad field on the impact of teacher characteristics on student achievements. Previous literature addressed demographic characteristics of teachers (e.g., Antecol et al., 2015; Feng & Li, 2016; Gong et al., 2018; Mullola et al., 2011; Xu & Li, 2018), teacher quality (e.g., Chu et al., 2015; Hanushek & Rivkin, 2010; Rivkin et al., 2005; Rockoff, 2004), teacher earnings (e.g., Eberts et al., 2002; Kingdon & Teal, 2010; Loyalka et al., 2019; Muralidharan & Sundararaman, 2011), teacher expectations (Farkas et al., 1990; Good, 1987), job satisfaction and self-efficacy of teacher (Caprara et al., 2006; Madigan & Kim, 2021), and so forth. Our findings based on the case of HTs might offer some insights into reducing the workload of Chinese teachers in compulsory education.

Third, we figured out three possible mechanisms by which HTs' administrative duties influenced students' achievements: teaching time and teachers' enthusiasm, teacher-student interaction and home-school connection, and peer effects. While previous literature analyzed mechanisms mainly in terms of teacher effort and teaching methods (Sun & Zhou, 2022), we further theorized the rationality of these three mechanisms. More importantly, this paper complemented mechanisms in terms of teacher-student interaction and negative peer effects, which have received widespread attention in the field of economics of education. Furthermore, in heterogeneity effects, we found that male and middle-aged HTs and HTs with higher administrative positions and job titles, and administrative duties had greater negative impacts on students' achievement. Also, boys were more likely to be influenced by HT administrative duties. These results provided some references for assigning and organizing administrative tasks in Chinese junior high schools and improving educational policies for coordinating teaching and administrative tasks.

The rest of the paper is organized as follows. Section 2 sets out the background on the HT system and teachers' administrative

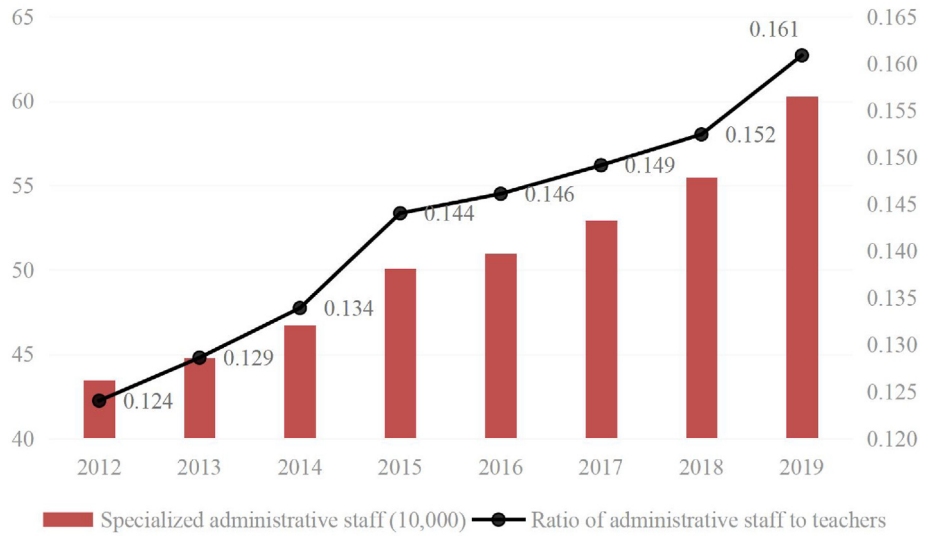


Fig. 1. Total administrative staff and ratio of administrative staff to teachers in secondary schools. Source: National education development statistics bulletin of Ministry of Education of China.

duties in Chinese middle schools. Section 3 outlines the potential theoretical mechanisms of HTs' administrative duties on students' achievements. Section 4 describes the data and variables. Section 5 exhibits the identification strategy. Section 6 presents and discusses the empirical results. Section 7 concludes this paper.

2. Background

2.1. The HT system in China

The HT system in China originated in the Qing Dynasty with a system of grade-appointed teachers. This model has been used since the founding of the People's Republic of China, and HTs have been assigned to all primary and secondary schools (including junior and senior high schools). The tasks of class teachers are clearly defined according to the Provisional Regulations on Homeroom Teacher Work in Primary and Secondary Schools. The basic task of HTs is to carry out class work in accordance with the requirements of all-around development of morality, intellect, physique, and aesthetics, and to educate, manage, and guide students comprehensively.

The HT system is not unique to China, but is also common in other countries, like Japan, the USA, UK, Australia, Canada, and Argentina (Chu et al., 2015; Shi & Leuwerke, 2010). However, the HTs in China and Japan are responsible for all aspects of the student's school experience (Kashy-Rosenbaum et al., 2018; Li et al., 2014). Specifically, HTs play several different roles in schools. They are leaders, organizers, and managers of the classes, bridges between school, parents, and society, effective aids to school leaders in carrying out educational programs in schools, and grassroots forces in implementing the national education policies (Liu & Barnhart, 1999; Shi & Leuwerke, 2010). Additionally, most HTs in China are also required to teach major subjects, especially in public elementary and middle schools. In summary, the HT is the core and overall connection of the students while they are in the school.

2.2. Teachers' administrative duties in Chinese secondary schools

Teachers who also have administrative duties are a special yet widespread phenomenon in Chinese compulsory education schools. The reasons are manifold. First, there is the conflict between the increasing complex administrative workload and the shortage of administrative staff in Chinese junior high school. Currently, Chinese schools are moving away from the traditional model of only teaching scientific knowledge. In the context of quality education reform, the interactions between schools and society are gradually increasing (Tan, 2019). In addition to the day-to-day teaching management, public relations, home-school cooperation, school construction, and diversified student activities have led to a steep increase in the administrative workload of secondary schools.

To reconcile teaching and administration, as early as 2001, the General Office of the State Council of China issued the "Opinions on the Formulation of Staffing Standards for Primary and Secondary Schools," which required that the management of primary and secondary schools should be performed by teachers as much as possible, and logistics services should be gradually socialized.[3] Subsequently, in 2002, the Ministry of Education further stipulated that in a standardized middle school class of 45–50 students, the average number of course instructors is 2.7 with 0.6–1 administrative personnel in urban schools, and secondary schools in towns or rural areas should have 0.1–0.4 administrative staff.[4] Put another way, the ratio of administrative staff to teachers in urban secondary schools should be between 22.22% and 37.03%. However, according to the national education development statistics bulletin released annually by the Ministry of Education, this ratio is far from being achieved.

In Fig. 1, it is clear that although the number of specialized administrative staff and the ratio to teaching faculty has climbed each year, in 2019, this ratio was only 16.1%. Even in some rural junior high schools, there are no dedicated administrative staff, and all administrative tasks are performed by the classroom teachers (Zheng & Sun, 2020). Taken together, subject to the objective reasons of increased administrative work, lack of professional administrators, and policy incentives, it is increasingly common for

[3] Information source: http://www.gov.cn/gongbao/content/2001/content_61159.htm.

[4] Data source: http://www.moe.gov.cn/jyb_xxgk/gk_gbgg/moe_0/moe_8/moe_26/tnull_316.html.

Chinese middle school teachers to hold administrative duties.

In addition to these outside options, access to extra part-time income from administrative work is an important reason. In China, public middle school teachers are not well paid (Liu, 2012; Liu, 2021); therefore, some teachers may volunteer for administrative duties because of the administrative stipends and subsidies and additional benefit packages. Moreover, administrative positions, as a special kind of political capital in Chinese schools, also bring more implicit benefits to teachers, such as more training opportunities, better research conditions and team mentoring, and preferential access to some of the school's physical and human resources (García-Gallego et al., 2015; Gigante & Firestone, 2008). More importantly, administrative duties are regarded in many schools as a form of performance appraisal and help teachers to advance their titles and expand their social networks.

3. Mechanisms

In this section, we discussed in detail the theoretical mechanisms of HTs' administrative duties on students' achievements. We classified potential channels into the following categories: teaching time and teachers' enthusiasm, teacher-student interaction and home-school connection, and peer effects.

3.1. Teaching time and teachers' enthusiasm

Administrative duties affect teachers' efforts and enthusiasm in two ways. On the one hand, holding administrative tasks make it necessary for teachers to complete regular teaching tasks as well as to handle various non-teaching school affairs, which is a great test of teachers' energy and schedule (García-Gallego et al., 2015; Li, 2018). For HTs in particular, they have to coordinate their time and energy between teaching, classroom management, and administrative duties. More importantly, administrative works require consideration of requests and orders from superiors (Bogler, 2001), and the satisfaction of superiors with the administrative work even partially determines teachers' income and promotion opportunities (Manion, 1985; Xie & Yang, 2021); thus, teachers are likely to prioritize administrative work. With the accumulation of work stress, teachers are in a state of high physical and mental exertion, which shortens their teaching time and disrupts teaching arrangement (Chen, 2016; Skaalvik & Skaalvik, 2011). Therefore, administrative duties may lead to a decline in HTs' teaching hours or even their ability to effectively perform their teaching duties, ultimately leading to lower quality teaching and lower student achievement (Ballet et al., 2006; Hargreaves, 2002; Lazarides et al., 2018).

On the other hand, teachers, especially HTs, combine multiple and complex roles in schools (Shi & Leuwerke, 2010), and administrative duties further lead to increased teacher role conflict, resulting in the need for HTs to meet multiple role expectations (Kim, 2019). Teachers with administrative duties need to seek role balance and resolve role conflicts in a complex role cluster, which poses a rigorous challenge to teachers' mental stamina and enthusiasm for teaching to maintain (Kim & Park, 2014; Tsang & Kwong, 2017). Additionally, in Chinese junior high schools, administrative duties held by teachers are mostly in middle management positions. These positions are usually in the key departments in school management (Zheng & Sun, 2020), involving complex interpersonal relationships consisting of school leaders, teachers, parents, alumni, and students. If administrative tasks are not handled properly, it is easy to fall into a bad emotional experience and weaken teachers' enthusiasm and teaching motivation (Bogler, 2001). Teacher enthusiasm and job satisfaction as the core factors affecting teaching (Caprara et al., 2006; Frenzel et al., 2019; Kunter et al., 2008), the burnout brought by complicated administrative duties may affect teachers' effectiveness and teaching quality (Ballet et al., 2006; Skaalvik & Skaalvik, 2011).

3.2. Teacher-student interaction and home-school connection

The costs for administrative duties are also reflected in reducing teacher-student interactions. First, the administrative workload may weaken teachers' innovation in teaching methods. Since teachers with administrative positions may not have much time to prepare lessons and refine lesson plans (Hargreaves, 1992; Kim, 2019), they are more inclined to adopt the traditional cramming system and reduce classroom interactions. Second, the contents of the main subjects (Chinese, Math, and English) are relatively stable in Chinese compulsory education. Teachers often spend a lot of time at the end of each day's lecture reviewing to reinforce learning, such as lecturing on exercises, tutoring and supervising course work, and helping students with memorization (He et al., 2021). However, the administrative workload may compress these opportunities. Third, there are many class activities in primary and secondary schools that require teachers and students to complete together, especially the organization and scheduling of HTs (Liu & Barnhart, 1999), but administrative duties are likely to crowd out HTs' extracurricular time, resulting in less teacher-student interactions in out-of-class time.

Another hazard of teachers' administrative duties is the reduction of home-school connections. Currently, home-school co-education is commonly adopted worldwide (McCarthey, 2000; Núñez et al., 2015; Urbina-Garcia, 2019). Strengthening contacts between course teachers or HTs and students' parents helps to reduce information asymmetry. On the one hand, parents can have a clearer understanding of their students' learning status and academic performance at school and take targeted measures to correct potential poor learning habits (Bauch & Goldring, 1995). On the other hand, teachers may also learn more about students' extracurricular learning conditions and get some feedback on teaching and students' homework through parents' descriptions (McCarthey, 2000; Núñez et al., 2015). Ways of home-school connections in Chinese primary and secondary schools include parent-teacher conferences, home visits, and communication via phone, WeChat, and SMS. Clearly, these additional tasks are equally time and energy intensive and may conflict with the teacher's administrative work, especially by compressing some face-to-face communication opportunities such as parent-teacher conferences and home visits, as they require more time to prepare (Jensen, 2006). Therefore, HTs' administrative duties are likely to decrease home-school connections and weaken the home-school co-education function.

3.3. Peer effects

Peer effects have been widely discussed in the field on the economics of education. Prior literature confirmed that class and grade characteristics have an impact on individual students' achievement and educational behaviors (e.g., Balestra et al., 2022; Duflo et al., 2011; Hu, 2018; Huang et al., 2021; Lavy et al., 2012; Sacerdote, 2001; Wang et al., 2018). Teachers, as an important component of instructional activities, can take several steps to play a more active and effective role in the operation of the peer effect (e.g., Duflo et al., 2011; Feng & Li, 2016; Hendrickx et al., 2016; Lavy et al., 2012; Li et al., 2014). Representative research came from Li et al. (2014). They conducted an experiment in China in which the 20 lowest-performing students were randomly assigned to the 10 highest-performing students in the class as peers for a semester. Results showed that the test scores of low-performing students

were significantly higher, while the high-performing students were not adversely affected. More importantly, HTs, as class managers, have more tools to intervene in the class peer effect, such as arranging student seating, organizing class activities, appointing class officers, and contacting home and arrangements (Feng & Li, 2016; Shi & Leuwerke, 2010). However, as mentioned earlier, the administrative workload may take up a lot of time and energy of HTs, making it difficult for them to take effective measures to intervene in class peer effects. Even the lack of effective classroom management may lead to classroom discipline disruptions (Kim, 2019), resulting in lower student achievement.

4. Data and variables

4.1. Data

The data used in this study is from the baseline wave of the CEPS in the 2013–2014 academic year, which is conducted by the National Survey Research Central (NSRC) of Renmin University of China. The CEPS database uses the average education level and the proportion of the floating population as stratified variables to randomly select about 20,000 secondary school students (in seventh and ninth grades) from 112 schools and 438 classes across the country and collects information about students, their parents, HTs, main subject teachers, classes, and schools. It is a major high-quality education micro-survey with national representativeness in China (e.g., Gong et al., 2018; Huang et al., 2021; Luo et al., 2021; Xu & Li, 2018).

Specifically, in the HT questionnaire, the CEPS investigates whether the HT concurrently holds administrative duties, which allows us to define the key explanatory variables in this paper.[5] The CEPS also contains the students' mid-term test scores in the following three main subjects: Chinese, Math, and English, the outcome of interest in this paper. Therefore, the CEPS provided an accessible and reliable source of data for our research on the HTs' administrative duties and students' academic achievements.

The CEPS also contains information on class assignment rules. School administrators were inquired about whether students were randomly assigned to class or in terms of other assignment rules.[6] Among the schools surveyed, 93 schools adopted the random student-classroom assignment, accounting for about 83%. Referring to previous related research (Huang et al., 2021; Huang & Zhu, 2020), we limited the sample to those schools whose administrators randomly or averagely assigned students and adopt the random student–classroom assignment as a quasi-natural experiment to identify the causal relationship between HTs' administrative duties and student achievements. In addition, we excluded samples with missing values for the explanatory variable, explained variables, and control variables. Finally, the sample used in the analysis includes 8286 students and 344 HTs in 344 classes of 93 schools.[7]

[5] In this survey, administrative duties mainly include three categories: grade team leader (responsible for managing various affairs of the grade), head of teaching and research section (responsible for managing the teaching and research of a certain subject in the school), school leader (vice principal and director of teaching affairs).

[6] In the CEPS data, school administrators were asked to report which one of the following assignment rules they use to place students: (i) Placement exams prior to the beginning of students' first academic year; (ii) Students' residential status; (iii) Random student-classroom assignment; (iv) Others.

[7] The final data is obtained by matching the data of students, parents, HTs, and schools.

4.2. Variables and summary statistics

Table A1 in the Appendix presents the definition and descriptive statistics of the main variables. Panel A shows the students' scores of three main subjects in Chinese secondary schools. Since HTs are responsible for almost all the affairs of the students in the class, their administrative duties may not only affect the courses they teach, but also other subjects and the overall academic achievement of students. Therefore, we took the average of three main subjects' scores and constructed the variable "*Average scores*" to reflect the overall academic achievement of students. In addition, in order to examine the impact of HTs' administrative duties on the courses they taught, we constructed the variable "*Scores taught by HT*", and fixed the subject effects additionally in the corresponding regression. It should be noted that the subjects taught by the HT include Chinese, Math, English, and other subjects four categories in the survey, but the questionnaire only contains the scores of the first three main subjects. Therefore, there is no data on "*Scores taught by HT*" if the HT taught other subjects, and this is why the observations of this variable are fewer than the others.

In terms of control variables, the main characteristics of the HT are shown in Panel B of Table A1. We controlled the HTs' age, years of teaching, gender, schooling, and title which may affect both students' performance and their administrative positions (Kim & Park, 2014; Zheng & Sun, 2020). Adequate control of HTs characteristics helps to minimize potential endogeneity problems. In addition, based on previous related research, we also selected control variables from the levels of students, their parents, and families, including gender of students, hukou status, grade, education expenditure, schooling of their father and mother (Chu et al., 2015; Gong et al., 2018; Hu, 2018; Loyalka et al., 2019), as shown in Panel C.

Furthermore, as the T-test results shown in Table A2 of the Appendix, compared with students whose HTs do not have administrative duties, various scores of students whose HTs have administrative duties are all relatively lower, which supports the negative effect of HTs' administrative duties. While the summary statistics are suggestive, we next further examined the causal effect of HTs' administrative duties on the academic performance of students.

5. Empirical strategy

To investigate the effect of HTs' administrative duties on academic performance of students, we estimated the following linear model:

$$Y_{ics} = \alpha + \beta\, HTA_{cs} + \gamma\, T_{cs} + \delta\, P_{ics} + \lambda\, H_{ics} + \theta\, S_s + \varepsilon_{ics} \quad (1)$$

where Y_{ics} denotes the average scores or the scores of the subject taught by the HT of student i from class c in school s. HTA_{cs} is a dummy variable indicating whether the HT from class c in school s has administrative duties. In control variables, T_{cs}, P_{ics}, and H_{ics} represent a row vector of control variables at the HT, student, and household levels. S_s is a school fixed effect that controls for the endogenous sorting of students across schools. ε_{ics} is the error term. Additionally, throughout the empirical analysis, we clustered the standard errors at the class level to allow for heteroskedasticity and arbitrary serial correlation across students within each class (Balestra et al., 2022; Huang et al., 2021).

However, it is usually difficult to investigate the causal effect of HTs' administrative duties on the academic performance of students, because the assignment of the HTs may not be random in some schools, which would lead to some potential correlation between the HTs' administrative duties and some characteristics of

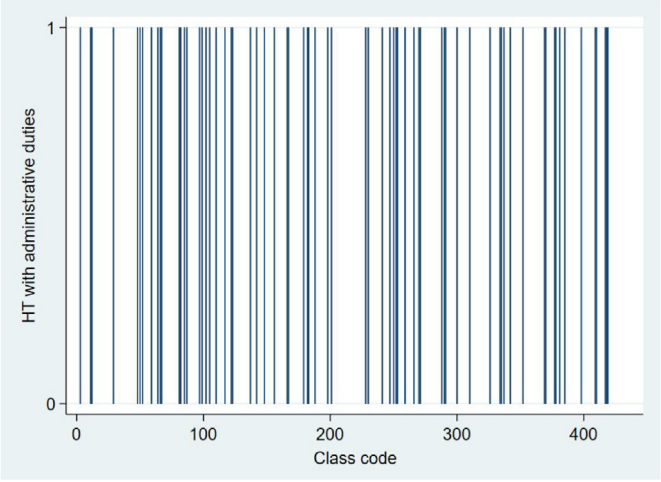


Fig. 2. The distribution of HTs with administrative duties in the sample.

classes, students and their households. Meanwhile, if the assignment of students is not random, such as assigning students with good grades to a class, it will also cause bias in the estimation. Referring to the empirical strategy commonly employed in previous literature (e.g., Gong et al., 2018; Hu, 2018; Huang et al., 2021; Huang & Zhu, 2020; Luo et al., 2021; Wang et al., 2018; Xu & Li, 2018), we employed the random student–classroom assignment as a quasi-natural experiment to address the possible endogeneity. Since each class has a unique HT, the characteristics of HT can also be considered as a kind of class characteristic. In the framework of random assignment of students to classes, the HTs and students are also randomly matched.[8] Therefore, we could obtain the unbiased estimate β of HTA_{cs} by running equation (1) using the OLS method.

To further prove the validity of the random assignment of HTs, we also plotted the changes of HT with administrative duties among the observed classes. As shown in Fig. 2, with the change of the class code, HT with administrative duties frequently changes between 0 and 1. This result shows that the distribution of HTs with administrative duties should be evenly dispersed throughout each class in this sample (Sun & Zhou, 2022). Taken together, we basically confirmed that the assignment of the HTs with administrative duties is relatively random, rather than with selection bias.

Before the baseline regressions, the validity of the underlying assumption used in this paper should be examined: the variation in whether the HT has administrative duties should be random conditional on school fixed effects. That means, whether the HT has administrative duties is expected to be uncorrelated with any predetermined background characteristics of the class and students in this class. We examined the above assumption by the T-tests and balancing test. The summary statistics of class-level variables in the T-tests and balancing tests are shown in Table A3 and Table A4 in the Appendix. First, we compared the class-level differences between the two groups whose HTs have administrative duties and those who have no administrative duties. Further, we run separate regressions with each background characteristics as outcomes to test this assumption at the class level (Wang et al., 2018).

The results of the T-tests are shown in Table A5 in the Appendix, indicating that most characteristics of two groups, HTs with administrative duties and HTs without administrative duties, are very similar. Although there are significant differences in class size, grade, average education expenditure, and rate of extracurricular tutoring, these are likely to come from the school or regional level, which cannot be controlled in the T-tests. Therefore, these results further confirmed that the students are randomly assigned. To further test for the randomness, we performed balancing tests on class-level ex-ante variables. In Table A6 of the Appendix, we found that whether the HT has administrative duties is not correlated with most background characteristics at the class-level, but only has a significant correlation with the students' grade. The results further support the randomness of student-classroom assignment in the sample, indicating that in the process of student-classroom assignment, we had greater confidence that HTs' administrative duties are also randomly distributed and independent within the school after controlling the grade, that is, HTs' administrative duties are not affected by other background characteristics at the class-level.

Therefore, by controlling the grade and school fixed effects, we were able to eliminate all observed correlations between the HTs' administrative duties and background characteristics at the class-level, lending support to the validity of the underlying assumption in this research. Additionally, we further controlled for a number of HT characteristics and student and family characteristics, indicating that the estimates in the following empirical

Table 1
The effect of HTs' administrative duties on students' achievements.

	(1)	(2)	(3)	(4)
	Average scores		Scores taught by HT	
Administrative duties	−4.5880***	−4.9388***	−3.5150*	−3.7402*
	(1.5888)	(1.5736)	(2.0894)	(2.1019)
Age of HT	0.1006	0.0771	1.1358**	1.1053**
	(0.3740)	(0.3703)	(0.5464)	(0.5409)
Years of teaching	−0.0304	−0.0394	−0.9902*	−0.9923*
	(0.3704)	(0.3714)	(0.5225)	(0.5218)
Gender of HT	−4.1470***	−3.9679***	−8.2504***	−7.8462***
	(1.5439)	(1.5223)	(2.2215)	(2.1937)
Schooling of HT	−1.0504	−1.0141	−0.2378	−0.1403
	(1.4686)	(1.3966)	(1.9829)	(1.9512)
Title	1.8788*	1.5708	3.5495***	3.4533***
	(1.0873)	(1.1123)	(1.2783)	(1.2779)
Gender of student		−7.1331***		−6.3285***
		(0.4489)		(0.6816)
Hukou		−0.4115		−0.6716
		(0.5829)		(0.7308)
Ninth grade		1.4656		0.7299
		(1.3745)		(1.8148)
Education expenditure		0.4000***		0.3254***
		(0.0991)		(0.1177)
Schooling of father		0.6663***		0.6183***
		(0.1123)		(0.1471)
Schooling of mother		0.2933***		0.2824**
		(0.0908)		(0.1230)
School fixed effects	Yes	Yes	Yes	Yes
Subject fixed effects	No	No	Yes	Yes
Constant	100.2383***	88.9229***	49.5565	38.1255
	(26.8711)	(25.1344)	(36.5054)	(35.3245)
R-squared	0.3693	0.4020	0.3527	0.3723
N	8286	8286	6985	6985

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses.

[8] We can use the "proof by contradiction" method to understand that the assignment of teachers is random in the process of random student–classroom assignment. The random assignment of students is the basic assumption. If it is assumed that in the process of random student–classroom assignment, HT A and students a, b, and c are not randomly matched. Then this also means that this part of students a, b, and c are not randomly matched. Because if students a, b, and c will definitely enter the class of HT A, that is, students a, b, and c are fixedly matched with HT A. It also indicates that student a must be matched with students b and c, which clearly violates the basic assumption of random student–classroom assignment. Therefore, random student–classroom assignment is not only a random match between students and classes, but also a random match of teaching resources, especially teachers.

analysis are unlikely to be biased due to the omission of main confounders and it should be safe to interpret the estimates as causal evidence.

6. Empirical results

6.1. HTs' administrative duties and students' academic achievements

Table 1 reports the coefficients from regressions of the HTs' administrative duties on students' average scores of three main subjects and scores taught by HT with the OLS estimation. All standard errors have been cluster-corrected at the class level. Control variables include HTs' characteristics and students and household characteristics. In particular, it should be noted that the grade is controlled based on the results of the balancing tests.

First, we examined the impact of HTs' administrative duties on students' average scores. In column (1), we only controlled for HTs' characteristics and school fixed effects, while all control variables are included in column (2). The results in these two columns show that the HTs' administrative duties have negative and statistically significant effects on the average scores of students. According to the preferred specification in column (2), the HTs' administrative duties reduce students' average scores by about 4.9388 points, which amounts to a 6.02% decrease in students' average test scores to the average value (81.9841).

Moreover, by limiting the sample of students whose HTs teach the main subjects and controlling the subject fixed effects additionally, we also investigated the effect on scores taught by HT.[9] The estimated coefficients of HTs' administrative duties are both negative in columns (3) and (4), but only significant at the level of 10%. As shown in column (4), for those students whose HTs have administrative duties, their scores taught by HT dropped by 3.7402 points. Relative to the average scores (83.6882), HTs' administrative duties reduced students' test scores in the subjects they taught by 4.47%.

From the results of the baseline regressions, the negative effect of HTs' administrative duties on students' academic achievement is comprehensive. Our estimates imply that it not only made the students' performance worse in the courses taught by the HT, but also had a negative effect on the overall academic achievement, and the impact on the latter is greater and more significant.

6.2. Robustness checks

We performed a number of robustness checks. First, we removed some extreme observations to ensure the generality of the sample, such as the worst and best-ranked schools in each county, students with highly educated parents,[10] very poor and very wealthy families, students with poor health status, schools in remote areas (Hu, 2018). Additionally, schools that assigned more administrative tasks to HTs may be fundamentally different from those with no tasks, which is also related to student achievement; so we further deleted schools without HTs in administrative positions. Considering that most private schools in China are headed by non-teachers, we excluded all students who attended private schools in our sample. Moreover, to avoid the influence of students' average scores by the administrative positions of other non-HTs, we excluded the sample of regular teachers who had administrative jobs. In addition, because of the effect of HTs' teaching experience on student outcomes and holding administrative positions, particularly less experienced teachers may interfere with the estimated results (Zheng & Sun, 2020), we eliminated the relevant samples of those HTs' experience is less than 5 years. Finally, we also investigated the impact caused by student's rearrangement. In the sample, some 9th grade classes experienced the student's rearrangement,[11] which may cause the biased results (Huang et al., 2021). Therefore, we excluded classes reclassified at the beginning of 9th grade for robustness checks. In Table A7 in the Appendix, no matter which extreme observations are eliminated in two Panels, the OLS estimates are qualitatively similar to the main estimates, although somewhat changes in magnitude.

Second, our identification strategy is mainly based on the quasi-natural experiment random student-classroom assignment, and the results of the balancing test proved that the key explanatory variable *Administrative duties* has no correlation with most class-level variables. Therefore, in the main specifications, we controlled for some variables at the HT, student, and household levels, with school fixed effects and cluster standard errors at the class level. We assessed the robustness of the findings to different specifications by adding some other control variables for characteristics of school, family, student, and HT (Feng & Li, 2016; Sun & Zhou, 2022), and the definition and descriptive statistics of these control variables are represented in Table A8 in the Appendix.

In Table A9 of the Appendix, on the basis of the main specifications, we added other control variables at the school, family, student, and HT levels in columns (1) to (4), respectively, and controlled all these control variables in column (5). In particular, in column (1), we added the proportion of HTs with administrative duties in school characteristics, which helps us capture schools' preference for HT appointments to administrative positions and the size of teachers with administrative positions in schools. Furthermore, in column (4), we controlled for some variables involving the time length of being HTs and the HT workload. On the one hand, it is possible to avoid challenges from the fact that the administrative duties of HTs may change with the time length of being an HT in a certain class (Zhong & Liu, 2021). On the other hand, we could determine that the negative impact of HTs' administrative duties on student achievement does not only come from the increase in workload (Li, 2018). As shown in Table A9, compared with baseline regressions in Table 1, the coefficients and significance of *Administrative duties* are virtually unchanged.

Third, we conducted placebo tests with students' previous academic performance in primary school as outcomes to check whether the results pick up some unobserved confounders (Hu, 2018; Huang et al., 2021). The questionnaire contains the students' achievement ranking in the class and self-reported learning difficulty degrees in Chinese, Math, and English in the sixth grade,[12] which can be used to measure students' previous academic performance. As reported in Table A10 in the Appendix, the placebo test results indicate that there is almost no association between HTs' administrative duties and students' previous academic performance, strengthening the interpretation that the estimates of baseline regressions do not pick up the effect of other unobserved confounders.

Fourth, since the CEPS only provides scores for one test, it may

[9] As mentioned in Table A1, the questionnaire only contains the scores of the three main subjects (Chinese, Math, and English). So, in columns (3) and (4) of Table 1, we excluded the sample whose HT teaches other subjects and control the differences between subjects by additional subject fixed effects.

[10] At least one of the parents has a master degree or above.

[11] Although these students were also randomly assigned, rearrangement would most likely replace HT. therefore, these student samples were excluded.

[12] Students' self-reported learning difficulty degrees in Math (Chinese or English) is divided into four categories (1 = very difficult; 2 = difficult; 3 = easy; 4 = very easy).

introduce contingency for our estimates. In addition, another concern is that although we controlled for several HT characteristics, we could not rule out the possibility that teachers with administrative positions may give unfairer scores. Thus, we attempted to replace our explained variable for further robustness tests. Two variables were *Easy to learn in main subjects* and *Low cognitive ability*.[13] Clearly, these two variables come from student self-assessment and CEPS team testing, respectively, and have nothing to do with teacher marking. More importantly, they are closely related to student performance (Gong et al., 2018; Zhao & Chen, 2022). In column (1) of Table A11 in the Appendix, we found that the coefficient on *Administrative duties* is significantly negative, while in column (2) it is significantly positive. These results suggest that our results remain convincing after excluding the potential interference from chance scores and teachers' unfair markings.

Last, we adopted the instrumental variable (IV) method to perform robustness tests. Although there is considerable confidence in our causal estimates in a framework of random student-classroom assignment, after all, teachers holding administrative positions may be selective.[14] Thus, we relied on the IV approach for further identification. More specifically, we adopted the method from Lewbel (2012) to construct IVs that depended on heteroskedasticity-based covariance restrictions (Baum & Lewbel, 2019; Wang & Cheng, 2022). Previous studies confirmed that the estimates obtained by this method are very close to those of traditional IV estimations (Umberger et al., 2015; Wang & Cheng, 2022). In columns (1) and (2) of Appendix Table A12, the results of the Breusch-Pagan test show that the original hypothesis of homoscedasticity is rejected, indicating that the IV constructed by Lewbel's (2012) method is appropriate in this paper. In addition, other statistics (e.g., Kleibergen-Paap rk LM statistic and Hansen J statistic) also imply that our IVs are valid. Consistent with the results of OLS estimates, the coefficients on *Administrative duties* keep significantly negative at the 1% level, indicating that the baseline regression results are robust.

6.3. Mechanism tests

In this subsection, we tested for multiple potential mechanisms, through which HTs' administrative duties reduced students' academic achievements. As noted in Section 3, these channels included teaching time and teachers' enthusiasm, teacher-student interaction and home-school connection, and peer effects.

6.3.1. Teaching time and teachers' enthusiasm

As emphasized in Section 3, administrative duties may squeeze HTs' teaching hours and weaken HTs' enthusiasm and teaching motivation, which would reduce their effectiveness and teaching quality (Kim, 2019), and then lead to a decline in students' academic achievements (Ballet et al., 2006; Caprara et al., 2006; Frenzel et al., 2019; Kunter et al., 2008; Skaalvik & Skaalvik, 2011). To confirm this mechanism, we selected three variables related to the HTs' teaching time, and four variables related to the enthusiasm as outcomes. The specific definitions and descriptions of these variables are shown in Panel A of Table A13 in the Appendix.

In Table 2, the results of columns (1) to (3) suggest that although administrative duties do not significantly affect the HTs' total number of lessons, they significantly reduce the proportion of preparing lessons and correcting homework in their working hours by 5.85% and 4.49% at the 5% level. In addition, the results from columns (4) to (7) indicate that administrative duties do not significantly affect the HTs' boredom with the work, responsibility and patience with students, and their assessment pressure, suggesting that there is no obvious relationship between administrative duties and HTs' enthusiasm. These results differ from some previous studies in Korea and Hong Kong (Kim, 2019; Kim & Park, 2014; Tsang & Kwong, 2017). The possible reason is that teachers in Chinese mainland may have gained more benefits from their administrative positions (Zheng & Sun, 2020), offsetting their job boredom. Taken together, our findings partly provide supportive evidence for previous expectations that administrative workload squeezes HTs' teaching time in their working hours, and the reducing teaching hours may translate into a decline in students' academic achievements.

6.3.2. Teacher-student interaction and home-school connection

Next, we considered the effects of administrative duties on teacher-student interaction and home-school connection. We selected four related variables to measure the teacher-student interaction and home-school connection, namely *HT-child communication*, *HT-parent communication*, *HT-child discussion in class*, *Familiarity with parents*, of which the details are in Panel B of Table A13 in the Appendix.

In Table 3, the results of column (1) indicate that administrative duties reduce the proportion of communication with students in HTs' working hours per week. Specifically, compared to the HTs without administrative duties, the HTs serving in administrative positions will spend 4.83% less of their working hours on communication with their students per week. However, based on results from the last three columns, we found that administrative duties have no significant impact on HT-parent communication, HT-child discussion in class, and familiarity with parents. These results confirm that although administrative duties did not affect the HTs' communication and familiarity with parents, and the frequency of discussion with students in teaching, they play an adverse role in communication between HTs and their students. Given that lacks of teacher-student interactions and teachers' guidance and supervision are detrimental to student performance (Bold et al., 2017; Duflo et al., 2012; He et al., 2021; Muralidharan & Sundararaman, 2011), we believed that teachers' administrative duties are likely to reduce student test scores by decreasing teacher-student interaction.

6.3.3. Peer effects

As noted earlier, the administrative workload may take up a lot of time and energy of HTs, making it difficult for them to take effective measures to intervene in class peer effects (Li et al., 2014). Even the lack of effective classroom management may lead to classroom discipline disruptions (Kim, 2019). Based on previous literature (e.g., Chu et al., 2015; Gong et al., 2018; Huang et al., 2021; Sacerdote, 2001), there are four variables measuring the peer effects, in terms of friends' bad behaviors and the average academic achievements of the class (other than the student observed). These variables include *Frequency of smoking and drinking of the student's friends*, *Frequency of going to net bar of the student's friends*, *Average scores of three subjects in the class*, *Average scores taught by HTs in the*

[13] *Easy to learn in main subjects* is a dummy variable, which is assigned to 1 if students feel at least two main subjects are easy to learn (student self-assessment); 0 otherwise. *Low cognitive ability* is also binary. It is equal to 1 if the standard score of the student's cognitive ability is ranked in the bottom 10% of the sample; otherwise, it is 0.

[14] This selectivity may come from self-selection by teachers and screening effects by school managers. Schools prioritize excellent, prestigious, or senior teachers for administrative positions who may improve student achievement. However, the baseline results indicated that administrative duties have a negative impact on student achievement; therefore, the school screening effect may not be a source of potential endogeneity.

Table 2
HTs' administrative duties and teaching time and teachers' enthusiasm.

	(1)	(2)	(3)	(5)	(6)	(7)	(8)
	Proportion of preparing lessons in working hours	Proportion of correcting homework in working hours	Total number of lessons	Boredom	Responsibility	Patience	Self-rated pressure
Administrative duties	−0.0585**	−0.0449**	−0.6553	0.0178	0.0276	0.0124	−0.1270
	(0.0276)	(0.0215)	(0.5703)	(0.1234)	(0.0310)	(0.0348)	(0.1390)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	8179	8204	8205	8268	8260	8258	8286

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses. Baseline control variables and school fixed effects are added in all regressions.

Table 3
HTs' administrative duties and communication with students and their parents.

	(1)	(2)	(3)	(4)
	HT-child communication	HT-parent communication	HT-child discussion in class	Familiarity with parents
Administrative duties	−0.0483*	−0.0640	0.0300	0.0851
	(0.0260)	(0.0600)	(0.1146)	(0.1466)
Control variables	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
N	8223	8194	8265	8257

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses. Baseline control variables and school fixed effects are added in all regressions.

class, of which the definitions are reported in Pane C of Table A13 in the Appendix.

In Table 4, the estimation results in columns (1) and (2) show that HTs' administrative duties have significant positive effects on the frequency of bad behaviors of the student's friends, including smoking and drinking, and going to the net bar, which may have negative peer effects on the students observed and interfere with their learning (Sacerdote, 2001; Stinebrickner & Stinebrickner, 2006). Additionally, as shown in columns (3) and (4), HTs' administrative duties also have a significant negative impact on the class average scores and the scores taught by HTs. Clearly, the decline in the average academic achievement of the classmates will further have a negative impact on the performance of the observed students through peer effects (e.g., Balestra et al., 2022; Duflo et al., 2011; Hu, 2018; Huang et al., 2021; Wang et al., 2018).

6.4. Heterogeneity analysis

We explored the heterogeneity of HTs' administrative duties on the students' academic achievements from four perspectives: types of administrative duties, the title of HTs, gender and age of HTs, and students' gender and *hukou*.

First, we analyzed the heterogeneity of types of HTs' administrative duties. Specifically, we divided the sample into four groups according to the type of HTs' administrative duties, including grade team leader, head of teaching and research section, school leader, and common HT without administrative duties, and then compared the first three types of positions with common HT. In Table 5, we found that the HT concurrently serving as the grade team leader or school leader had a significant negative effect on the average scores of students, because these two types of positions require a lot of time for management work (Qian & Walker, 2019). In contrast, the impact of serving as the head of teaching and research section on students' scores was not significant. The main reason is that this position is mainly engaged in administrative work related to teaching and research, which has a direct connection with the students' learning.

Second, according to the estimated results in Table 6, in terms of the impact on students' average scores, HTs' administrative work had significant effects among teachers with all titles. Specifically,

Table 4
HTs' administrative duties and peer effects.

	(1)	(2)	(3)	(4)
	Frequency of smoking and drinking of the student's friends	Frequency of going to net bar of the student's friends	Class average of students' average scores	Class average of scores taught by HT
Administrative duties	0.0209**	0.0260*	−4.7269***	−3.7932*
	(0.0101)	(0.0144)	(1.5923)	(2.2072)
Control variables	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
N	8226	8223	8286	6985

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses. Baseline control variables and school fixed effects are added in all regressions. Subject fixed effects are additionally controlled for in column (4). Class average of students' average scores and class average of scores taught by HT are both other than the observation.

Table 5
Heterogeneity effects by types of administrative duties.

	(1)	(2)	(3)
Panel A. Average scores			
Grade team leader	−5.2142***		
	(1.8272)		
Head of teaching and research section		−2.5521	
		(2.7828)	
School leaders			−5.5714*
			(3.1416)
N	7443	6989	7178
Panel B. Scores taught by HT			
Grade team leader	−4.5848		
	(2.9633)		
Head of teaching and research section		0.3072	
		(2.9622)	
School leader			−4.7646
			(3.3959)
N	6316	5935	6132

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses. Baseline control variables and school fixed effects are added in all regressions. Subject fixed effects are additionally controlled for in Panel B. Administrative duties mainly include three categories: grade team leader, head of teaching and research section, school leader (vice principal and director of teaching affairs).

Table 6
Heterogeneity effects by HTs' titles.

	(1)	(2)	(3)
	Junior title	Middle title	Senior title
Panel A. Average scores			
Administrative duties	−6.3876**	−9.9321***	−14.4741**
	(2.5330)	(3.0603)	(6.7960)
N	2971	3938	1377
Panel B. Scores taught by HT			
Administrative duties	−4.4881	−9.7241**	−9.5194**
	(4.0104)	(4.2551)	(4.1246)
N	2417	3290	1278

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses. Baseline control variables and school fixed effects are added in all regressions. Subject fixed effects are additionally controlled for in Panel B. In the CEPS, secondary school teacher's titles include three categories: junior title (non-grade, third-grade, and second-grade), middle title (first-grade), and senior title (senior and professor).

the results in Panel A suggest that HTs' administrative duties resulted in students' scores decreased by approximately 6.39, 9.93, and 14.47 points in three title grades, respectively. In Panel B, the coefficient of administrative duties is significant in columns (2) and (3), indicating that HTs with middle and senior grade titles holding administrative tasks instead had a greater negative impact on students' performance. Seniority and teaching experience are important criteria for Chinese junior high school teachers to evaluate job titles. The higher the title, the heavier their teaching and administrative work may be (Zheng & Sun, 2020; Zhong & Liu, 2021), which is not conducive to improving teaching quality.

Third, we tried to test the heterogeneity brought by gender and age of HTs (Antecol et al., 2015; Gong et al., 2018). In Table 7, the results of columns (1) and (2) indicate that for female HTs, administrative duties only affect students' average scores significantly, while coefficients are both significant in Panels A and B for male HTs. It can be explained by the fact that more administrative positions in Chinese junior high schools are also held by men (Ning & Zhang, 2018; OECD, 2015), which may crowd out their teaching quality. In addition, according to Chinese practice, teachers under the age of 35 are treated as young teachers and we used it as a dividing line to explore the HT age heterogeneity effect. As columns (3) and (4) shown, compared with middle-aged HTs, administrative workload has a very limited influence on youthful HTs, which may be mainly due to the more energy and time of young teachers.[15]

Last, in columns (1) and (2) of Table 8, we found that the HTs' administrative duties had a greater effect on boys' scores, while it had a limited effect on girls' achievement, especially for the scores taught by HT. The reason may be that girls are more serious in their studies at the compulsory education level, while boys are more susceptible to the bad behaviors of their surrounding peers (Gong et al., 2018; Luo et al., 2021; Xu & Li, 2018). As pointed out in the mechanism analysis, constrained by administrative duties, HTs may have to spend less time intervening in students' undesirable behaviors, which has a negative impact on boys' scores in particular. Moreover, as shown in columns (3) and (4), urban-rural heterogeneity is insignificant, suggesting that the negative impact of HT administrative positions on student achievement is combined and not influenced by the unequal distribution of teaching resources between urban and rural areas in China.

7. Conclusions and discussion

In some Asian countries, such as China, Korea, and Japan, many teachers in private and secondary schools are required to hold administrative duties, especially in China (Li, 2018; Zhong & Liu, 2021). However, few studies investigated the negative impact of HTs' administrative duties in Chinese middle schools (Sun & Zhou, 2022). By using the nationally representative data from CEPS, we employed the random student-classroom assignment as a quasi-natural experiment to identify the causal relationship between HTs' administrative duties and students' academic achievements and explored potential mechanisms.

We drew several conclusions from our findings. First, the baseline results suggested that HTs' administrative duties are detrimental to students' outcomes, which implied that more duties for teachers in Chinese secondary schools produced worse educational outcomes. Specifically, we found that HTs' administrative duties reduce students' average scores by about 4.94 points and scores taught by HTs by 3.74 points. The main results are supported by a series of robustness tests, such as removing extreme observations, using alternative specifications, placebo tests, and IV estimations. Second, we explored several potential mechanisms, and empirical results showed that administrative duties squeezed HTs' teaching time, reduced the teacher-student interaction, and produced undesirable class peer effects. However, HT burnout and decreased enthusiasm for teaching and reduction in home-school contacts were not mechanisms by which administrative work is detrimental to student achievement. Third, the results of heterogeneity analysis indicated that for male and middle-aged HTs and HTs with higher administrative positions and job titles, administrative duties had greater negative impacts on students' achievement. Moreover, boys are more likely to be negatively affected by the HTs' administrative duties, while urban-rural heterogeneity was insignificant.

At present, many studies examined the relationship between teacher characteristics and student academic achievement in compulsory education, but few focused on the effect of teachers' administrative duties (Kim & Park, 2014; Sun & Zhou, 2022; Zheng & Sun, 2020). Our study highlights the impact of teachers' teaching time and effort on student achievement, particularly as class managers of HTs (Ballet et al., 2006; Caprara et al., 2006; Frenzel et al., 2019; Kim, 2019). Moreover, this paper further enriches the

[15] According to various award criteria for youthful teachers in China, youthful teachers are usually under 35 years old.

Table 7
Heterogeneity effects by gender and age of HTs.

	(1)	(2)	(3)	(4)
	Female	Male	Youthful teachers	Middle-aged teachers
Panel A. Average scores				
Administrative duties	−5.9652**	−4.3773*	−4.7713*	−6.5120***
	(2.5588)	(2.5487)	(2.6752)	(2.1923)
N	5107	3179	3575	4711
Panel B. Scores taught by HT				
Administrative duties	−4.2277	−8.0857*	−4.0227	−8.6342**
	(2.7748)	(4.3951)	(3.8766)	(3.3691)
N	4447	2538	2619	3993

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses. Baseline control variables and school fixed effects are added in all regressions. Subject fixed effects are additionally controlled for in Panel B.

Table 8
Heterogeneity effects by students' gender and *Hukou*.

	(1)	(2)	(3)	(4)
	Girls	Boys	Rural	Urban
Panel A. Average scores				
Administrative duties	−3.8824**	−6.1001***	−4.9484***	−5.1215**
	(1.6073)	(1.7941)	(1.6794)	(2.2805)
N	4218	4068	4658	3628
Panel B. Scores taught by HT				
Administrative duties	−2.6566	−5.0495**	−4.1717*	−4.2557*
	(2.1788)	(2.2882)	(2.4847)	(2.5191)
N	3582	3403	3848	3137

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses. Baseline control variables and school fixed effects are added in all regressions. Subject fixed effects are additionally controlled for in Panel B.

research related to teacher workload and burden and their potential adverse effects (e.g., Kim, 2019; Park & Sung, 2013; Tsang & Kwong, 2017; Yang et al., 2009; Zhong & Liu, 2021). However, there is a recent study that is similar to our work (Sun & Zhou, 2022). In comparison, we focused more on the role of HTs and the negative effects of HTs' administrative duties. Also, in both theoretical and empirical terms, our study contributes to the addition of some new mechanisms. In addition, the rich heterogeneity analysis in our study provides insight into how different schools avoid the adverse impact of HTs' administrative duties. Taken together, our study implied that reducing teachers' administrative workload could be an important and effective strategy to improve students' achievements, and could have large benefits in terms of child development and human capital accumulation.

Our findings provide some insights into reducing teachers' workload and improving related educational policies in China and other Asian countries and regions. We should pay special attention to the workload of administrative duties on the HTs and its negative impacts on the students' academic achievements. Possible feasible measures include increasing the proportion of professional school administrators, having more young teachers and non-HT in administrative positions, and reducing the importance of administrative work in teacher evaluation. In recent years, with the implementation of the double reduction policy in China, the workload of Chinese teachers in compulsory education has further increased. Therefore, we need to pay more attention to the adverse effects of additional administrative work by teachers, especially for HTs.

There are some limitations of this paper and recommendations for further research. First, limited by the CEPS questionnaire, we could only know whether the HT holds administrative duties, but not the length of time the HT spends on administrative tasks and its proportion in the total working time. If the above-mentioned data are available, we can further quantify the impact of administrative work time on student performance. Second, the questionnaire does not include the scores of other subjects such as Physics and Chemistry, so we cannot examine the impact on these subjects, which may be more likely to be ignored and be negatively affected by the HT compared with the three main subjects. Third, our study focused on junior middle school students in the seventh and ninth grades, therefore, the applicability of our findings to primary school and high school students requires further validation.

Credit statement

Boou Chen: Data curation, Conceptualization, Formal analysis, Software, Methodology, Writing – original draft, Writing-Reviewing and Editing, Funding acquisition, **Chunkai Zhao**: Conceptualization, Methodology, Validation, Writing – original draft, Writing-Reviewing and Editing, Funding acquisition

Conflict of interest (COI) statement

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Declaration of competing interest

The authors have no conflicts of interest.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities, China (No.QCDC-2020-10; No. QCDC-2020-21), and the Tsinghua Rural Studies PhD Scholarship (202110).

Appendix

Table A1
Variable definition and descriptive statistics

Variables	Definitions	N	Mean	S.D.
Panel A. Scores				
Chinese	The mid-term scores of Chinese	8286	84.14	19.68
Math	The mid-term scores of Math	8286	80.38	31.26
English	The mid-term scores of English	8286	81.43	29.39
Average scores	Average scores of three main subjects (Chinese, Math, and English)	8286	81.98	24.07
Scores taught by HT	Scores of the subject taught by the HT (only include Chinese, Math and English teachers)	6985	83.69	27.31
Panel B. HT characteristics				
Administrative duties	= 1 if an HT had administrative duties; 0 otherwise	8286	0.196	0.397
Age of HT	The age of the HT	8286	37.24	6.932
Years of teaching	The years of teaching of the HT	8286	15.60	7.532
Gender of HT	0 = female; 1 = male	8286	0.384	0.486
Schooling of HT	0 = unschooled; 6 = primary school; 9 = junior middle school; 12 = senior high school; 15 = junior college; 16 = bachelor degree; 19 = master's degree	8286	15.70	0.535
Title	The title of the HT (0 = non-grade; 1 = third-grade; 2 = second-grade; 3 = first-grade; 4 = senior; 5 = professor)	8286	2.666	0.991
Panel C. Student and household characteristics				
Gender of student	0 = girl; 1 = boy	8286	0.491	0.500
Hukou	0 = rural; 1 = urban	8286	0.438	0.496
Ninth grade	= 1 if the student is in ninth grade; 0 otherwise	8286	0.468	0.499
Education expenditure	The logarithm of total family education expenditure (yuan)	8286	5.393	3.324
Schooling of father	0 = unschooled; 6 = primary school; 9 = junior middle school; 12 = senior high school; 15 = junior college; 16 = bachelor degree; 19 = master's degree	8286	10.29	3.085
Schooling of mother	0 = unschooled; 6 = primary school; 9 = junior middle school; 12 = senior high school; 15 = junior college; 16 = bachelor degree; 19 = master's degree	8286	9.474	3.549

Notes: In the sample, the subjects taught by the HT include four categories: Chinese, Math, English, and other subjects. However, the questionnaire only contains the scores of the first three main subjects. There is no score of the subject taught by the HT, if the HT taught other subjects; therefore, the observations of the variable *Scores taught by HT* are fewer than the others.

Table A2
Summary statistics of students' scores

	Full sample	HT with administrative duties	HT without administrative duties	Diff.
	Mean	Mean	Mean	T-tests
Chinese	84.1412	80.1988	85.1022	4.9034***
	(0.2162)	(0.4394)	(0.2453)	(0.5421)
Math	80.3793	75.5160	81.5648	6.0488***
	(0.3435)	(0.7468)	(0.3851)	(0.8627)
English	81.4318	72.8855	83.5151	10.6296***
	(0.3229)	(0.6652)	(0.3629)	(0.8050)
Average scores	81.9841	76.2001	83.3940	7.1939***
	(0.2645)	(0.5424)	(0.2987)	(0.6616)
N	8286	1624	6662	
Scores taught by HT	83.6882	76.2298	85.3712	9.1414***
	(0.3268)	(0.6987)	(0.3645)	(0.8361)
N	6985	1286	5699	

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively.

Table A3
Basic characteristics of classes in the sample

Class characteristics	Number of classes	Percent (%)
Panel A: HT with administrative duties		
Yes	69	20.06
No	275	79.94
Panel B: The subject taught by HT		
Chinese	102	26.74
Math	92	29.65
English	89	25.87
Other subjects	61	17.73
Panel C: Class size (number of students)		
<30	57	16.57
[30, 40)	73	21.22
[40, 50)	91	26.45
[50, 60)	75	21.80
> = 60	48	13.95
Panel D: Proportion of local students		
<30%	18	5.23
[30%, 50%)	16	4.65
[50%, 70%)	67	19.48
> = 70%	243	70.64
Total number of classes	344	100.00

Table A4
Descriptive statistics of class-level variables

Variables	Definitions	N	Mean	S.D.
Ratio of student gender	Ratio of boys in the class	344	0.489	0.121
Class size	Number of students in the class	344	43.96	13.98
Proportion of local students	Proportion of students with local *Hukou*	344	0.783	0.228
Proportion of rural students	Proportion of students with rural *Hukou*	344	0.490	0.305
Proportion of only-child	Proportion of students who are only-child in families	344	0.437	0.295
Ninth Grade	= 1 if the class is ninth grade; 0 otherwise	344	0.494	0.501
Proportion of teachers with bachelor degree	Proportion of teachers with bachelor degree (full-time)	344	0.468	0.346
Average years of teaching	Average years of teaching of teachers in the class	344	15.57	5.919
Senior (or professor) teacher	= 1 if there is at least one senior (or professor) teacher in the class; 0 otherwise	344	0.360	0.481
Proportion of parents with bachelor degree or above	Proportion of parents with bachelor degree or above in the class	344	0.153	0.200
Average education expenditure	Average household education expenditure of students in the class	344	5.641	2.221
Rate of extracurricular tutoring	Proportion of students participating in extracurricular tutoring in the class	344	7.169	9.206

Table A5
Comparison of class characteristics

	Full sample	HTs with administrative duties	HTs without administrative duties	Diff.
	Mean	Mean	Mean	T-tests
Ratio of student gender	0.4894	0.4795	0.4919	0.0124
	(0.0065)	(0.0162)	(0.0071)	(0.0163)
Class size	43.9622	41.1884	44.6582	3.4698*
	(0.7536)	(1.5737)	(0.8523)	(1.8754)
Proportion of local students	0.7831	0.7842	0.7828	−0.0014
	(0.0123)	(0.0298)	(0.0135)	(0.0307)
Proportion of rural students	0.4899	0.5286	0.4802	−0.0484
	(0.0165)	(0.0371)	(0.0184)	(0.0411)
Proportion of only-child	0.4374	0.4098	0.4443	0.0346
	(0.0159)	(0.0360)	(0.0178)	(0.0398)
Ninth grade	0.4942	0.5942	0.4691	−0.1251*
	(0.0270)	(0.0595)	(0.0301)	(0.0672)
Proportion of teachers with bachelor degree	0.4675	0.4300	0.4770	0.0470
	(0.0187)	(0.0408)	(0.0210)	(0.0466)
Average years of teaching	15.5727	15.9565	15.4764	−0.4802
	(0.3191)	(0.6839)	(0.3608)	(0.7977)
Senior or professor teacher	0.3605	0.2899	0.3782	0.0883
	(0.0259)	(0.0550)	(0.0293)	(0.0647)
Proportion of parents with bachelor degree or above	0.1527	0.1336	0.1575	0.0240
	(0.0108)	(0.0226)	(0.0122)	(0.0270)
Average education expenditure	5.6406	5.0727	5.7832	0.7105**
	(0.1198)	(0.3073)	(0.1273)	(0.2970)
Rate of extracurricular tutoring	0.3237	0.2521	0.3417	0.0896**
	(0.0155)	(0.0319)	(0.0175)	(0.0384)
N	344	69	275	

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively.

Table A6
Balancing tests for HTs' administrative duties at the class-level

Class characteristics	Administrative duties	School fixed effects	Number of Classes
Ratio of student gender	0.0029	Yes	344
	(0.0175)		
Class size	−0.8882	Yes	344
	(1.0311)		
Proportion of local students	−0.0033	Yes	344
	(0.0183)		
Proportion of rural students	−0.0048	Yes	344
	(0.0211)		
Proportion of only-child	0.0212	Yes	344
	(0.0164)		
Ninth grade	0.1857**	Yes	344
	(0.0907)		
Proportion of teachers with bachelor degree	−0.0069	Yes	344
	(0.0393)		
Average years of teaching	1.1591	Yes	344
	(0.8419)		
Senior (or professor) teacher	−0.0316	Yes	344
	(0.0696)		
Proportion of parents with bachelor degree or above	−0.0048	Yes	344
	(0.0198)		
Average education expenditure	0.1340	Yes	344
	(0.1782)		
Rate of extracurricular tutoring	−0.0061	Yes	344
	(0.0231)		

Notes: The numbers are the estimated coefficients on administrative duties, with class-level variables as outcomes, while school fixed effects were controlled. Robust standard errors are presented in parentheses. ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively.

Table A7
Robustness checks by removing some extreme observations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Exclude the worst and best ranked schools in each county	Exclude students with highly educated parents	Exclude very poor and very wealthy families	Exclude students with poor health status	Exclude schools in remote areas	Exclude schools where all HTs do not have administrative duties	Exclude private schools	Exclude classes where non-HTs have administrative duties	Exclude HTs with less than 5 years of teaching experience	Exclude classes reclassified at the beginning of 9th grade
Panel A. Average scores										
Administrative duties	−5.7009***	−4.8602***	−5.5131***	−5.1532***	−6.1427***	−4.7649***	−4.4982***	−3.2461*	−3.9836**	−4.9610***
	(1.7460)	(1.5702)	(1.5880)	(1.5661)	(1.8487)	(1.7075)	(1.6672)	(1.8056)	(1.6940)	(1.7656)
N	6046	8107	7937	7972	6978	4195	7460	6253	6848	7451
Panel B. Scores taught by HT										
Administrative duties	−4.0378*	−3.5387*	−4.3305**	−3.9029*	−5.5419**	−5.0193**	−4.3759**	−5.4519**	−4.2608*	−4.4099**
	(2.4026)	(2.1147)	(2.1205)	(2.0837)	(2.4633)	(2.1925)	(2.2021)	(2.1963)	(2.2463)	(2.2358)
N	4953	6823	6681	6734	5814	3506	6320	5212	5829	6365

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses. Baseline control variables and school fixed effects are added in all regressions. Subject fixed effects are additionally controlled in Panel B.

Table A8
Descriptive statistics of variables in robustness checks

Variables	Definitions	N	Mean	S.D.
Panel A. School characteristics				
Proportion of HTs with administrative duties	Proportion of HTs with administrative duties in the school (%)	8286	19.44	22.02
Financial expenditure	Financial expenditure per student on education (logarithm)	8063	6.422	1.391
Computers for teaching	Computers for teaching in each class (1 = yes; 0 = no)	8286	0.688	0.463
School reconstruction	The school has been rebuilt in the last 5 years (1 = yes; 0 = no)	8276	0.572	0.495
School ranking	The school ranking in the county (district) (1 = worst; 2 = not good; 3 = common; 4 = good; 5 = best)	8286	4.039	0.789
School location	Distance from city center (1 = very close; 2 = relatively close; 3 = normal; 4 = far; 5 = very far)	8286	2.638	1.534
Panel B. Family characteristics				
Father's Occupation	= 1 if the father is in a high-level job (personnel of national enterprises and institutions, enterprise managers, teachers, engineers, doctors, and lawyers); 0 otherwise (ordinary workers, self-employed, farmers, and unemployed)	8286	0.190	0.392
Computer available	Computer available at home (1 = yes; 0 = no)	8286	0.679	0.467
Internet available	Internet available at home (1 = yes; 0 = no)	8286	0.593	0.491
Home location	Distance from city center (1 = very close; 2 = relatively close; 3 = normal; 4 = far; 5 = very far)	8207	3.158	1.766
Living environment	Self-rated living environment (1 = good; 0 = bad)	8286	0.680	0.467
Panel C. Student characteristics				
Cognitive ability	Standardized scores of students' cognitive ability	8286	0.046	0.852
Kindergarten	= 1 if children have been to kindergarten; 0 otherwise	8286	0.798	0.401

Table A8 (*continued*)

Variables	Definitions	N	Mean	S.D.
Only child	Only child = 1; non-only child = 0	8286	0.416	0.493
Rank of child	Rank of child in your family	8286	1.351	0.532
Boarding	Boarding = 1; non-boarding = 0	8286	0.324	0.468
Panel D. Other HT characteristics				
Working hours	Working hours of HT in a week (h)	8223	49.53	19.37
Number of classes taught	Total number of classes currently taught by the HT	7887	2.297	1.214
Total number of lessons	Total number of HT's lessons in a week	8205	14.27	7.898
Time length as a HT	Time length since the first time as a HT (years)	8163	15.19	8.137
HT for the current class	Time length as a HT for the current class (years)	8271	2.778	1.682

Table A9
Robustness checks by adding more control variables

	(1)	(2)	(3)	(4)	(5)
	Add school characteristics	Add other family characteristics	Add other student characteristics	Add other HT characteristics	Add all variables
Panel A. Average scores					
Administrative duties	−5.2414***	−5.1268***	−4.4304***	−4.8154***	−4.2786***
	(1.5886)	(1.5765)	(1.5290)	(1.5997)	(1.5767)
N	8053	8207	8286	7664	7367
Panel B. Scores taught by HT					
Administrative duties	−3.7193*	−3.9488*	−3.5569*	−4.8926**	−4.4121*
	(2.1042)	(2.1289)	(2.1011)	(2.3588)	(2.4001)
N	6853	6930	6985	6426	6252

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses. Baseline control variables and school fixed effects are added in all regressions. Subject fixed effects are additionally controlled for in Panel B. The description of these added control variables is shown in Table A8 in the appendix.

Table A10
Placebo tests with the academic performance at 6th grade as outcomes

	(1)	(2)	(3)	(4)
	Achievement ranking in the class	Math	Chinese	English
Administrative duties	0.2163	−0.0374	0.0230	−0.0283
	(0.4551)	(0.0414)	(0.0264)	(0.0412)
Control variables	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
N	7875	8266	8259	8072

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses.

Table A11
Robustness checks by replacing the explained variable

	(1)	(2)
	Easy to learn in main subjects	Low cognitive ability
Administrative duties	−0.0158*	0.0231*
	(0.0084)	(0.0135)
Control variables	Yes	Yes
School fixed effects	Yes	Yes
N	8084	7821

Notes: Marginal effects are reported in the column. ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Standard errors clustered at the class level are reported in parentheses.

Table A12
IV estimations

	(1)	(2)
	Average scores	Scores taught by HT
Administrative duties	−4.1365***	- 3.7079***
	(0.6790)	(0.8865)
Control variables	Yes	Yes
School fixed effects	Yes	Yes
Subject fixed effects	No	Yes
Breusch-Pagan test	2641.58***	2801.41***
Kleibergen-Paap rk LM statistic	3599.990***	2788.738***
Cragg-Donald Wald F statistic	1437.243	4170.404
Hansen J statistic	920.711***	735.629***
N	8286	6985

Notes: ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are presented in parentheses.

Table A13
Descriptive statistics of variables in mechanism tests

Variables	Definitions	N	Mean	S.D.
Panel A. Teaching time and teachers' enthusiasm				
Proportion of preparing lessons in working hours	Hours of preparing lessons/working hours	8179	0.267	0.169
Proportion of correcting homework in working hours	Hours of correcting homework/working hours	8204	0.256	0.229
Total number of lessons	Total number of HT's lessons in a week	8205	14.27	7.898
Boredom	Boredom of the HT work (1 = never; 2 = occasionally; 3 = sometimes; 4 = always)	8268	2.280	0.804
Responsibility	Be responsible to students (1 = very irresponsible; 2 = irresponsible; 3 = common; 4 = responsible; 5 = very responsible)	8260	4.346	0.769
Patience	Be patient with students (1 = very impatient; 2 = impatient; 3 = common; 4 = patient; 5 = very patient)	8258	4.232	0.805
Self-rated pressure	Self-rated pressure from the school (1 = None; 2 = low; 3 = general; 4 = high; 5 = very high)	8286	3.648	0.974
Panel B. Teacher-student interaction and home-school connection				
HT-child communication	Proportion of communication with students in working hours (%)	8223	0.208	0.180
HT-parent communication	Frequency of communication with the students' parents in the last semester (1 = never; 2 = 1 times; 3 = 2–4 times; 4 = more than 4 times)	8194	2.128	1.012
HT-child discussion in class	Frequency of discussion with students in teaching (1 = never; 2 = occasionally; 3 = sometimes; 4 = often; 5 = always)	8265	3.838	0.782
Familiarity with parents	How many students' parents in the class are the HT familiar with (1 = almost none; 2 = few; 3 = about half; 4 = most; 5 = all)	8257	3.778	1.016
Panel C. Peer effects				
Frequency of smoking and drinking of the student's friends	1 = never; 2 = occasionally; 3 = frequently	8226	1.077	0.304
Frequency of going to net bar of the student's friends	1 = never; 2 = occasionally; 3 = frequently	8223	1.135	0.394
Class average scores	The class average test scores other than the student observed	8286	81.98	17.92
Class average score taught by HT	The class average test scores other than the student observed	6985	83.69	19.69

References

Antecol, H., Eren, O., & Ozbeklik, S. (2015). The effect of teacher gender on student achievement in primary school. *Journal of Labor Economics, 33*(1), 63–89.

Balestra, S., Eugster, B., & Liebert, H. (2022). Peers with special needs: Effects and policies. *The Review of Economics and Statistics, 104*(3), 602–618.

Ballet, K., Kelchtermans, G., & Loughran, J. (2006). Beyond intensification towards a schotarship of practice: Analysing changes in teachers' work lives. *Teachers and Teaching, 12*(2), 209–229.

Bauch, P., & Goldring, E. B. (1995). Parent involvement and school responsiveness: Facilitating the home–school connection in schools of choice. *Educational Evaluation and Policy Analysis, 17*(1), 1–21.

Baum, C. F., & Lewbel, A. (2019). Advice on using heteroskedasticity-based identification. *STATA Journal, 19*(4), 757–776.

Bogler, R. (2001). The influence of leadership style on teacher job satisfaction. *Educational Administration Quarterly, 37*(5), 662–683.

Bold, T., Filmer, D., Martin, G., Molina, E., Stacy, B., Rockmore, C., … Wane, W. (2017). Enrollment without learning: Teacher effort, knowledge, and skill in primary schools in Africa. *The Journal of Economic Perspectives, 31*(4), 185–204.

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology, 44*(6), 473–490.

Chen, J. (2016). Understanding teacher emotions: The development of a teacher emotion inventory. *Teaching and Teacher Education, 55*, 68–77.

Chen, X., Huang, Q., Rozelle, S., Shi, Y., & Zhang, L. (2014). Effect of migration on children's educational performance in rural China. In *China's economic development*. London: Palgrave Macmillan.

Chu, J. H., Loyalka, P., Chu, J., Qu, Q., Shi, Y., & Li, G. (2015). The impact of teacher credentials on student achievement in China. *China Economic Review, 36*, 14–24.

Droogenbroeck, F. V., & Spruyt, B. (2014). To stop or not to stop: An empirical assessment of the determinants of early retirement among active and retired senior teachers. *Research on Aging, 36*(6), 753–777.

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *The American Economic Review, 101*(5), 1739–1774.

Duflo, E., Hanna, R., & Ryan, S. P. (2012). Incentives work: Getting teachers to come to school. *The American Economic Review, 102*(4), 1241–1278.

Eberts, R., Hollenbeck, K., & Stone, J. (2002). Teacher performance incentives and student outcomes. *Journal of Human Resources, 37*(4), 913–927.

Farkas, G., Grobe, R. P., Sheehan, D., & Shuan, Y. (1990). Cultural resources and school success: Gender, ethnicity, and poverty groups within an urban school district. *American Sociological Review, 55*(1), 127–142.

Feng, H., & Li, J. (2016). Head teachers, peer effects, and student achievement. *China Economic Review, 41*, 268–283.

Frenzel, A. C., Taxer, J. L., Schwab, C., & Kuhbandner, C. (2019). Independent and joint effects of teacher enthusiasm and motivation on student motivation and experiences: A field experiment. *Motivation and Emotion, 43*(2), 255–265.

García-Gallego, A., Georgantzís, N., Martín-Montaner, J., & Pérez-Amaral, T. (2015). How) Do research and administrative duties affect university professors' teaching? *Applied Economics, 47*(45), 4868–4883.

Gerrard, J., & Farrell, L. (2014). Remaking the professional teacher: Authority and curriculum reform. *Journal of Curriculum Studies, 46*(5), 634–655.

Gigante, N. A., & Firestone, W. A. (2008). Administrative support and teacher leadership in schools implementing reform. *Journal of Educational Administration, 46*(3), 302–331.

Gong, J., Lu, Y., & Song, H. (2018). The effect of teacher gender on students' academic and noncognitive outcomes. *Journal of Labor Economics, 36*(3), 743–778.

Good, T. L. (1987). Two decades of research on teacher expectations: Findings and future directions. *Journal of Teacher Education, 38*(4), 32–47.

Gutman, M. (2018). Ethical dilemmas in senior teacher educators' administrative work. *European Journal of Teacher Education, 41*(5), 591–603.

Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *The American Economic Review, 100*(2), 267–271.

Hargreaves, A. (1992). Time and teachers' work: An analysis of the intensification thesis. *Teachers College Record, 94*, 87–108.

Hargreaves, A. (2002). Teaching and betrayal. *Teachers and Teaching, 8*(3), 393–407.

Hendrickx, M. M., Mainhard, M. T., Boor-Klip, H. J., Cillessen, A. H., & Brekelmans, M. (2016). Social dynamics in the classroom: Teacher support and conflict and the peer ecology. *Teaching and Teacher Education, 53*, 30–40.

He, Y., Zhang, Y., Ma, X., & Wang, L. (2021). Does private supplementary tutoring matter? The effect of private supplementary tutoring on mathematics achievement. *International Journal of Educational Development, 84*, Article 102402.

Hu, F. (2018). Migrant peers in the classroom: Is the academic performance of local students negatively affected? *Journal of Comparative Economics, 46*(2), 582–597.

Huang, B., Lu, H., & Zhu, R. (2021). Disabled peers and student performance: Quasi-experimental evidence from China. *Economics of Education Review, 82*, Article 102121.

Huang, B., & Zhu, R. (2020). Peer effects of low-ability students in the classroom: Evidence from China's middle schools. *Journal of Population Economics, 33*(4), 1343–1380.

Jensen, D. A. (2006). Using newsletters to create home–school connections. *The Reading Teacher, 60*(2), 186–190.

Kashy-Rosenbaum, G., Kaplan, O., & Israel-Cohen, Y. (2018). Predicting academic achievement by class-level emotions and perceived homeroom teachers' emotional support. *Psychology in the Schools, 55*(7), 770–782.

Keller, M. M., Hoy, A. W., Goetz, T., & Frenzel, A. C. (2016). Teacher enthusiasm: Reviewing and redefining a complex construct. *Educational Psychology Review, 28*(4), 743–769.

Kim, K. N. (2019). Teachers' administrative workload crowding out instructional activities. *Asia Pacific Journal of Education, 39*(1), 31–49.

Kim, K. N., & Park, J. S. (2014). The relationship between teachers' administrative

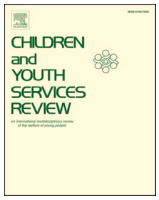
workload and classroom preparation time for instruction: Identifying a causal relationship and a moderating effect by school establishment type. *Korean Society for the Study of Sociology of Education, 24*(1), 29–54.
Kingdon, G., & Teal, F. (2010). Teacher unions, teacher pay and student performance in India: A pupil fixed effects approach. *Journal of Development Economics, 91*(2), 278–288.
Kunter, M., Tsai, Y. M., Klusmann, U., Brunner, M., Krauss, S., & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. *Learning and Instruction, 18*(5), 468–482.
Lasky, S. (2005). A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teaching and Teacher Education, 21*(8), 899–916.
Lavy, V., Paserman, M. D., & Schlosser, A. (2012). Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *Economic Journal, 122*(559), 208–237.
Lazarides, R., Buchholz, J., & Rubach, C. (2018). Teacher enthusiasm and self-efficacy, student-perceived mastery goal orientation, and student motivation in mathematics classrooms. *Teaching and Teacher Education, 69*, 1–10.
Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics, 30*(1), 67–80.
Li, Y. (2018). A study on job satisfaction of "double-shouldered" teachers in primary and secondary schools. *Journal of Teaching and Management, 28*(4), 31–33.
Li, T., Han, L., Zhang, L., & Rozelle, S. (2014). Encouraging classroom peer interactions: Evidence from Chinese migrant schools. *Journal of Public Economics, 111*, 29–45.
Liu, S. (2012). The influences of school climate and teacher compensation on teachers' turnover intention in China. *Educational Psychology, 32*(5), 553–569.
Liu, J. (2021). Exploring teacher attrition in urban China through interplay of wages and well-being. *Education and Urban Society, 53*(7), 807–830.
Liu, J. Q., & Barnhart, R. (1999). Homeroom teacher and homeroom class: The key to classroom management in China's schools. *The Educational Forum, 63*(4), 380–384.
Loyalka, P., Sylvia, S., Liu, C., Chu, J., & Shi, Y. (2019). Pay by design: Teacher performance pay design and the distribution of student achievement. *Journal of Labor Economics, 37*(3), 621–662.
Luo, R., Teng, K., Xu, L., & Zhu, Y. (2021). A decomposition of student achievement gap by gender in China: Evidence from random class assignment. *International Journal of Educational Research, 106*, Article 101721.
Madigan, D. J., & Kim, L. E. (2021). Does teacher burnout affect students? A systematic review of its association with academic achievement and student-reported outcomes. *International Journal of Educational Research, 105*, Article 101714.
Manion, M. (1985). The cadre management system, post-Mao: The appointment, promotion, transfer and removal of party and state leaders. *China Quarterly, 102*, 203–233.
McCarthey, S. J. (2000). Home–school connections: A review of the literature. *Journal of Educational Research, 93*(3), 145–153.
Mullola, S., Jokela, M., Ravaja, N., Lipsanen, J., Hintsanen, M., Alatupa, S., & Keltikangas-Järvinen, L. (2011). Associations of student temperament and educational competence with academic achievement: The role of teacher age and teacher and student gender. *Teaching and Teacher Education, 27*(5), 942–951.
Muralidharan, K., & Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from India. *Journal of Political Economy, 119*(1), 39–77.
Ning, B., & Zhang, M. (2018). The change of gender structure and leadership of principals in junior high schools: An exploration in Shanghai. *Global Education, 47*(2), 31–38.
Núñez, J. C., Suarez, N., Rosário, P., Vallejo, G., Cerezo, R., & Valle, A. (2015). Teachers' feedback on homework, homework-related behaviors, and academic achievement. *Journal of Educational Research, 108*, 204–216.
OECD. (2015). *Shanghai country note: Results from TALIS 2013-2014.* http://www.oecd.org/edu/school/TALIS-2014-coutry -note-Shanghai.pdf.
Okumura, S. (2017). Homeroom teachers or specialist teachers?: Considerations for the workforce for teaching English as a subject at elementary schools in Japan. *Asian Journal of Education and Training, 3*(1), 1–5.
Park, M., & Sung, Y. K. (2013). Teachers' perceptions of the recent curriculum reforms and their implementation: What can we learn from the case of Korean elementary teachers? *Asia Pacific Journal of Education, 33*(1), 15–33.
Qian, H., & Walker, A. (2019). Reconciling top-down policy intent with internal accountability: The role of Chinese school principals. *Educational Assessment, Evaluation and Accountability, 31*(4), 495–517.
Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.
Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review, 94*(2), 247–252.
Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics, 116*(2), 681–704.
Shi, Q., & Leuwerke, W. C. (2010). Examination of Chinese homeroom teachers' performance of professional school counselors' activities. *Asia Pacific Education Review, 11*(4), 515–523.
Skaalvik, E. M., & Skaalvik, S. (2011). Teacher job satisfaction and motivation to leave the teaching profession: Relations with school context, feeling of belonging, and emotional exhaustion. *Teaching and Teacher Education, 27*(6), 1029–1038.
Stinebrickner, R., & Stinebrickner, T. R. (2006). What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics, 90*(8–9), 1435–1454.
Sun, Y., & Zhou, X. (2022). The effect of teacher's concurrent administrative position on students' academic outcomes: Evidence and mechanisms. *China Economic Review, 71*, Article 101735.
Tan, C. (2019). PISA and education reform in Shanghai. *Critical Studies in Education, 60*(3), 391–406.
Tsang, K. K., & Kwong, T. L. (2017). Teachers' emotions in the context of education reform: Labor process theory and social constructionism. *British Journal of Sociology of Education, 38*(6), 841–855.
Umberger, W. J., He, X., Minot, N., & Toiba, H. (2015). Examining the relationship between the use of supermarkets and over-nutrition in Indonesia. *American Journal of Agricultural Economics, 97*(2), 510–525.
Urbina-Garcia, A. (2019). Preschool transition in Mexico: Exploring teachers' perceptions and practices. *Teaching and Teacher Education, 85*, 226–234.
Wang, H., & Cheng, Z. (2022). Kids eat free: School feeding and family spending on education. *Journal of Economic Behavior & Organization, 193*, 196–212.
Wang, H., Cheng, Z., & Smyth, R. (2018). Do migrant students affect local students' academic achievements in urban China? *Economics of Education Review, 63*, 64–77.
Wang, J., & Rao, N. (2019). Classroom goal structures: Observations from urban and rural high school classes in China. *Psychology in the Schools, 56*(8), 1211–1229.
Xie, L., & Yang, L. (2021). The influence of perceptions of promotion opportunities on job performance and its mechanisms: A case study of Chinese junior civil servants. *Journal of Contemporary China, 30*(127), 118–135.
Xu, D., & Li, Q. (2018). Gender achievement gaps among Chinese middle school students and the role of teachers' gender. *Economics of Education Review, 67*, 82–93.
Yang, X., Ge, C., Hu, B., Chi, T., & Wang, L. (2009). Relationship between quality of life and occupational stress among teachers. *Public Health, 123*(11), 750–755.
Zhao, C., & Chen, B. (2022). Parental migration and non-cognitive abilities of left-behind children in rural China: Causal effects by an instrumental variable approach. *Child Abuse & Neglect, 123*, Article 105389.
Zheng, L., & Sun, L. (2020). Does the part-time administrative weaken the teaching work? A psm estimate based on the data of 730 junior high school teachers. *Contemporary Educational Science, 35*(11), 52–58.
Zhong, J., & Liu, Y. (2021). The paradox of reducing teachers' burden: The dilemma of de-professionalization and the trap of re-specialization. *Tsinghua Journal of Education, 42*(6), 80–90.

Children and Youth Services Review 164 (2024) 107879

Contents lists available at ScienceDirect

Children and Youth Services Review

journal homepage: www.elsevier.com/locate/childyouth



The long-term effect of childhood left-behind experiences on entrepreneurial activities in adulthood: Empirical evidence from China

Binhui Wei [a], Runtao Wang [a], Xiakaifeng Lin [a], Chunkai Zhao [a,b,*], Mingzhong Luo [a]

[a] College of Economics and Management, South China Agricultural University, No.483, Wushan Street, Tianhe District, Guangzhou, Guangdong Province, China
[b] Agricultural and Rural Policy and Reform Innovation Laboratory, South China Agricultural University, No.483, Wushan Street, Tianhe District, Guangzhou, Guangdong Province, China

ARTICLE INFO

Keywords:
Left-behind experiences
Childhood
Entrepreneurial activities
Long-term effect
China

ABSTRACT

Although the issue of left-behind children has received academic attention, little is known about the potential long-term effect. This study was conducted in China, aiming to explore the impact of left-behind experiences in childhood on entrepreneurial activities in adulthood. The study design was quantitative and we used secondary data with samples from the China Labor-force Dynamic Survey. To mitigate potential endogeneity problems, we adopted the conditional mixed process approach with instrumental variable. The empirical results show that left-behind experiences decreased the likelihood of adulthood entrepreneurship, which is supported by some robustness tests. This negative impact can be partially explained by disadvantages in human capital and social networks as well as risk aversion. Moreover, we found that the exogenous shocks from *hukou* transfer and WTO accession weaken the negative impact of left-behind experiences and lack of entrepreneurship can further explain the income penalization effect of left-behind experiences. Our study provides new evidence for the life course theory and reveals the long-term adverse effects of left-behind experiences in childhood. Also, we shed light on safeguarding the welfare of left-behind children and young people with left-behind experiences.

1. Introduction

The phenomenon of left-behind children remains a prominent social issue in China (Nguyen, 2016; Meng & Yamauchi, 2017; Zhao & Chen, 2022). According to the newly released "China's Population and Children in 2020: Facts and Bases," jointly published by the National Bureau of Statistics of China, the United Nations International Children's Emergency Fund, and the United Nations Population Fund, in 2020, 108 million children in China were not able to reside permanently with their parents, and the number of left-behind children in rural areas was 41.77 million, accounting for 37.9 % of all rural children. The issue of left-behind children has long been the focus of scholars. Numerous studies recognized that left-behind children are disadvantaged in multiple fields, such as mental health (e.g., Xiao et al., 2019; Wang et al., 2019; Wu et al., 2015), cognitive and non-cognitive abilities (Zhang et al., 2014; Zhao & Chen, 2022), physical development (e.g., Chang et al., 2019; Li et al., 2015; Wu & Guo, 2020; Zhou et al., 2015), academic outcomes (e.g., Qiu et al., 2024; Wang et al., 2019; Zhao et al., 2014), and quality of life (Chen & Chan, 2016; Guo et al., 2019). There are even some studies that point to a greater likelihood of delinquency and delinquent behaviors among left-behind children (Chen & Chan, 2016; Zhang et al., 2019).

In recent years, some scholars have begun to focus the long-term effect of the left-behind experience. They found that adults with left-behind experiences continued to face psychological and behavioral problems (Li M et al., 2021; Liang et al., 2020; Sun et al., 2017; Yang et al., 2023; Zhang et al., 2021). More importantly, a few studies have focused on the relationship between left-behind experiences and labor market outcomes in adulthood (Lyu & Chen, 2019; Wang et al., 2021). Yet there is little literature on the impact of the left-behind experience on entrepreneurial activities in adulthood. Clearly, entrepreneurial activities are engines for promoting labor force employment, stimulating innovation, and driving economic growth. For individuals and households, entrepreneurship is an important way to alleviate poverty, promote income growth, and improve household welfare (Deng et al., 2020; Lyons & Zhang 2017; Stephan et al., 2023).

We focused on the long-term impact of special left-behind experiences on entrepreneurship for two main reasons. From a theoretical

* Corresponding author at: College of Economics and Management, South China Agricultural University, No.483, Wushan Street, Tianhe District, Guangzhou, Guangdong Province, China.
E-mail address: ckz_zhao@163.com (C. Zhao).

https://doi.org/10.1016/j.childyouth.2024.107879
Received 27 February 2024; Received in revised form 20 August 2024; Accepted 21 August 2024
Available online 30 August 2024
0190-7409/© 2024 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

perspective, the life course theory reveals the long-term impact of structural, social, and cultural factors in the course of life development on an individual's life (Elder, 1998; George, 2014). Childhood experiences are the foundation for development throughout the life cycle, or some special childhood experiences are decisive for many important decisions in adulthood (Schoon & Duckworth, 2012; White & Wu, 2014). Thus, particular left-behind experiences may be linked to entrepreneurial behaviors in adulthood. Moreover, from a practice perspective, unlike previous literature that were interested in some sudden and singular events, such as natural disasters and public health events (e.g., Churchill et al., 2021; Yi et al., 2022), left-behind experiences tend to last longer in childhood and are more prevalent (Yang et al., 2023; Zhang et al., 2021). Exploring the impact of left-behind experiences on entrepreneurship may have stronger generalizability and policy implications. Taken together, in this study, in terms of the left-behind experiences, we attempted to offer new insights into the entrepreneurial effects of adverse childhood experiences.

There are some interesting findings in this study which contribute to the existing research. First, based on a quantitative research design, we confirmed that childhood left-behind experiences inhibited entrepreneurial activities in adulthood in China, which provided an important addition to the literature on the life course theory and supports the view that adverse childhood experiences are detrimental to entrepreneurship (Prakash & Kumar, 2021; Zhao & Li, 2022; Yu et al., 2023). Prior research mainly dealt with the long-term effects of specific event shocks (Cheng et al., 2021; Churchill et al., 2021; Yi et al., 2022) and different types of childhood adversities (Zhao & Li, 2022; Yu et al., 2023). In contrast to these studies, we paid particular attention to left-behind experiences in China, offering new evidence that adverse childhood experiences inhibited entrepreneurial activities in adulthood.

Second, based on the particular perspective of entrepreneurial activities, our study further enriches the research related to the long-term effect of the left-behind experience in childhood. As mentioned earlier, several psychological studies suggested that the left-behind experience may lead to mental problems or illnesses for adults, but only a little research addresses labor market outcomes and economic welfare (Lyu & Chen, 2019; Wang et al., 2021). Clearly, this research revealed the potential vulnerability of left-behind children in future business activities. Importantly, our study helps to clarify the potential controversies of the left-behind experience in childhood (e.g., Aregbeshola, 2022; Zhang et al., 2019). At least at the adulthood entrepreneurial level, we provided evidence that the potential positive impact of left-behind experiences is far less than its negative impact in China.

Third, we explored the mechanisms by which the left-behind experience in childhood influenced entrepreneurial behaviors in adulthood. Prior literature on left-behind experiences in childhood and later labor market outcomes has focused primarily on human capital mechanisms such as education, cognitive abilities, skills, and physical and mental health (Lyu & Chen, 2019; Wang et al., 2021), ignoring the potential role of social networks and risk appetite. Results of the mediation effect show that left-behind experiences in childhood lead to deficiencies in social network constructions and the development risk averse traits, which further inhibits entrepreneurial activities. Moreover, heterogeneity effect results show that left-behind experiences at all school ages were found to have a negative effect on entrepreneurship, but this effect was more pronounced for the younger people. These results provide potential policy insights on how to minimize the long-term negative effects of the left-behind experience.

Last, we further investigated some of the moderating effects, showing that *hukou* conversion and WTO accession helped weaken the negative impact of left-behind experiences in childhood on entrepreneurial behaviors in adulthood. These findings further validate the life course theory that emphasizes that some special turning points may have a milestone effect, which in turn eliminates the impact of the original experience (George, 2014; Zhao & Guo, 2022). More importantly, we found that entrepreneurial behaviors even explained the income-penalizing effect of left-behind experiences, revealing the socioeconomic status and well-being conditions of adults with left-behind experiences in China.

2. Theoretical mechanisms and hypothesis development

In this section, we addressed the theoretical mechanisms by which childhood left-behind experiences influence entrepreneurial behaviors, including human capital, social networks, and risk appetite.

2.1. Human capital

The life course theory implies a long-run effect of early life shocks on later economic behaviors, which has well documented in sociology, psychology, and other social science fields. These outcomes, including earnings, labor market performance, business activities, and well-beings, all provide strong support (e.g., Cheng et al., 2021; Kwon & Ruef, 2017; Wang et al., 2021; Zhao & Guo, 2022); thus, human capital status in adulthood may also be influenced by particular experiences in childhood (Lyu & Chen, 2019). In addition, health and education outcomes are typically shaped in a cumulative manner, in which current and previous investments are combined with an individual's inherent abilities and endowments (Marvel et al., 2016; Qin & Kong, 2021). Thus, human capital in adulthood may depend heavily on investments and accumulations made during childhood (Caucutt & Lochner, 2020; Wang et al., 2021). For left-behind children, their human capital status in adulthood is closely related to left-behind experiences.

Existing studies mainly analyzed the impact of parental migration on the human capital of left-behind children from two aspects: the remittance effect and the separation effect. On the one hand, from the remittance effect, parents migrate to urban areas wages can get higher economic returns, which in turn increase family income. Therefore, through the remittance effect, parental migration can effectively increase the household investment in nutrition, medical care, and health (Ajefu & Massacky, 2023; Cuadros-Meñaca, 2020), which promotes children human capital development (Aregbeshola, 2022). In addition, by migration to urban areas, parents may acquire more health knowledge and parenting skills, which in turn improve their children's academic performance and health (Zheng et al., 2023).

On the other hand, in terms of separation effects, parental migration may be detrimental to children human capital development, leading to poorer academic performance and health outcomes for children left behind. First, most left-behind children are cared for by their grandparents, who are at a disadvantage in terms of education and health concepts (McFarland et al., 2017; Wu & Guo, 2020; Zhao & Chen, 2022), and grandparents may even spoil these left-behind children. Second, there is evidence that left-behind children are more burdened with household chores, which can affect their physical development and delay learning, especially in rural areas (Chang et al., 2019; Li et al., 2015). Third, owing to the lack of parental supervision, the probability of left-behind children becoming addicted to the Internet and developing bad habits is elevated (Chen & Chan, 2016; Qiu et al., 2024; Zhang et al., 2019). Moreover, it is worrisome that previous studies implied that left-behind children are more likely to suffer from psychological problems or illnesses, such as depression, low self-esteem, sensitivity, emotional outbursts and aggressive behaviors (e.g., Chang et al., 2019; Liang et al., 2020; Wang et al., 2019; Xiao et al., 2019). Worse, such psychological imprints may persist into adulthood (Li X et al., 2021; Sun et al., 2017; Zhang et al., 2021).

As highlighted earlier, most of the literature based on the Chinese context endorsed that the separation effect plays a major role, emphasizing the negative effect of left-behind experiences on human capital. Further, human capital accumulation is one of the necessary elements to become an entrepreneur. A large body of previous studies confirmed that better education, physical fitness, emotional stability, and strong psychological capital contribute to entrepreneurial success (e.g., Ahn &

Winters, 2023; Baron et al., 2016; Marvel et al., 2016; Van Ness & Seifert, 2016), especially in China (Cheng & Smyth, 2021; Qin & Kong, 2021; Zhao & Li, 2022; Zou et al., 2016). Thus, we put forward the first hypothesis:

Hypothesis 1 Left-behind experiences in childhood is detrimental to entrepreneurial activities in adulthood owing to deficiencies in human capital, such as impaired physical and mental health, and inadequate education and skills.

2.2 Social networks

The social network is a relatively stable relationship system between individuals by social interactions. In economic and management studies, social networks are often considered as components of social capital, which profoundly influence people's economic behaviors (e.g., Nason & Bothello, 2023; Gee et al., 2017; Van Burg et al., 2022). More importantly, social networks are often associated with entrepreneurial behaviors (Leyden et al., 2014; Zhao et al., 2024), especially in typical relations-based societies like China (Burt, 2019; Zhao & Li, 2022). Social networks may promote entrepreneurship in several ways. First, social networks can accelerate the dissemination of entrepreneurial information and knowledge and help potential entrepreneurs with more market information, which reduces the information asymmetry (Leyden et al., 2014; Zhao et al., 2024). Second, social networks play a financing function by providing informal credit to alleviate entrepreneurial capital constraints (Banerjee et al., 2024; Gee et al., 2017). Especially in areas with less developed financial markets, the social network is one of the central factors for entrepreneurial success (Van Burg et al., 2022; Zhao & Li, 2022). Third, social networks reduce the search costs for potential entrepreneurs to find suitable partners to share entrepreneurial risks (Burt, 2019). Moreover, social networks also provide informal institutional protection (Mumtaz, 2022; Nason & Bothello, 2023), compensating for the additional risk of deficiencies in the formal system.

However, the left-behind experience can be a negative signal for the construction of high-quality social networks. As mentioned earlier, there is a greater likelihood that left-behind children have character deficits and psychological distress (Guo et al., 2021; Xiao et al., 2019; Wang et al., 2019; Wu et al., 2015). They are more likely to be insecure and emotionally unstable owing to parental absence (Chang et al., 2019; Liang et al., 2020). In social interactions, left-behind children are more inclined to adopt a negative or violent approach (Wen et al., 2021; Zhang et al., 2021), which exacerbates the social isolation (Liu & Villa, 2020; Yang et al., 2021). Second, there is evidence that emotional supports from family and friends are directly related to children's interpersonal deficits (Li M et al., 2021). Apparently, owing to the lack of emotional support from parents, it is difficult for left-behind children to confide in or seek help from their parents when they encounter negative events in their daily interactions (Li M et al., 2021; Nguyen, 2016). Worse, left-behind children may have difficulty integrating into the collective life of the school and the community and become gradually marginalized (Sun et al., 2017). In addition, the risks of being left behind include violent behaviors, Internet addiction, and even infectious diseases (Chen & Chan, 2016; Wen et al., 2021; Zhang et al., 2019), and these risky elements create negative feedback through social networks, hindering the formation of social trust (Guo et al., 2019; Zhao & Li, 2022).

Childhood is a critical period for character formation and initial social network construction (Li X et al., 2021; Sun et al., 2017), and these particular resource endowments can profoundly affect the quality of social networks in adulthood (Leung et al., 2022; Zhao & Li, 2022). As a result, adults with left-behind experiences may have relative deficits in social network constructions, especially in high-quality social networks related to entrepreneurship. In summary, we proposed our second hypothesis:

Hypothesis 2 Owing to disadvantages in social networks, left-behind experiences in childhood are unlikely to promote entrepreneurship.

2.3 Risk appetite

Left-behind experiences may indirectly affect entrepreneurial activities in adulthood by influencing risk appetite. There is a strong link between the left-behind experience and risk appetite in adulthood for the following reasons. First, the psychological shock created by left-behind experiences is likely to create an irrational precautionary motive, leading adults more sensitive to the consequences of risky events (Benmelech & Frydman, 2015; Bernile et al., 2017), which makes their entrepreneurial decisions more guarded. Perhaps adults with left-behind experiences tend to take some conservative approach to risk, as they experienced in their childhood (Zhang et al., 2021; Zhao & Li, 2022). Second, previous research suggested that negative experiences may cause individuals to have a fearful emotional response to any negative events (Cassar et al., 2017), as risk attitude can be shifted between domains (Xu & Cheng, 2021). Thus, left-behind experiences are likely to lead adults to behave in ways that are more risk-averse. Third, realistically, parental migration is more of a helpless choice, which means that most families with left-behind children are not wealthy or have even fallen into poverty (Zhao & Chen, 2022). As a result, left-behind children may recognize the difficulties of life too early, making them more likely to become conservative in adulthood (Zhao & Li, 2022).

Entrepreneurs are inherently brave, often facing myriad market and non-market uncertainties. Their propensity for risk-taking sets them apart, a sentiment echoed in academic research. Kihlstrom & Laffont (1979) surmised that individuals with a higher tolerance for risk lean into entrepreneurship, whereas the more risk-averse find solace in employment. This idea was supported by subsequent studies, showing that a predilection for risk can fuel entrepreneurial endeavors (Rosado-Cubero et al., 2022; Zhao et al., 2024). Moreover, embracing risks in entrepreneurship can usher in benefits, from expanding business scales and fostering innovation to improving long-term performance (Putniņš & Sauka, 2020; Rosado-Cubero et al., 2022). Taken together, we developed the fourth hypothesis:

Hypothesis 3 By developing risk aversion traits, left-behind experiences in childhood are likely to hinder entrepreneurial behaviors in adulthood.

3 Methodology

3.1. Data and sample

We used secondary data in this study, namely the China Labor-force Dynamic Survey (CLDS), which is a nationally representative survey conducted by the Center for Social Survey of Sun Yat-sen University in 2018. The CLDS was granted ethical approval from the Research Ethics Committees of Sun Yat-sen University. All subjects signed written informed consent before the interview. This database is publicly available and we strictly adhere to the norms and ethical guidelines for the use of this database.

We adopted the fourth wave of CLDS.[1] In this wave, the sample covered 28 provinces (autonomous regions or municipalities) in Mainland China, including 368 communities or villages, 13,501 households, and 16,537 individuals. Besides its main focus on the status and

[1] The CLDS is a biennial survey since 2012. However, information about the left-behind experience was recorded only in the fourth round of surveys in 2018.

transition of Chinese labor force, this database provides rich information of entrepreneurial activities. More importantly, this wave of CLDS also provides information on whether post-1980 respondents[2] had left-behind experiences at different school age (Liu et al., 2022). Taken together, it has a prominent advantage to explore the effect of the left-behind experience in childhood on entrepreneurial behaviors in adulthood in China.

Moreover, we further screened the sample as follows. First, we excluded observations that were younger than 18 years old in the survey year, as we cared about entrepreneurial activities in adulthood (Saud & Margono, 2021; Wang et al., 2021). Second, we retained only observations that were born with a rural *hukou* because rural left-behind children are dominant in China and previous literature also has focused on the rural left-behind children (e.g., Chang et al., 2019; Chen & Chan, 2016; Li et al., 2015; Liu & Villa, 2020; Zhao & Chen, 2022).

3.2. Variables

3.2.1. Entrepreneurial activities

In this study, we measured individual entrepreneurship activities in terms of employment behavior, implying that, unlike wage earners, the act of establishing a new business is defined as entrepreneurship (Cheung & Tsang, 2015; Yang & Zhang, 2023; Zhao & Li, 2022). Thus, according to the CLDS database and the established conventions in prior research (e.g., Cheng et al., 2021; Yang & Zhang, 2023; Zhao & Li, 2022), we classified respondents in employer roles as entrepreneurs, denoting the variable *Entrepreneurship* as 1, and 0 otherwise. Moreover, the database also provides other valuable information on entrepreneurial outcomes, such as entrepreneurial income, intentions, and survival entrepreneurship. For additional analyses below, we considered these variables to deepen our comprehension of the influence of childhood left-behind experiences on entrepreneurial outcomes in the Chinese context.

3.2.2. The left-behind experience

The explanatory variable is *Left-behind experience*, which is measured by whether the respondent had left-behind experiences in childhood. In the CLDS questionnaire, respondents born after 1980 were asked who they primarily lived with at each school age.[3] Options included both parents, father, mother, grandparents, other relatives and friends, siblings, living alone, and others. According to previous studies, a child is considered to be left behind if at least one parent has migrated (e.g., Chang et al., 2019; Liu et al., 2022; Zhao & Chen, 2022). Thus, in at least one of the school years, respondents with answers other than "both parents" were defined as having left-behind experiences in adulthood (Liu et al., 2022), and the variable *Left-behind experience* was set to 1, and 0 otherwise. This information can help us explore the heterogeneous effect of left-behind experiences at different school ages on entrepreneurial activities in adulthood. In addition, in the robustness tests, we further considered the case where both parents migrate.

3.2.3. Control variables

Owing to the selective nature of left-behind experiences in childhood, controlling for specific individual and family attributes can alleviate potential endogeneity concerns, especially when some pre-left-behind experience characterization variables are included (Yi et al., 2022; Zhao & Guo, 2022). Specifically, we controlled for gender, age, parents' *hukou* at birth, parents' education, and siblings.[4] Obviously, these variables may be related to individuals' decision to left-behind experiences in childhood. For example, parents' education can influence their employment status and family income, which may further affect an individual's later entrepreneurial decisions (Cheng & Smyth, 2021). Incorporating these controls is instrumental in ensuring our estimates discern the causal relationship between left-behind experiences and entrepreneurial activities. Table A1 in the Appendix contains the pertinent definitions for above variables.

3.3. Empirical design

As highlighted earlier, we used a quantitative study design to explore the impact of left-behind experiences in childhood and entrepreneurial activities in adulthood. The estimation method we used the conditional mixed process (CMP) approach with instrumental variable, which is in consideration of potential endogeneity issues. More specifically, we chose proportion of left-behind children at the birth cohort level as an instrumental variable. In China, there are differences in the probability of children being left behind across birth cohorts (MacLachlan & Gong, 2022). This instrumental variable helps to capture exogenous shocks to macro-level parental migration events (Wang et al., 2021; Zhao et al., 2022). Moreover, the left-behind proportion reflects the overall state of parental migration across time in China, which affected parental migration decisions of individual families owing to peer group effects. Furthermore, parental migration rates (left-behind children ratio) are unlikely to be directly related to an individual's entrepreneurial decisions in the early stages of life (Wang et al., 2021). To further safeguard the exogeneity of this instrumental variable, we used another micro-database, China Family Panel Studies, to compute the left-behind proportion (Zhao & Guo, 2022). Taken together, we had considerable confidence that this instrumental variable effectively satisfies the correlation assumption and exclusion restriction assumption.

One potential threat is that information at the birth cohort level may influence an individual's entrepreneurial decisions (Cheng et al., 2021; Marquis & Qiao, 2020). Therefore, we strictly controlled for the 2-year birth cohort fixed effect (FE)[5] and used birth cohort-level clustering standard errors for correcting. Consequently, we formulated the following endogenous bivariate probit model that integrates probits for both entrepreneurship and the left-behind experience. Similar to previous literature (Hou et al., 2020; Zhao et al., 2024), we used the CMP method to estimate this complex model.[6]

$$Entrepreneurship_{ic}^{*} = \alpha_0 + \beta LBE_{ic} + \alpha_2 X_{ic} + \phi(c) + \lambda_p + \mu_{ic} \tag{1}$$

$$Entrepreneurship_{ic} = \begin{cases} 1 if Entrepreneurship_{ic}^{*} > 0; \\ 0 otherwise \end{cases} \tag{2}$$

$$LBE_{ic}^{*} = \omega_0 + \omega_1 Z_{ic} + \omega_2 X_{ic} + \phi(c) + \lambda_p + \varepsilon_{ic} \tag{3}$$

[2] As mentioned above, with the reform and opening up, China has seen a wave of migrant workers since the 1980s and the problem of left-behind children gradually came to the fore in rural China (Wang et al., 2019; Zhao & Chen, 2022). Therefore, it is quite reasonable for the CLDS database to ask only respondents born after 1980 (Liu et al., 2022).

[3] These school age stages include preschool (0–6 years old), primary school (7–12 years old), junior high school (13–15 years old), and senior high school (16–18 years old).

[4] Since the explained variable is a dummy variable, we mainly used probit and endogenous bivariate probit models in the empirical analysis below, reporting the marginal effects. Considering that the marginal effects are meaningful only when all predictors have scores between 0 and 1. Therefore, continuous variables such as age, parents' education, and siblings were normalized. Similarly, other related continuous variables assigned a value greater than 1, were also normalized to satisfy the conditions for reporting marginal effects.

[5] The 2-year birth cohort level is about as tight a form as we can control for. Otherwise, the full covariance problem arises if birth cohort FE is controlled in the model.

[6] In the following empirical analysis, both the endogenous treatment effect model and endogenous bivariate probit model can be estimated using the CMP method.

$$LBE_{it} = \begin{cases} 1 if LBE_{it}^{*} > 0; \\ 0 otherwise \end{cases} \quad (4)$$

where $Entrepreneurship_{ic}^{*}$ is a latent variable representing whether adult i started a business in birth cohort c. Whether respondent i born in year c had left-behind experiences is indicated by the dummy variable LBE_{ic}. A group of control variables is referred to as X_{ic}. As mentioned above, the inclusion of some exogenous control variables helps potential endogeneity. $\phi(c)$ and λ_p denote 2-year birth cohort FE and province FE, respectively.LBE_{ic}^{*} is a latent variable that depicts the continuously changing inclination of person i to be left-behind in childhood. The left-behind percentage is the instrumental variable, expressed by Z_{ic} in equation (3). μ_{ic} and ε_{ic} are random error terms.

Furthermore, in the mechanism analysis, to further test the hypotheses presented above, we constructed the following empirical models:

$$Mech_{ic} = \varphi_0 + \varphi_1 LBE_{ic} + \varphi_2 X_{ic} + \phi(c) + \lambda_p + \xi_{ic} \quad (5)$$

$$Entrepreneurship_{ic} = \theta_0 + \theta_1 LBE_{ic} + \theta_2 Mech_{ic} + \theta_3 X_{ic} + \phi(c) + \lambda_p + \zeta_{ic} \quad (6)$$

In equation (5), $Mech_{ic}$ is the set of mediating variables. If the estimated coefficient of LBE_{ic} φ_1 is in line with theoretical expectations and the absolute value of θ_1 in Eq. (6) is less than the absolute value of baseline estimated coefficient β, we could assume that the mediating effect holds. That is, the mechanism is effective in this study. Definitely, since endogenous explanatory variable LBE_{ic} is also included in both equations, the CMP approach with an instrumental variable was also be adopted. Incidentally, ξ_{ic} and ζ_{ic} are similar to the previous equations, denoting the random error terms.

4. Results

4.1. Summary statistics

Before the regression analysis, we began with a brief descriptive analysis in Table 1. From our final sample of 2918 observations, 326 respondents identified as entrepreneurs, representing 11.17 % of the full sample. In addition, we found that 35.06 % of the respondents had left-behind experiences, which is close to the previous statistics in the literature (Liu et al., 2022; Wang et al., 2021). The percentage of those whose parents both migrated in childhood is 30.67 %. These statistics are consistent with our intuition that rural parents are more likely to migrate as their children get older in China. Moreover, the results of the T-tests show that compared to adults without left-behind experiences, adults with left-behind experiences are 5.01 percentage points less likely to start a business. These statistics tentatively suggest that the left-behind experience in childhood is likely to be detrimental to entrepreneurial activity in adulthood.

In terms of the instrumental variable, we found that compared to individuals without left-behind experiences, the mean value of the birth cohort-level proportion of left-behind is 1.77 percentage points higher for these with left-behind experiences in Table 1, which is in line with our expectations and implies that the correlation hypothesis for this instrumental variable is statistically satisfied. Moreover, the T-test results show significant mean differences for most of the control variables, suggesting the appropriateness of our selected controls. Furthermore, Table 2 represents the correlation coefficients of variables. Notably, the coefficient between *Left-behind experience* and *Entrepreneurship* is both negative and statistically significant, offering preliminary evidence that supports our conjecture.

4.2. Baseline results

We explored the impact of childhood left-behind experiences on entrepreneurial activities in adulthood and Table 3 shows our main results. In column (1), the results by using the linear probability model pinpoint a negative correlation between left-behind experiences and entrepreneurship (*marginal effect* = -0.0557, $p < 0.05$). Also, the results remain unchanged in column (2) by using the probit method (*marginal effect* = -0.0564, $p < 0.05$). While these results are consistent with our theoretical expectations, we should proceed with caution given the potential endogeneity issues discussed previously. The concluding two columns of Table 3 lay out the outcomes from the endogenous bivariate probit model. Column (3) unveils the first-stage results, highlighting a positive correlation between our chosen instrumental variable, *Left-behind percentage*, and *Left-behind experience* (*marginal effect* = 0.9793, $p < 0.01$). This inference bolsters the correlation assumptions behind instrumental variable estimation and dispels concerns of a weak instrumental variable (Zhao & Guo, 2022; Zhao et al., 2024). The statistical significance of the correlation coefficient ($\rho_{\varepsilon\mu} = 0.3855, p < 0.01$) further emphasizes the existence of a selectivity bias from unobserved variables influencing childhood left-behind experiences (Han & Vytlacil, 2017), thereby reinforcing the validity of deploying the endogenous bivariate probit model.

Column (4) of Table 3 reports the second-stage findings. We found that the coefficient on *Left-behind experience* remains significantly negative (*marginal effect* = -0.1755, $p < 0.01$), hinting that childhood left-behind experiences reduce the likelihood of entrepreneurial pursuits in adulthood by a substantial 17.55 percentage points. When converted into the standard deviation metrics, this estimate parallels a 0.2658 standard deviation decrease in entrepreneurship. Compared to previous literature on the impact of childhood experiences on entrepreneurship in adulthood in China (e.g., Cheng et al., 2021; Qin & Kong, 2021; Yi et al., 2022; Zhao & Li, 2022), our estimates are economically significant. For instance, Zhao & Li (2022) found that for Chinese adults, adverse childhood experiences resulted in a 0.087 standard deviation decrease in the likelihood of being engaged in entrepreneurship. Estimates from Cheng et al., (2021) show that on average, the childhood experience of the famine increased the probability of a rural resident becoming an entrepreneur by 0.2–1.2 percentage points. In sum, our main findings underscore that left-behind experiences in childhood are detrimental to fostering entrepreneurial initiatives in adulthood in China.

4.3. Robustness checks

In this subsection, we reinforced the credibility of our baseline findings with various robustness checks. We employed alternative estimated methods and key variables, introduced supplementary control variables, and excluded potential confounding childhood experiences linked to left-behind experiences.

4.3.1. Using alternative specifications

In the main specifications, we acknowledged a potential limitation: the sample might not ideally represent the broader population, thereby making our estimates vulnerable to sample selection bias (Zhao et al., 2024). Adults with left-behind experiences are also more likely to suffer from mental illness and criminal tendencies (Li X et al., 2021; Liang et al., 2020; Sun et al., 2017), making it possible for the CLDS to adequately capture respondents with left-behind experiences. To counteract this, the first two columns of Appendix Table A2 apply the Heckman two-step method as a corrective measure (Cheung & Chow, 2011; Wolfolds & Siegel, 2019). In column (1), we used all exogenous control variables to derive the Inverse Mills Ratio during the preliminary phase. Meanwhile, column (2) introduces an augmentation by integrating the instrumental variable to generate a modified Inverse Mills Ratio (Wolfolds & Siegel, 2019; Zhao et al., 2024).[7] We found that although the coefficient on *Inverse Mills Ratio* is insignificant in column

[7] In columns (1) and (2), the probit estimation is used in the second stage of the Heckman two-step method.

Table 1
Descriptive statistics.

Variables	Full Sample			Respondents without left-behind experiences	Respondents with left-behind experiences	Difference (T-tests)
	N	Mean	S.D.	Mean	Mean	Mean
Panel A. Entrepreneurship activities						
Entrepreneurship	2918	0.1117	0.3151	0.1293	0.0792	0.0501***
Entrepreneurial income	2835	0.3749	2.3997	0.4000	0.3304	0.0697
Entrepreneurial intentions	2918	0.2584	0.4378	0.2327	0.3060	−0.0732***
Panel B. Left-behind experiences in childhood						
Left-behind experience	2918	0.3506	0.4772			
Both parents migrate	2918	0.3067	0.4612			
Left-behind at preschool	2259	0.1611	0.3677			
Left-behind at primary school	2302	0.1768	0.3816			
Left-behind at junior high school	2440	0.2234	0.4166			
Left-behind at senior high school	2787	0.3201	0.4666			
Panel C. IV						
Left-behind percentage	2918	0.3147	0.0558	0.3085	0.3262	−0.0177***
Panel D. Control variables						
Age	2918	0.5859	0.2821	0.6207	0.5214	0.0993***
Gender	2918	0.5490	0.4977	0.5451	0.5562	−0.0111
Father's *hukou* at birth	2918	0.0154	0.1232	0.0100	0.0254	−0.0154***
Mother's *hukou* at birth	2918	0.0065	0.0804	0.0063	0.0068	−0.0005
Education of father	2918	0.8727	0.1523	0.8687	0.8799	−0.0112**
Education of mother	2918	0.7918	0.1895	0.7910	0.7903	−0.0024
Siblings	2918	0.8706	0.1604	0.8666	0.8779	−0.0113*

Notes: ***, **, * denote the significance levels of 1%, 5%, and 10%, respectively.

Table 2
Correlation coefficient matrix.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Entrepreneurship (1)	1								
Left-behind experience (2)	−0.0759***	1							
Age (3)	0.0046	−0.1679***	1						
Gender (4)	−0.0393**	0.0106	0.0166	1					
Father's *hukou* at birth (5)	−0.0179	0.0596***	−0.0233	−0.0319*	1				
Mother's *hukou* at birth (6)	−0.0152	0.0030	−0.0080	−0.0379**	0.4394***	1			
Education of father (7)	0.0594***	0.0352*	−0.1204***	−0.0086	0.0452**	0.0207	1		
Education of mother (8)	0.0404**	0.0059	−0.1617***	0.0029	0.0421**	0.0516***	0.4955***	1	
Siblings (9)	−0.0417**	0.0336*	0.1588***	0.1853***	−0.0393**	−0.0295	−0.0523***	−0.1588***	1

Notes: ***, **, * denote the significance levels of 1%, 5%, and 10%, respectively.

(1), it is significantly positive in column (2). This hints at potential sample selection bias. Fortunately, this estimate (*marginal effect* = -0.1415, $p < 0.01$) is generally consistent with the baseline result. Thus, the foundational assertion that childhood left-behind experiences deter entrepreneurial pursuits in adulthood remains unchallenged even with this alternative analytical approach.

Second, another concern arises from differences in the characteristics and sample sizes of adults with and without childhood left-behind, which are likely to affect our estimates (Li et al., 2015; Yang et al., 2023). Although we used the instrumental variable to mitigate endogeneity, we further considered the propensity score matching approach in robustness tests (Wang et al., 2024; Zhao et al., 2024). Specifically, we used nearest neighbor matching to obtain a subsample, and then we applied the endogenous bivariate probit model for estimations. Column (3) of Table A2 reports the results of the second stage. We found that the coefficient on *Left-behind experience* remains significantly negative (*marginal effect* = -0.1043, $p < 0.01$), indicating that our baseline results are quite robust by adopting this alternative estimation method.

Third, in the main results, we used the endogenous bivariate probit model to treat the explained variable as a dummy variable. In fact, some of the literature also adopted a linear approach for estimation (e.g., Wang et al., 2021; Zhao & Chen, 2022; Zhao & Guo, 2022). Therefore, in the robustness test, we substituted our estimation approach with the endogenous treatment effect model with the instrumental variable (Hou et al., 2020; Zhao et al., 2024). Column (4) of Table A2 revealed that the endogenous treatment effect model estimate for left-behind experience consistently manifests as significantly negative (*marginal effect* = -0.1248, $p < 0.01$), underlining that the main results in Table 3 exhibit stability irrespective of the chosen estimation technique.

Fourth, we attempted to further demonstrate the robustness of the instrumental variable, confirming that our estimation results are not sensitive to how the instrumental variable is selected and measured (Cheung, 2024; Zhao & Guo, 2022). We chose an alternative instrumental variable, which is as measured by the overall left-behind rate at the village level. Prior literature discussed the rationale for this instrumental variable. They recognized that village-level migration rates, as a proxy for social networks, can influence parental migration decisions and the formation of left-behind children by providing information on employment and reducing the costs and risks of migration (Wang et al., 2021). The results are represented in column (1) of Table A3 in the Appendix.These estimates indicate that even with the use of this alternative instrumental variable, the conclusion that left-behind experiences in childhood are detrimental to entrepreneurship in adulthood remains fairly robust (*marginal effect* = -0.1373, $p < 0.01$).

Last, we further tested the sensitivity of the core explained and explanatory variables to ensure the reliability of our main results. Consistent with previous literature (e.g., Liu & Villa, 2020; Yi et al., 2022; Zhao & Li, 2022), we used alternative variables for robustness testing. In column (2) of Table A3, we replaced the explanatory variable with the number of entrepreneurial endeavors because some respondents had more than just one entrepreneurial experience in adulthood. Clearly, the estimation results are not sensitive to how the explained variable is measured (*marginal effect* = -0.0517, $p < 0.05$). Furthermore, in the main specification, one parent migrating was

Table 3
The effect of left-behind experiences on entrepreneurial behaviors.

	(1)	(2)	(3)	(4)
	Linear probability model	Probit	Endogenous bivariate probit model	
	Entrepreneurship	Entrepreneurship	Left-behind experience (probit)	Entrepreneurship (probit)
Left-behind experience	−0.0557**	−0.0564**		−0.1755***
	(0.0236)	(0.0231)		(0.0340)
Left-behind percentage			0.9793***	
			(0.1710)	
Age	0.1363	0.1585	−0.2358	0.1321
	(0.1832)	(0.1870)	(0.1923)	(0.1884)
Gender	−0.0187	−0.0187*	0.0113	−0.0177
	(0.0118)	(0.0108)	(0.0151)	(0.0110)
Father's *hukou* at birth	−0.0401	−0.0511	0.2481***	−0.0229
	(0.0517)	(0.0680)	(0.0769)	(0.0666)
Mother's *hukou* at birth	−0.0552	−0.0684	−0.1529	−0.0882
	(0.0745)	(0.1123)	(0.1035)	(0.1211)
Education of father	0.1145**	0.1307***	0.0870	0.1467***
	(0.0422)	(0.0468)	(0.0687)	(0.0489)
Education of mother	0.0137	0.0149	−0.0393	0.0108
	(0.0431)	(0.0444)	(0.0557)	(0.0473)
Siblings	−0.0787*	−0.0742**	0.0302	−0.0734*
	(0.0391)	(0.0354)	(0.0735)	(0.0376)
$\rho_{\epsilon\mu}$			0.3855***	
			(0.0848)	
Province FE	Yes	Yes	Yes	Yes
2-year birth cohort FE	Yes	Yes	Yes	Yes
Observations	2918	2918	2918	2918

Notes: The last three columns report the marginal effects for probit regressions. Columns (3) and (4) report the results using the endogenous bivariate probit model. Among them, column (3) shows the first-stage results and the second-stage results are presented in column (4). ***, **, * denote the significance levels of 1%, 5%, and 10%, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

considered to be left behind. We now strictly define *Left-behind experience* as both parents migrating for at least one school-age stage. In column (3) of Table A3, we found that the estimate keeps robust (*marginal effect* = -0.2283, $p < 0.01$). Additionally, according to Table 1, a higher rate of parental migration at the high school level may take into account the fact that high school students are more independent and have less need for parental care and supervision (Chang et al., 2019). Therefore, we redefined the variable *Left-behind experience* by considering only the preschool, primary school, and junior high school. The results in column (4) of Appendix Table A3 show that this estimate is qualitatively similar to the main estimates, although somewhat change in magnitude (*marginal effect* = -0.2810, $p < 0.01$).

4.3.2. Considering addition controls

As emphasized above, the main threat comes from omitting unobserved variables in our specifications. Thus, here we perform robustness tests by adding some additional control variables, which is in line with prior research (Chen & Zhao, 2022; Wang et al., 2024; Yang et al., 2023). On one hand, we additionally controlled for some village characteristics, including *Migrant population*, *Topography*, *Surnames*, *Clan*, *Ethnic minority*, *Distance from the county seat*, *Village infrastructure*. These village level variables are defined as shown in Appendix Table A5. In column (1) of Appendix Table A4, we found that the coefficient on *Left-behind experience* keeps significantly negative (*marginal effect* = -0.1380, $p < 0.01$). Moreover, we further strictly controlled for village FE in column (2) of Table A4. The estimates show a similar negative effect of left-behind experiences in childhood on entrepreneurial activities in adulthood (*marginal effect* = -0.4642, $p < 0.01$).

On the other hand, to further ensure the exogeneity of the instrumental variable, we controlled for two high-dimensional FEs related to birth cohorts, namely province FE*birth cohort trend and porovince-by-2-year birth cohort FE, as the instrumental variable in main results also involve information at the birth cohort level. These high-dimensional FEs further capture provincial characteristics that vary with birth cohort, reduces endogeneity, ensuring the exclusion restriction hypothesis (Hou et al., 2020; Zhao & Guo, 2022). In columns (3) and (4) of Table A4, we could see that the results remain unchanged (*marginal effect* = -0.1472, $p < 0.01$; *marginal effect* = -0.1791, $p < 0.01$), which means our main specification is satisfactory in capturing these relevant FEs.

4.3.3. Excluding the effects of other childhood experiences

Previous studies found that some similar childhood-specific experiences may also affect entrepreneurial activities in adulthood, which could be a shock to our baseline results (Cheng et al., 2021; Churchill et al., 2021; Yi et al., 2022; Zhao & Li, 2022). We therefore excluded the effect of these experiences in our robustness tests. At the micro level, we considered three experiences: childhood parental divorce, childhood poverty, and childhood migration. The CLDS records some information about respondents at age 14, such as parental marital status, family socioeconomic status (self-rated), and migration status.[8] At the macro level, we mainly considered shocks from natural disasters. Since our sample was born after 1980, the major natural disasters they may have experienced include the Wenchuan earthquake in 2008 and the China floods in 1998.[9] In the first four columns of Table A6 in the Appendix, we found that the coefficient on *Left-behind experience* remains unchanged after excluding the effects of these particular childhood experiences. Further, in column (5), we considered all these childhood experiences simultaneously. As expected, the estimate for left-behind experience is

[8] Specifically, first, observations of parental divorce at age 14 are excluded from column (1) of Table A6. Second, families were considered poor when the self-assessed socioeconomic status of the family at age 14 was a score of 1 (out of a total score of 10). These observations are dropped in column (2) of Table A6. Third, observers are removed from column (3) of Table A6 if they migrated before 14.

[9] Specifically, respondents were considered to have experienced the Wenchuan earthquake if their place of birth was Sichuan and they were born after 1990 (they were younger than 18 in 2008). In addition, if respondents were born in Hunan, Jiangxi, Hubei, and Heilongjiang, which were most affected by the floods, and were born before 1998, they were considered to have experienced the China floods in 1998 (Zhao et al., 2024).

virtually unchanged (*marginal effect* = -0.2269, $p < 0.01$).

4.4. Mechanism tests

In this subsection, we tested multiple potential channels through which left-behind experiences in childhood decreases entrepreneurial activities in adulthood. As noted in above hypotheses, these channels include human capital, social networks, and risk appetite.

4.4.1. Human capital

Hypothesis 1 posits that left-behind experiences can influence two primary dimensions of human capital: education and skills, as well as physical and mental health. To test this mechanism, drawing on the CLDS questionnaire and insights from earlier studies (e.g., Meng & Yamauchi, 2017; Zhao & Li, 2022), we firstly identified five mediating variables to gauge respondents' educational and skill levels, *Education*, *College*, *Language skill*, *Vocational certificate*, and *Entrepreneurship training*. In addition, to assess respondents' physical and mental well-being, we devised four mediating variables. Physical health is captured through *Self-rated health* and *Injury*, while variables *Emotional issues* and *Mental issues* reflect mental health. Comprehensive definitions and statistics for these human capital measures can be found in Panels A and B of Appendix Table A7.

The results of the effect of left-behind experiences in childhood on education and skills in adulthood are reported in Panel A of Table 4. We could see that the left-behind experience in childhood significantly leads to a lower probability of attending college (*marginal effect* = -0.1160, $p < 0.05$). These results are similar to previous findings in the literature, emphasizing the negative impact of left-behind on education (Wang et al., 2019; Zhao et al., 2014). Moreover, in column (3), we found that the left-behind experience is not correlated with language skills in adulthood. In columns (4) and (5), the coefficients of *Left-behind experience* are both negative in estimates by using endogenous bivariate probit model (*marginal effect* = -0.1282, $p < 0.05$; *marginal effect* = -0.1145, $p < 0.1$), indicating that left-behind experiences lead to adults being less likely to have vocational certificates and entrepreneurial training. Clearly, owing to the lack of parental supervision and tutoring, Chinese left-behind children may perform worse academically, which is not conducive to their long-term human capital accumulation (Lyu & Chen, 2019; Wang et al., 2021).

We further explored the influence of left-behind experiences in childhood on physical and mental health and the results are shown in Panel A of Table 5. The results suggest that the left-behind experience is significantly associated with self-rated health and mental illness. More specifically, in columns (1) and (4), it is indicated that left-behind experiences resulted in a decrease in self-rated health scores (*marginal effect* = -0.0182, $p < 0.01$) and an increase in mental illness scores (*marginal effect* = 0.0893, $p < 0.01$). Our findings echo previous research, emphasizing the detrimental health impact of left-behind experiences in childhood, as evidenced by Li X et al. (2021), Wu et al. (2015), and Zhang et al. (2021). Moreover, we found that adults with left-behind experiences also had a higher probability of experiencing work-related injuries (*marginal effect* = 0.0507, $p < 0.05$). Prior literature confirmed that labors with left-behind experiences are more likely to work in lower-end and riskier jobs in China, which may adversely affect their somatic health (Yang et al., 2023; Zhou et al., 2015). All in all, these results supported the view that left-behind experiences in childhood is indeed detrimental to physical and mental health in adulthood.

To obtain rigorous procedures of mechanism testing, according to the Eqs. (5) and (6), we further added each of these mediating variables to the baseline results, respectively, and observed the changes in estimated coefficients on the core explanatory variable *Left-behind experience*. In Panel B of Table 4, compared to the baseline estimate in column (4) of Table 3 (*marginal effect* = -0.1755, $p < 0.01$), we found that the absolute value of the coefficient on *Left-behind experience* decreases in the last four columns. Especially in columns (9) and (10), it fells by almost 30 % (*marginal effect* = -0.1235, $p < 0.01$; *marginal effect* = -0.1224, $p < 0.01$). In addition, the coefficients of these mediating variables are significantly positive except for column (8), which is similar to the findings in previous literature, emphasizing the positive role of education and skill training in promoting entrepreneurship (e.g., Ahn & Winters, 2023; Baron et al., 2016; Qin & Kong, 2021). Moreover, in Panel B of Table 5, we found that left-behind experiences are detrimental to adulthood entrepreneurship by decreasing self-rated health and developing mental illnesses. As the previous literature implies, good physical and mental health is one of the important foundations of entrepreneurship (Zhao & Li, 2022). However, although there are

Table 4
Human capital: education and skills (CMP methods).

Panel A. The effect of left-behind experience on education and skills					
	(1) Education	(2) College	(3) Language skill	(4) Vocational certificate	(5) Entrepreneurial training
Left-behind experience	−0.0113	−0.1160**	0.0037	−0.1282**	−0.1145*
	(0.0154)	(0.0467)	(0.0437)	(0.0591)	(0.0591)
Observations	2918	2918	2918	2918	2918
Panel B. Mediation effects of education and skills					
	(6) Entrepreneurship	(7) Entrepreneurship	(8) Entrepreneurship	(9) Entrepreneurship	(10) Entrepreneurship
Left-behind experience	−0.1732***	−0.1757***	−0.1722***	−0.1235***	−0.1224***
	(0.0353)	(0.0380)	(0.0333)	(0.0311)	(0.0311)
Education	0.4364***				
	(0.1026)				
College		0.0463***			
		(0.0153)			
Language skill			−0.0188		
			(0.0143)		
Vocational certificate				0.0312*	
				(0.0160)	
Entrepreneurial training					0.0333**
					(0.0167)
Observations	2918	2918	2918	2918	2918

Notes: In column (1), the endogenous treatment effect model is used in Panel A, as *Education* is a continuous variable (normalization). In the last four columns, the endogenous bivariate probit model is employed in Panel A, with all columns reporting marginal effects for the second-stage results, as these four variables are binary. The endogenous bivariate probit model is employed in all columns of Panel B, reporting marginal effects for the second-stage results. The baseline control variables, province FE, and 2-year birth cohort FE are added in all regressions. ***, **, * denote the significance levels of 1%, 5%, and 10%, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

Table 5
Human capital: physical and mental health (CMP methods).

Panel A. The effect of left-behind experience on physical and mental health				
	(1)	(2)	(3)	(4)
	Self-rated health	Injury	Emotional issues	Mental illness
Left-behind experience	−0.0182***	0.0507**	0.0066	0.0893***
	(0.0028)	(0.0218)	(0.0631)	(0.0043)
Observations	2918	2918	2918	2918
Panel B. Mediation effects of physical and mental health				
	(5)	(6)	(7)	(8)
	Entrepreneurship	Entrepreneurship	Entrepreneurship	Entrepreneurship
Left-behind experience	−0.1760***	−0.1852***	−0.1759***	−0.1675***
	(0.0393)	(0.0346)	(0.0338)	(0.0387)
Self-reported health	0.5627			
	(0.6902)			
Injury		0.0320**		
		(0.0153)		
Emotional issues			0.0317	
			(0.0249)	
Mental illness				−0.0604*
				(0.0310)
Observations	2918	2918	2918	2918

Notes: In columns (1) and (4), the endogenous treatment effect model is used in Panel A, as *Self-reported health* and *Mental illness* are continuous variables (normalization). In columns (2) and (3), the endogenous bivariate probit model is employed in Panel A, with all columns reporting marginal effects for the second-stage results, as these two variables are binary. The endogenous bivariate probit model is employed in all columns of Panel B, reporting marginal effects for the second-stage results. The baseline control variables, province FE, and 2-year birth cohort FE are added in all regressions. ***, **, * denote the significance levels of 1%, 5%, and 10%, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

somewhat changes in the coefficient of *Left-behind experience*, especially for column (8) (*marginal effect* = -0.1675, $p < 0.01$), these changes are modest compared to the baseline results. This suggests that the mediating effect of health may not have been evident in our study.

In summary, our results provide supportive evidence for Hypothesis 1 that the left-behind experiences in childhood is likely to decrease the possibility of entrepreneurial behaviors owing to the destruction of human capital accumulation, particularly reflected in vocational certificate and entrepreneurship training.

4.4.2. Social networks

Next, we considered the mediation effects of social networks, as laid out in Hypothesis 2. To test this hypothesis, we selected five related mediating variables to measure the social networks, namely *Entrepreneurial opportunities from friends*, *Start-up capital comes from relatives and friends*, *Local friends with lending money*, *Social activities*, and *Gift money*. Among them, the first two are closely related to entrepreneurial activities. However, since the CLDS only asks about entrepreneurs, it cannot be regressed on mediating effects (otherwise it faces the full covariance problem). The last three are social network variables in a general form, as highlighted in previous literature (Burt, 2019; Gee et al., 2017). Detailed information and descriptions pertaining to these social capital-related variables can be found in Panel C of Table A7 in the Appendix.

Table 6 presents the findings examining the social capital channel via the CMP method. In Panel A, the coefficients on *Left-behind experience* in the last four columns are notably negative and statistically significant.

Table 6
Social networks (CMP methods).

Panel A. The effect of left-behind experience on social networks					
	(1) Entrepreneurial opportunities from friends	(2) Start-up capital comes from relatives and friends	(3) Local friends with lending money	(4) Social activities	(5) Gift money
Left-behind experience	−0.0880	−0.1097***	−0.0697***	−0.7747***	−0.3249***
	(0.1989)	(0.0298)	(0.0155)	(0.0809)	(0.0169)
Observations	2918	2918	2918	2918	2837
Panel B. Mediation effects of social networks					
		(6) Entrepreneurship	(7) Entrepreneurship		(8) Entrepreneurship
Left-behind experience		−0.1674***	−0.1756***		−0.1731***
		(0.0357)	(0.0340)		(0.0353)
Local friends with lending money		0.0420*			
		(0.0229)			
Social activities			0.4770***		
			(0.1649)		
Gift money					0.0015
					(0.0311)
Observations		2918	2918		2837

Notes: In columns (4) and (5), the endogenous treatment effect model is used in Panel A, as *Social activities* and *Gift money* are continuous variables (normalization). In columns (1)-(3), the endogenous bivariate probit model is employed in Panel A, with all columns reporting marginal effects for the second-stage results, as these two variables are binary. The endogenous bivariate probit model is employed in all columns of Panel B, reporting marginal effects for the second-stage results. The baseline control variables, province FE, and 2-year birth cohort FE are added in all regressions. ***, **, * denote the significance levels of 1%, 5%, and 10%, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

This aligns with the prevailing notion that adults with left-behind experiences might face challenges in their social networks, as supported by studies such as Li X et al. (2021). We further tested mediating effects of social networks in Panel B of Table 6. Clearly, in columns (6) and (7), coefficients on these social network variables are significantly positive (*marginal effect* = 0.0420, $p < 0.1$; *marginal effect* = 0.4770, $p < 0.01$), indicating that social network did play a positive role in entrepreneurial activities (e.g., Burt, 2019; Leyden et al., 2014; Zhao et al., 2024). In addition, we found that compared to the baseline estimate (*marginal effect* = -0.1755, $p < 0.01$), the absolute value of the coefficient of *Left-behind experience* (*marginal effect* = -0.1674, $p < 0.01$) also decreases in column (6), implying that the social network channel is feasible. Taken together, our estimation validates the reliability of Hypothesis 2 that by hindering social network development, left-behind experiences in childhood are detrimental to entrepreneurial behaviors in adulthood.

4.4.3. Risk appetite

The final channel is risk appetite. In Hypothesis 3, we posited that left-behind experiences in childhood might foster risk-averse characteristics, potentially inhibiting entrepreneurial pursuits in adulthood. However, one challenge, consistent with previous research, is the inadequate measurement of risk appetite (Benmelech & Frydman, 2015; Zhao & Li, 2022). Drawing from the CLDS database, we opted for two proxies related to risk appetite: *Non-essential debt* and *Risky financial assets*, as shown in Panel D of Appendix Table A7. It is clear that accruing more debt inherently indicates higher risk, applicable across individuals, households, or businesses (Huang et al., 2023; Zhao et al., 2024). Such non-essential debts can provide insights into an individual's risk tolerance. Conversely, investing in risky financial assets carries the potential for financial loss, suggesting a stronger inclination towards risk-taking (Wen et al., 2023).

Table 7 shows the tests examining the channel of risk appetite. In line with our anticipations, the coefficients on *Left-behind experience* in columns (1) and (3) are both notably negative (*marginal effect* = -0.0729, $p < 0.05$; *marginal effect* = -0.3489, $p < 0.05$). This suggests that adults with left-behind experiences tend to exhibit a reduced appetite for risk, a conclusion that resonates with findings from prior studies such as Bernile et al. (2017) and Cassar et al. (2017). Delving into the mediating role of risk appetite, as depicted in columns (2) and (5), there is a discernible negative association between non-essential debt and risky financial assets with entrepreneurship. This underscores the influence and positive role that risk appetite plays in entrepreneurial ventures (Putniņš & Sauka, 2020), implying the plausibility of these two variables in representing risk appetite. Additionally, the results suggest the presence of a mediating effect of risk appetite, this effect is prominently visible. Compared to column (4) (*marginal effect* = -0.1399, $p < 0.01$),[10] the coefficient on *Left-behind experience* (*marginal effect* = -0.0960, $p < 0.05$) decreases by about 31.38 % in column (5). However, the estimated coefficient on *Non-essential debt* is insignificant in column (2). Thus, our results are suggestive and partially verify Hypothesis 3.

4.5. Heterogeneity effects

In addition to baseline results and potential mechanisms, we are also interested in the heterogeneous effects of left-behind experiences affecting entrepreneurship in adulthood. In this subsection, we explored the heterogeneity in three ways: left-behind experiences at different stage, gender, and age.

The life course theory describes the impact of differences in experiences on outcomes, such as exposure frequency, chronology, and age stages (George, 2014). Thus, we first analyzed the effect of left-behind experiences at different stages on entrepreneurial activities in adulthood. In Table 8, we found that the estimated coefficients on *Left-behind experience* are significantly negative in all four columns, suggesting that left-behind experiences at any stages have adverse effects on entrepreneurial activities in adulthood. Specifically, the estimates show that left-behind experiences in preschool, primary school, junior high school, and senior high school lead to a decrease in the likelihood of entrepreneurship in adulthood by 0.4590, 0.4285, 0.4477, and 0.3077 standard deviations, respectively. These results imply that the negative effects of left-behind experiences at preschool and middle school stages on entrepreneurship in adulthood seem to be even more pronounce (Li M et al., 2021). Importantly, these estimates tell us that the negative impact of childhood left-behind experiences on later entrepreneurial activities is quite comprehensive.

We further focused on the heterogeneity effects by gender. Evidence from previous studies suggested that left-behind has inconsistent effects on boys and girls. Girls are more likely to be traumatized and lack self-confidence (Xiao et al., 2019; Wang et al., 2019; Zhao & Chen, 2022); boys may develop avoidant personalities and are more likely to be addicted to the Internet and violence (Wen et al., 2021; Zhang et al., 2019). Yet, the question remains: can these gender-driven disadvantages directly influence entrepreneurial activities in adulthood? To further investigate this, we segmented the complete sample into two subsamples based on the respondents' gender. The findings from this analysis can be observed in columns (1) and (2) of Table 9. The coefficient of *Left-behind experience* is significantly negative in both columns (*marginal effect* = -0.1736, $p < 0.01$; *marginal effect* = -0.2083, $p < 0.01$), indicating that there is no much gender difference in the negative impact of left-behind experiences in childhood on entrepreneurial behavior in adulthood. It is indicated that although the manifestation of the negative impact of left-behind on boys and girls is not identical, the difference in their negative performance on later entrepreneurship in adulthood is not significant from a long-term perspective.

Moreover, we were interested in the heterogeneity effect of left-behind experiences on adulthood entrepreneurial behaviors of different ages, as adults have different entrepreneurial trends at different ages (Yi et al., 2022; Zhao et al., 2021). We analyzed the heterogeneity effect using the median 30 years old a cut-off due to the fact that younger than 30 is usually recognized as the youth group (Ida et al., 2020; Saud et al., 2020). Based on the findings presented in columns (3) and (4) of Table 9, although the coefficients on *Left-behind experience* are significantly negative in the last column (*marginal effect* = -0.2663, $p < 0.01$), it appears that the negative effect of left-behind experiences in childhood on entrepreneurial behaviors is more pronounced for younger groups. These findings suggest that younger adults with left-behind experiences may gain more disadvantages from left-behind experiences, which in turn discourages their entrepreneurial motivations.

4.6. Further studies

In this subsection, we attempted to further enrich the understanding of how left-behind experiences affect entrepreneurial behaviors in adulthood, including the introduction of moderating effects analyses and exogenous shocks, observing entrepreneurial performance as well as explaining the income-penalizing effect of the left-behind experience.

4.6.1. Moderating effects on left-behind experiences

As emphasized above, the life course theory not only emphasizes the long-term effects of particular experiences during critical periods, but also endorses the milestone effect, which means the occurrence of important events or shocks that may alter the original effect (George, 2014; Zhao & Guo, 2022). That is, some later changes may have a moderating effect, or even a reverse moderating effect. In the Chinese setting, both the formation of left-behind childhood and entrepreneurial incentives are deeply embedded in this particular institutional context (Qin & Kong, 2021; Wang et al., 2021). Thus, we argued that

[10] Excessive missing values for the variable *Financial risk assets*. Column (4) of Table 7 shows the results of baseline results after excluding these missing values.

Table 7
Risk appetite (CMP methods).

	(1) Non-essential debt	(2) Entrepreneurship	(3) Risky financial assets	(4) Entrepreneurship	(5) Entrepreneurship
Left-behind experience	−0.0729**	−0.1761***	−0.3489**	−0.1399***	−0.0960**
	(0.0342)	(0.0371)	(0.1715)	(0.0504)	(0.0455)
Non-essential debt		0.0360			
		(0.0495)			
Risky financial assets					0.0640**
					(0.0285)
Baseline controls	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
2-year birth cohort FE	Yes	Yes	Yes	Yes	Yes
Observations	2847	2847	1028	1028	1028

Notes: The endogenous bivariate probit model is employed in all columns, reporting the marginal effects of second-stage results. ***, **, * denote the significance levels of 1%, 5%, and 10%, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

Table 8
Left-behind experiences at different stages (CMP methods).

	(1) Preschool	(2) Primary school	(3) Junior high school	(4) Senior high school
Left-behind experience	−0.3933***	−0.3539***	−0.3386***	−0.2078***
	(0.0874)	(0.0567)	(0.0516)	(0.0411)
Baseline controls	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
2-year birth cohort FE	Yes	Yes	Yes	Yes
Observations	2259	2302	2440	2787

Notes: The endogenous bivariate probit model is employed in all columns, reporting the marginal effects of second-stage results. ***, **, * denote the significance levels of 1%, 5%, and 10%, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

Table 9
Heterogeneity effects by gender and age (CMP methods).

	(1) Female	(2) Male	(3) Age ≥ 30	(4) Age ≤ 29
Left-behind experience	−0.1736***	−0.2083***	−0.1570	−0.2663***
	(0.0496)	(0.0688)	(0.0993)	(0.0574)
Baseline controls	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
2-year birth cohort FE	Yes	Yes	Yes	Yes
Observations	1602	1316	1611	1307

Notes: The endogenous bivariate probit model is employed in all columns, reporting the marginal effects of second-stage results. ***, **, * denote the significance levels of 1%, 5%, and 10%, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

environmental and institutional changes are likely to have a moderating effect on the unfavorable entrepreneurial effects of left-behind experiences in China. We chose three particular shocks with Chinese characteristics to test this milestone effect, *hukou* transfer, clans, and WTO accession (Wang et al., 2023; Zhang, 2020; Zhao et al., 2024).

First, we focused on *hukou* transitions, as our sample are all born with rural *hukou* status. Prior literature found potential benefits of *hukou* switching from rural to urban areas on entrepreneurial activities (Chen & Hu, 2021; Wang et al., 2023). Left-behind children may experience a range of *hukou* switching events in their later life course, which may have a moderating effect on their entrepreneurial activities in adulthood. We constructed a moderating variable *Hukou transfer*, which is equal to 1 if the respondent had an urban *hukou* in the survey year, and 0 otherwise. In columns (1) and (2) of Table A8 in the Appendix, moderating effects results show that the coefficients on the interaction term *Left-behind experience* × *Hukou transfer* is significantly positive by using the probit model (*marginal effect* = 0.0575, $p < 0.1$). These results suggest that *hukou* transitions may weaken the negative impact of left-behind experiences in childhood on entrepreneurial behaviors in adulthood. The *hukou* switching shock undermines the long-term effect of left-behind experiences.

Second, we further explored the moderating effects of clans. Clans are a particular cultural symbol in China that help to expand social networks, enhance resource channeling, and share risks (Zhang, 2020). As such, clans are also important drivers of entrepreneurial behaviors in China. We constructed an interaction term between *Left-behind experience* and *Clan* and the results are presented in columns (3) and (4) of Table A8. We found that although the estimated coefficients of this interaction term are positive, they are insignificant. These estimates hint at the potential of clan in weakening the adverse effects of left-behind experiences in childhood on entrepreneurial behaviors in adulthood. However, this moderating effect is not statistically significant in our sample.

Third, our study accounted for the exogenous shock resulting from China's WTO accession in 2001. Prior research validated that WTO accession fast-tracked the evolution of the Chinese market and catalyzed a surge in international demand (Moore et al., 2021; Qin & Kong, 2021). Soon, China underwent a notable entrepreneurial upswing (Moore et al., 2021; Zhao et al., 2024). Consequently, this pivotal event, which significantly shaped entrepreneurial undertakings, could have implications for our primary observed effects. We assigned the exogenous moderator variable *WTO accession* to 1 if in 2001, respondents had completed their formal education and were not less than 18 years old, implying that they were hit by the WTO accession event, and 0 otherwise. By constructing an interaction term, we could see positive moderating effects of WTO accession shocks on the entrepreneurial effects of left-behind experiences in the last two columns of Table A8 (*marginal effect* = 0.1014, $p < 0.05$; *marginal effect* = 0.1037, $p < 0.01$). Clearly, more entrepreneurial opportunities brought about by WTO accession effectively attenuate the negative effect of left-behind experiences on entrepreneurship in adulthood, further suggesting that changes in the macro-environment may lead to the disappearance of the effect of the left-behind experiences.

4.6.2. Entrepreneurial performances

In addition to entrepreneurial behaviors, we further focused on entrepreneurial performance to strengthen the comprehensive understanding on entrepreneurial impacts of the childhood left-behind experiences (Kwon & Ruef, 2017; Zhao et al., 2024). The CLDS database provides information related to entrepreneurial performances, such as entrepreneurial income, future entrepreneurial intentions, and survival entrepreneurship (Cheng & Smyth, 2021; Nguyen et al., 2019). Among them, survival entrepreneurship is measured by a binary variable accordingly to determine the forced entrepreneurial behaviors of

respondents. We were concerned that owing to the adverse circumstances faced by adults with left-behind experiences, it likely that they have survival entrepreneurial incentives that challenges on our main results (Yu et al., 2023). In Table A9, the estimates show that all coefficients on *Left-behind experience* are insignificant, suggesting that left-behind experiences in childhood are not only detrimental to entrepreneurial performance and future entrepreneurial intentions, but also do not create survival entrepreneurial motives. As we analyzed above, owing to disadvantages in human capital, social networks, and risk appetite, the adults with left-behind experiences lack entrepreneurial support, awareness, and passion.

4.6.3. *Entrepreneurial behaviors and the long-term income penalty effect of left-behind experiences*

As discussed earlier, prior research found that there are long-term income-penalizing effects of left-behind experiences in childhood (Liu et al., 2022; Lyu & Chen, 2019; Wang et al., 2021). That is, adults with left-behind experiences have lower average incomes than these without in China. Yet these studies did not address the contribution of entrepreneurial behaviors to this income penalty effects. Thus, here we attempt to analyze whether entrepreneurial behaviors could explain this effect.

Based on CLDS data, we selected several income-related variables, including *Individual income*, *Total household income*, *Household operating income*, and *Low income*. The definitions of these variables appear in Appendix Table A11. In Panel A of Table A10, the estimates tell us that left-behind experiences significantly reduced individual income and household operating income, which is consistent with the findings on the income-penalizing effects of left-behind experiences in prior research (Lyu & Chen, 2019; Wang et al., 2021). Moreover, in Panel B of Appendix Table A10, we further included *Entrepreneurship* as a mediating variable, finding that the coefficients are positive in the first three columns and significantly negative in column (4), which indicates the positive role of entrepreneurship in improving income (Prado et al., 2022; Zhao et al., 2014). More importantly, comparing the corresponding columns in Panel A, the absolute value decreases in all columns, suggesting that the mediating effect of entrepreneurial activities may hold. Clearly, these results mean that entrepreneurial behaviors may explain the long-run income penalizing effect of left-behind experiences in China.

5. Discussion

5.1. *Theoretical implications*

We sought to make three major contributions to understanding the long-term effect of left-behind experiences on entrepreneurial activities in adulthood. First, from the perspective of entrepreneurial activities, we further enriched the life course theory. The life course theory not only emphasizes the influence of particular experiences during sensitive periods on behaviors in later life course (Elder, 1998; Zhao & Guo, 2022), but also suggests that this effect may be weakened or even disappear as circumstances change owing to the milestone effect (George, 2014). Based on the sensitive period of childhood, we found a negative impact of left-behind experiences on entrepreneurship in adulthood. More importantly, considering the particular Chinese context of left-behind experiences, we attempted to test which environmental changes may weaken the negative effect of left-behind experiences in multiple ways. The results show that the particular shocks of *hukou* transfer and WTO accession make the negative impact of left-behind experiences in childhood on entrepreneurial activities in adulthood less pronounced. Thus, we offered sufficient and plausible evidence for the life course theory from multiple aspects.

Second, our study added the research field related to adverse childhood experiences and adulthood entrepreneurial activities. There are two divergent views. Some scholars found that childhood adverse experiences may make individuals more resilient and self-reliant, which helps them to undertake entrepreneurial activities in adulthood (Cheng et al. 2021; Churchill et al. 2021; Yi et al. 2022). Another set of studies emphasized that adverse experiences in childhood undermine human and social capital and discourage later entrepreneurship (Prakash & Kumar, 2021; Zhao & Li, 2022). Our results favor the latter. One potential reason for this is that most of the first type of literature is based on specific negative events such as wars, famines, and natural disasters (Cheng et al., 2021; Churchill et al., 2021; Yi et al., 2022). After these events, a large number of entrepreneurial opportunities emerged, which may further reinforce the positive impact of adverse childhood experiences on entrepreneurial activity (Churchill et al., 2021; Yu et al., 2023). We revealed that left-behind experiences not only weaken the entrepreneurship-related resource accumulations, but also develop risk averse traits. In conclusion, based on the left-behind experiences, our study provides new insights into how special childhood experiences impact entrepreneurial outcomes.

Third, this study further complemented related research on the welfare of left-behind children. The issue of left-behind children has long received attention from social science fields. We explored the long-term negative impact of left-behind experiences on entrepreneurial behaviors in adulthood, further finding that lack of entrepreneurship can explain the income-penalizing effect of left-behind experiences. This is a valid interpretation of the welfare of left-behind children from a new perspective. While there is a considerable body of research emphasizing the negative short-term effects of parental migration on left-behind children, some studies addressed the potential benefits in terms of remittance effects (Aregbeshola, 2022; Cuadros-Meñaca, 2020), as migrate parents may receive higher incomes. Thus, in the debate about left-behind children development, our study supports the view that left-behind experiences lead to long-term losses in individual welfare from an entrepreneurial perspective (Deng et al., 2020). Anyway, the long-term effects of left-behind experiences continue to deserve more academic focus.

5.2. *Practical implications*

Our study has important policy implications. It is essential to reduce the incidence of left-behind children at the source. One reason for the problem of rural left-behind children is the large number of migrant children who are unable to attend school in the cities. Therefore, by further easing institutional barriers to the enrollment of migrant children in urban schools, the prevalence of left-behind children can be diminished, as highlighted in existing studies (Wang et al., 2021; Zhao & Chen, 2022). Moreover, another reason for children left behind is the inability to provide suitable employment locally. So, better public policies should be formulated to encourage more migrant parents to return to their hometowns for employment and entrepreneurship.

For adults with left-behind experiences, more supports from family, community, and government are particularly important, especially as they related to entrepreneurial activities. First, proactive efforts by parents to enhance family bonds and restructure family dynamics could potentially counteract the enduring consequences of the stigma associated with left-behind experiences. Second, there is a need to provide more psychosocial centers and reduce societal discrimination against left-behind children. Third, The government sector could offer more entrepreneurial training programs to compensate for the vulnerability of adults with left-behind experiences in terms of human capital and social networks. Also, strengthening the reform of the *hukou* system and creating more entrepreneurial opportunities is also considered an effective measure in our results.

5.3. *Limitations and suggestions for future research*

Our research still comes with some limitations. First, the gauging of certain mechanism variables could benefit from refinement in

subsequent studies. One potential concern with the two proxy variables for risk attitude pertains to the unique subset of individuals who might have access to enhanced social security provisions. We postulated that individuals with elevated personal or household incomes would be more inclined to engage in high-risk financial markets and accrue non-essential debts. It would be beneficial for future research to validate our findings using more objective measures for risk attitude.

Second, our analysis did not capture all potential channels. Notable omissions include elements like physical capital and financial constraints. Prior research emphasized that adverse childhood experiences may lead to negative effects in adulthood in terms of physical capital accumulation and credit constraints (e.g., Bernile et al., 2017; Chen & Hu, 2019). However, as mentioned earlier, parental migration may help left-behind children accumulate primary entrepreneurial capital (Ajefu & Massacky, 2023). Therefore, theoretically this mechanism is not clear enough and more empirical evidence is needed. However, constrained by the information in the CLDS, we were unable to test this mechanism.

Third, the CLDS only provides information about left-behind experiences in the fourth wave, which allows us to use only cross-sectional data, not panel data. While cross-sectional data is appropriate in the literature exploring the long-term effects of past experiences (Yi et al., 2022), using panel data captures more dynamic information (Zhao & Guo, 2022). Therefore, subsequent studies might try to employ panel data to further validate the reliability of our results.

Last, similar to previous studies involving childhood experiences, the potential for memory error exists because these particular experiences or events were self-reported by the respondent in the survey, as public survey databases were used in our study. Therefore, potential self-selection bias is difficult to avoid in this type of research (Prakash & Kumar, 2021; Zhao & Li, 2022). Also, the CLDS also lack information on the duration of left-behind experiences, which is not conducive to detailed mining for the entrepreneurial effects of left-behind experiences. Follow-up studies may provide more information on left-behind experiences and minimize this potential bias by using a self-designed questionnaire and setting some constraint conditions.

6. Conclusion

Our conclusions are summarized below. First, left-behind experiences in childhood are detrimental to entrepreneurial activity in adulthood in China. This result is supported by some robustness tests, such as alternative specifications, additional controls, and the exclusion of other childhood experiences. Second, through mediating effects, we found that left-behind experiences are detrimental to entrepreneurial behaviors in adulthood owing to disadvantages in human capital and social networks as well as risk averse traits. Third, in the heterogeneity analysis, the results show that left-behind experiences at different school ages have negative impacts on entrepreneurial activities in adulthood; entrepreneurial behaviors of younger adults are more likely to receive negative shocks from left-behind experiences. Last, we found that *hukou* transfer and WTO accession events weaken the negative impact of left-behind experiences on entrepreneurship. The lack of entrepreneurial activities may further explains the long-term income penalty effect of left-behind experiences. Also, our results imply that adults with left-behind experiences do not have survival entrepreneurial motives.

Ethical approval

This paper does not contain any studies with human participants performed by any of the authors. The data we used in this study come from the China Labor-force Dynamics Survey and the China Family Panel Studies. These surveys and data are publicly released. We strictly adhere to the rules governing the use of these databases.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

The authors acknowledge the support from the National Natural Science Foundation of China (No. 72304102; No.72373043), National Social Science Fund of China (No. 23&ZD112), and the Philosophy and Social Science Program Discipline Co-construction Project of Guangdong Province (No. GD23XYJ68).

Appendix

Table A1
Definitions of main variables.

Variables	Definitions
Panel A. Entrepreneurship	
Entrepreneurship	= 1 if the respondent was an employer; 0 otherwise
Entrepreneurial income	Total business income (10,000 yuan, RMB)
Entrepreneurial intentions	= 1 if the respondent intends to start a business in the next two years; 0 otherwise
Panel B. Left-behind experiences in childhood	
Left-behind experience	= 1 if the respondent had left-behind experiences (at least one parent migrated) in childhood; 0 otherwise
Both parents migrate	= 1 if both parents migrated in childhood; 0 otherwise
Left-behind at preschool	= 1 if at least one parent migrated at preschool (0–6 years old); 0 otherwise
Left-behind at primary school	= 1 if at least one parent migrated at primary school stage (7–12 years old); 0 otherwise
Left-behind at junior high school	= 1 if at least one parent migrated at junior high school stage (13–15 years old); 0 otherwise
Left-behind at senior high school	= 1 if at least one parent migrated at senior high school stage (16–18 years old); 0 otherwise
Panel C. Instrumental variable	
Left-behind percentage	Proportion of left-behind children at the birth cohort level from the other database, China Family Panel Studies
Panel D. Control variables	
Age	Age of the respondents, normalization

(*continued on next page*)

Table A1 (*continued*)

Variables	Definitions
Gender	Female = 1; male = 0
Father's *hukou* at birth	Respondent's father's *hukou* status at birth. Rural = 1; non-rural = 0
Mother's *hukou* at birth	Respondent's mother's *hukou* status at birth. Rural = 1; non-rural = 0
Education of father	Normalization. Unschooled = 0; primary school = 1; junior high school = 2; senior high school = 3; college or bachelor's degree = 4; Master's or doctoral degrees = 5
Education of mother	Normalization. Unschooled = 0; primary school = 1; junior high school = 2; senior high school = 3; college or bachelor's degree = 4; Master's or doctoral degrees = 5
Siblings	Number of siblings of the respondent, normalization

Notes: ***, **, * denote the significance levels of 1 %, 5 %, and 10 %, respectively.

Table A2
Alternative estimated methods.

	(1)	(2)	(3)	(4)
	Heckman two-step estimations		Propensity score matching	Endogenous treatment effect model
Left-behind experience	−0.0560**	−0.1415***	−0.1043**	−0.1248***
	(0.0231)	(0.0335)	(0.0506)	(0.0310)
Inverse Mills Ratio	−0.8750	0.3326***		
	(0.7855)	(0.1024)		
Baseline controls	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
2-year birth cohort FE	Yes	Yes	Yes	Yes
Observations	2918	2918	1872	2918

Notes: The first three columns report the marginal effects. All columns report the second-stage results. ***, **, * denote the significance levels of 1 %, 5 %, and 10 %, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

Table A3
Alternative core variables.

	(1)	(2)	(3)	(4)
	Alternative instrumental variable	Alternative explained variable	Alternative explanatory variable	
	Left behind proportion at the village level	Number of business start-ups	Both parents migrate	Migration before junior high school
Left-behind experience	−0.1373***	−0.0517**	−0.2283***	−0.2810***
	(0.0292)	(0.0205)	(0.0502)	(0.0386)
Baseline controls	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
2-year birth cohort FE	Yes	Yes	Yes	Yes
Observations	2918	2918	2918	2918

Notes: Columns (1), (3), and (4) report the marginal effects by using the endogenous bivariate probit model. the endogenous treatment effect model is used in column (2). All columns report the second-stage results. ***, **, * denote the significance levels of 1 %, 5 %, and 10 %, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

Table A4
Additional controls (CMP methods).

	(1)	(2)	(3)	(4)
	Village characteristics	Village FE	Province FE*Birth cohort trend	Province-by-2-year birth cohort FE
Left-behind experience	−0.1380***	−0.4642***	−0.1472***	−0.1791***
	(0.0355)	(0.0502)	(0.0405)	(0.0511)
Baseline controls	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
2-year birth cohort FE	Yes	Yes	Yes	Yes
Village FE	No	Yes	Yes	No
Province FE*Birth cohort trend	No	No	Yes	No
Province-by-2-year birth cohort FE	No	No	No	Yes
Observations	1662	2918	2918	2918

Notes: The endogenous bivariate probit model is employed in all columns, reporting the marginal effects of second-stage results. ***, **, * denote the significance levels of 1 %, 5 %, and 10 %, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

Table A5
Definitions of village-level variables.

Variables	Definitions
Migrant population	Proportion of migrant population to total population at the village level
Topography	Normalization. Plains = 1; hills = 2; mountains = 3
Surnames	Normalization. Number of surnames in the village
Clan	= 1 if the village had a clan organization or genealogy; 0 otherwise
Ethnic minority	= 1 if there is an ethnic minority in the village; 0 otherwise

(*continued on next page*)

Table A5 (*continued*)

Variables	Definitions
Distance from the county seat	Normalization. Distance of the villages from the nearest county town (kilometers, logarithmic)
Village infrastructure	Normalization. Sum scores across dimensions (yes = 1; no = 0): primary schools, middle schools, playgrounds, libraries, plazas or parks, and senior citizen activity rooms

Table A6
Other childhood experiences (CMP methods).

	(1)	(2)	(3)	(4)	(5)
	Parental divorce in childhood	Poverty in childhood	Migration in childhood	Natural disasters in childhood	All these childhood experiences
Left-behind experience	−0.2010***	−0.1665***	−0.2166***	−0.1755***	−0.2269***
	(0.0396)	(0.0390)	(0.0490)	(0.0340)	(0.0678)
Baseline controls	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
2-year birth cohort FE	Yes	Yes	Yes	Yes	Yes
Observations	2781	2558	2278	2918	1918

Notes: The endogenous bivariate probit model is employed in all columns, reporting the marginal effects of second-stage results. ***, **, * denote the significance levels of 1 %, 5 %, and 10 %, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

Table A7
Definitions of mediating variables.

Variables	Definitions
Panel A. Education and skills	
Education	Normalization. 1 = no formal education; 2 = primary education; 3 = junior high school; 4 = senior high school (including vocational school); 5 = college/bachelor's degree; 6 = master's or doctoral degree.
College	= 1 if educational level of the respondent is at college or bachelor's degree or above; 0 otherwise
Language skill	= 1 if the respondent was good at a foreign language; 0 otherwise
Vocational certificate	= 1 if the respondent had an occupational certificate; 0 otherwise
Entrepreneurship training	= 1 if the respondent experienced entrepreneurial training; 0 otherwise
Panel B. Physical and mental health	
Self-reported health	Normalization. "How would you rate your current health status?" 5 = "excellent"; 4 = "good"; 3 = "neutral/average"; 2 = "below average"; 1 = "poor"
Injury	= 1 if the respondent suffered a work-related injury; 0 otherwise
Emotional issues	= 1 if the respondent experienced emotional problems (such as depressed or anxious) in the past month; 0 otherwise
Mental health impairment	Normalization. Rate the average scores across four aspects of evaluating life: failure, hopelessness, meaninglessness, and unsustainability. Use the scale: 1 for "never", 2 for "rarely", 3 for "often", and 4 for "almost always"
Panel C. Social networks	
Entrepreneurial opportunities from friends	Yes = 1; no = 0
Start-up capital from relatives or friends	Yes = 1; no = 0
Local friends with lending money	= 1 if the respondent had local friends who can borrow more than 5,000 yuan (RMB); 0 otherwise
Social interactions	Normalization. Calculate the sum of scores across eight dimensions: neighborhood committee, social work organizations, owners' committee, recreational or sports clubs, training institutions, hometown associations, clan organizations, charitable entities, and religious organizations.
Money gift	Normalization. Expenditure on household money gifts last year (yuan, RMB, logarithm)
Panel D. Risk appetite	
Non-essential debts	= 1 if the household had non-essential debt such as credit cards; 0 otherwise
Risky financial assets	= 1 if the household bought financial risk assets (stocks, funds, or bonds); 0 otherwise

Table A8
Some moderating effects.

	(1) Linear probability model	(2) Probit	(3) Linear probability model	(4) Probit	(5) Linear probability model	(6) Probit
Left-behind experience	−0.0633**	−0.0651***	−0.0617**	−0.0624***	−0.0783**	−0.0832***
	(0.0237)	(0.0233)	(0.0248)	(0.0236)	(0.0278)	(0.0282)
Hukou transfer	−0.0442	−0.0412				
	(0.0267)	(0.0251)				
Left-behind experience × *Hukou* transfer	0.0504	0.0575*				
	(0.0304)	(0.0298)				
Clan			−0.0196	−0.0172		
			(0.0140)	(0.0126)		
Left-behind experience × Clan			0.0163	0.0169		
			(0.0206)	(0.0222)		
WTO accession					−0.1387	−0.1390
					(0.1816)	(0.1844)
Left-behind experience × WTO accession					0.1014**	0.1037***

(*continued on next page*)

Table A8 (*continued*)

	(1) Linear probability model	(2) Probit	(3) Linear probability model	(4) Probit	(5) Linear probability model	(6) Probit
					(0.0388)	(0.0359)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
2-year birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2918	2918	2918	2918	2918	2918

Notes: The probit model is employed in columns (2), (4), and (6), reporting the marginal effects. ***, **, * denote the significance levels of 1 %, 5 %, and 10 %, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

Table A9
Entrepreneurial performances (CMP methods).

	(1) Entrepreneurial income	(2) Entrepreneurial intentions	(3) Survival entrepreneurship
Left-behind experience	0.0043	−0.0765	0.0502
	(0.0882)	(0.0807)	(0.0902)
Baseline controls	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
2-year birth cohort FE	Yes	Yes	Yes
Observations	2835	2918	2610

Notes: The endogenous treatment effect model is employed in column (1) and the endogenous bivariate probit model is employed in columns (2) and (3), reporting the marginal effects. All columns report the second-stage results. ***, **, * denote the significance levels of 1 %, 5 %, and 10 %, respectively. The standard errors clustered at the birth cohort level are shown in parentheses.

Table A10
Entrepreneurship activities and long-term economic outcomes of left-behind experience (CMP methods).

	(1) Individual income	(2) Total household income	(3) Household operating income	(4) Low income
Panel A. Left-behind experiences and income				
Left-behind experience	−0.8008*	−0.6731	−0.1505*	0.1402
	(0.4390)	(0.7763)	(0.0871)	(0.1408)
Observations	2197	2846	2847	2918
Panel B. Mediation effects of entrepreneurial activities				
Left-behind experience	−0.7429	−0.5571	−0.1449	0.0983
	(0.5329)	(0.9039)	(0.1171)	(0.1266)
Entrepreneurship	0.4975	0.6891	0.0488	−0.0541**
	(0.3396)	(0.6123)	(0.0943)	(0.0217)
Observations	2197	2846	2847	2918

Notes: In two Panels, the endogenous treatment effect model is employed in columns (1)-(3) and the endogenous bivariate probit model is employed in column (4), reporting the marginal effects. All columns report the second-stage results. The baseline control variables, province FE, and 2-year birth cohort FE are added in all regressions for two Panels. ***, **, * denote the significance levels of 1 %, 5 %, and 10 %, respectively. The standard errors clustered at the individual level are shown in parentheses.

Table A11
Definitions of variables in further studies.

Variables	Definitions
Individual income	Total individual income from the previous year in units of 10,000 yuan (RMB)
Total household income	Total household income, including all family members, from the previous year in units of 10,000 yuan (RMB)
Household operating income	Total income from operating activities, such as rental income and sales income from the previous year in units of 1,000 yuan (RMB)
Low income	= 1 if the family received relief or basic living allowance; 0 otherwise

References

Ahn, K., & Winters, J. V. (2023). Does education enhance entrepreneurship? *Small Business Economics, 61*(2), 717–743.

Ajefu, J. B., & Massacky, F. (2023). Mobile money, child labour and school enrolment. *Telecommunications Policy, 47*(10), Article 102648.

Aregbeshola, R. A. (2022). Interplay of poverty, remittances and human capital development: Panel evidence from selected Sub-Saharan African countries. *International Migration, 60*(6), 250–271.

Banerjee, A., Breza, E., Chandrasekhar, A. G., Duflo, E., Jackson, M. O., & Kinnan, C. (2024). Changes in social network structure in response to exposure to formal credit markets. *Review of Economic Studies, 91*(3), 1331–1372.

Baron, R. A., Franklin, R. J., & Hmieleski, K. M. (2016). Why entrepreneurs often experience low, not high, levels of stress: The joint effects of selection and psychological capital. *Journal of Management, 42*(3), 742–768.

Benmelech, E., & Frydman, C. (2015). Military CEOs. *Journal of Financial Economics, 117*(1), 43–59.

Bernile, G., Bhagwat, V., & Rau, P. R. (2017). What doesn't kill you will only make you more risk-loving: Early-life disasters and CEO behavior. *Journal of Finance, 72*(1), 167–206.

Burt, R. S. (2019). The networks and success of female entrepreneurs in China. *Social Networks, 58*, 37–49.

Cassar, A., Healy, A., & Von Kessler, C. (2017). Trust, risk, and time preferences after a natural disaster: Experimental evidence from Thailand. *World Development, 94*, 90–105.

Caucutt, E. M., & Lochner, L. (2020). Early and late human capital investments, borrowing constraints, and the family. *Journal of Political Economy, 128*(3), 1065–1147.

Chang, F., Jiang, Y., Loyalka, P., Chu, J., Shi, Y., Osborn, A., & Rozelle, S. (2019). Parental migration, educational achievement, and mental health of junior high school students in rural China. *China Economic Review, 54*, 337–349.
Chen, M., & Chan, K. L. (2016). Parental absence, child victimization, and psychological well-being in rural China. *Child Abuse & Neglect, 59*, 45–54.
Chen, J., & Hu, M. (2019). What types of homeowners are more likely to be entrepreneurs? The evidence from China. *Small Business Economics, 52*, 633–649.
Chen, J., & Hu, M. (2021). City-level hukou-based labor market discrimination and migrant entrepreneurship in China. *Technological and Economic Development of Economy, 27*(5), 1095–1118.
Chen, B., & Zhao, C. (2022). More is less: Homeroom teachers' administrative duties and students' achievements in China. *Teaching and Teacher Education, 119*, Article 103857.
Cheng, Z., Guo, W., Hayward, M., Smyth, R., & Wang, H. (2021). Childhood adversity and the propensity for entrepreneurship: A quasi-experimental study of the Great Chinese Famine. *Journal of Business Venturing, 36*(1), Article 106063.
Cheng, Z., & Smyth, R. (2021). Education and migrant entrepreneurship in urban China. *Journal of Economic Behavior & Organization, 188*, 506–529.
Cheung, C. K. (2024). Housing Welfare, Personal Responsibility, and Morale in Hong Kong. *Social Indicators Research, 171*(3), 905–919.
Cheung, C. K., & Tsang, E. Y. H. (2015). Political-economic coalition among entrepreneurs, professionals, and cadres in Guangdong. *China. Sociological Research Online, 20*(4), 45–57.
Cheung, C. K., & Chow, E. O. W. (2011). Interpersonal influences between the care provider's burnout and the older care recipient's quality of life. *Administration in Social Work, 35*(4), 425–445.
Churchill, S. A., Munyanyi, M. E., Smyth, R., & Trinh, T. A. (2021). Early life shocks and entrepreneurship: Evidence from the Vietnam War. *Journal of Business Research, 124*, 506–518.
Cuadros-Meñaca, A. (2020). Remittances, health insurance, and pension contributions: Evidence from Colombia. *World Development, 127*, Article 104766.
Deng, W., Liang, Q., Fan, P., & Cui, L. (2020). Social entrepreneurship and well-being: The configurational impact of institutions and social capital. *Asia Pacific Journal of Management, 37*, 1013–1037.
Elder, G. H., Jr (1998). The life course as developmental theory. *Child development, 69*(1), 1–12.
Gee, L. K., Jones, J., & Burke, M. (2017). Social networks and labor markets: How strong ties relate to job finding on Facebook's social network. *Journal of Labor Economics, 35*(2), 485–518.
George, L. K. (2014). Taking time seriously: A call to action in mental health research. *Journal of Health and Social Behavior, 55*(3), 251–264.
Guo, S., Cheung, C. K., Ning, X., & Low, Y. T. A. (2019). Moderation effect of social capital on flourishing among left-behind and other children in rural China. *Asia Pacific Journal of Social Work and Development, 29*(4), 278–293.
Guo, S., Cheung, C. K., Qin, T., & Ning, X. (2021). Intrinsic motivation of rural adolescents in China: How individuation influences? *Children and Youth Services Review, 120*, Article 105778.
Han, S., & Vytlacil, E. J. (2017). Identification in a generalization of bivariate probit models with dummy endogenous regressors. *Journal of Econometrics, 199*(1), 63–73.
Hou, B., Liu, H., & Wang, S. X. (2020). Returns to military service in off-farm wage employment: Evidence from rural China. *China Economic Review, 59*, Article 101361.
Huang, G. Y., Shen, C. H. H., & Wu, Z. X. (2023). Firm-level political risk and debt choice. *Journal of Corporate Finance, 78*, Article 102332.
Ida, R., Saud, M., & Mashud, M. I. (2020). Persistence of social media on political activism and engagement among Indonesian and Pakistani youths. *International Journal of Web Based Communities, 16*(4), 378–395.
Kihlstrom, R. E., & Laffont, J. J. (1979). A general equilibrium entrepreneurial theory of firm formation based on risk aversion. *Journal of Political Economy, 87*(4), 719–748.
Kwon, S. W., & Ruef, M. (2017). The imprint of labor markets on entrepreneurial performance. *Journal of Business Venturing, 32*(6), 611–626.
Leung, D. Y., Chan, A. C., & Ho, G. W. (2022). Resilience of emerging adults after adverse childhood experiences: A qualitative systematic review. *Trauma, Violence, & Abuse, 23*(1), 163–181.
Leyden, D. P., Link, A. N., & Siegel, D. S. (2014). A theoretical analysis of the role of social networks in entrepreneurship. *Research Policy, 43*(7), 1157–1163.
Li, M., Duan, X., Shi, H., Dou, Y., Tan, C., Zhao, C., & Zhang, J. (2021). Early maternal separation and development of left-behind children under 3 years of age in rural China. *Children and Youth Services Review, 120*, Article 105803.
Li, Q., Liu, G., & Zang, W. (2015). The health of left-behind children in rural China. *China Economic Review, 36*(6), 367–376.
Li, X., Coid, J. W., Tang, W., Lv, Q., Zhang, Y., Yu, H., & Li, T. (2021). Sustained effects of left-behind experience during childhood on mental health in Chinese university undergraduates. *European Child & Adolescent Psychiatry, 30*, 1949–1957.
Liang, L., Yang, Y., & Xiao, Q. (2020). Young people with left-behind experiences in childhood have higher levels of psychological resilience. *Journal of Health Psychology, 25*(13–14), 2453–2463.
Liu, M., & Villa, K. M. (2020). Solution or isolation: Is boarding school a good solution for left-behind children in rural China? *China Economic Review, 61*, Article 101456.
Liu, Z., Yang, S., & Wang, Y. (2022). Window of time and trauma of time: The temporal effect of left-behind experience on subjective well-being. *Chinese Journal of Sociology, 42*(6), 188–213. in Chinese.
Lyons, E., & Zhang, L. (2017). The impact of entrepreneurship programs on minorities. *American Economic Review, 107*(5), 303–307.
Lyu, L., & Chen, Y. (2019). Parental migration and young migrants' wages in urban China: An exploratory analysis. *Urban Studies, 56*(10), 1968–1987.
MacLachlan, I., & Gong, Y. (2022). China's new age floating population: Talent workers and drifting elders. *Cities, 131*, Article 103960.
Marquis, C., & Qiao, K. (2020). Waking from Mao's dream: Communist ideological imprinting and the internationalization of entrepreneurial ventures in China. *Administrative Science Quarterly, 65*(3), 795–830.
Marvel, M. R., Davis, J. L., & Sproul, C. R. (2016). Human capital and entrepreneurship research: A critical review and future directions. *Entrepreneurship Theory and Practice, 40*(3), 599–626.
McFarland, M. J., McLanahan, S. S., Goosby, B. J., & Reichman, N. E. (2017). Grandparents' education and infant health: Pathways across generations. *Journal of Marriage and Family, 79*(3), 784–800.
Meng, X., & Yamauchi, C. (2017). Children of migrants: The cumulative impact of parental migration on children's education and health outcomes in China. *Demography, 54*(5), 1677–1714.
Moore, E. M., Dau, L. A., & Mingo, S. (2021). The effects of trade integration on formal and informal entrepreneurship: The moderating role of economic development. *Journal of International Business Studies, 52*, 746–772.
Mumtaz, Z. (2022). Informal social protection: A conceptual synthesis. *Social Policy & Administration, 56*(3), 394–408.
Nguyen, A. T., Do, T. H. H., Vu, T. B. T., Dang, K. A., & Nguyen, H. L. (2019). Factors affecting entrepreneurial intentions among youths in Vietnam. *Children and Youth Services Review, 99*, 186–193.
Nguyen, C. V. (2016). Does parental migration really benefit left-behind children? Comparative evidence from Ethiopia, India, Peru and Vietnam. *Social Science & Medicine, 153*, 230–239.
Nason, R., & Bothello, J. (2023). Far from void: How institutions shape growth in informal economies. *Academy of Management Review, 48*(3), 485–503.
Prado, A. M., Robinson, J. A., & Shapira, Z. (2022). Serving rural low-income markets through a social entrepreneurship approach: Venture creation and growth. *Strategic Entrepreneurship Journal, 16*(4), 826–852.
Prakash, K., & Kumar, S. (2021). "Smoking your child's job away": Parental smoking during one's childhood and the probability of being employed in adulthood. *Journal of Business Research, 136*, 86–98.
Putniņš, T. J., & Sauka, A. (2020). Why does entrepreneurial orientation affect company performance? *Strategic Entrepreneurship Journal, 14*(4), 711–735.
Qin, N., & Kong, D. (2021). Human capital and entrepreneurship. *Journal of Human Capital, 15*(4), 513–553.
Qiu, H., Liang, X., & Sun, D. (2024). Parental migration, sibling migration, and the educational outcomes of children left behind in rural China. *Children and Youth Services Review, 157*, Article 107392.
Rosado-Cubero, A., Freire-Rubio, T., & Hernández, A. (2022). Entrepreneurship: What matters most. *Journal of Business Research, 144*, 250–263.
Saud, M., Ida, R., & Mashud, M. I. (2020). Democratic practices and youth in political participation: A doctoral study. *International Journal of Adolescence and Youth, 25*(1), 800–808.
Saud, M., & Margono, H. (2021). Indonesia's rise in digital democracy and youth's political participation. *Journal of Information Technology & Politics, 18*(4), 443–454.
Schoon, I., & Duckworth, K. (2012). Who becomes an entrepreneur? Early life experiences as predictors of entrepreneurship. *Developmental Psychology, 48*(6), 1719.
Stephan, U., Rauch, A., & Hatak, I. (2023). Happy entrepreneurs? Everywhere? A meta-analysis of entrepreneurship and wellbeing. *Entrepreneurship Theory and Practice, 47*(2), 553–593.
Sun, M., Xue, Z., Zhang, W., Guo, R., Hu, A., Li, Y., & Rosenheck, R. (2017). Psychotic-like experiences, trauma and related risk factors among "left-behind" children in China. *Schizophrenia Research, 181*, 43–48.
Van Burg, E., Elfring, T., & Cornelissen, J. P. (2022). Connecting content and structure: A review of mechanisms in entrepreneurs' social networks. *International Journal of Management Reviews, 24*(2), 188–209.
Van Ness, R. K., & Seifert, C. F. (2016). A theoretical analysis of the role of characteristics in entrepreneurial propensity. *Strategic Entrepreneurship Journal, 10*(1), 89–96.
Wang, C., Shen, J., & Liu, Y. (2023). Hukou transfer intention of rural migrants with settlement intention in China: How cities' administrative level matters. *Journal of Rural Studies, 99*, 1–10.
Wang, H., Cheng, Z., Wang, B. Z., & Chen, Y. (2021). Childhood left-behind experience and labour market outcomes in China. *Journal of Business Research, 132*, 196–207.
Wang, L., Zheng, Y., Li, G., Li, Y., Fang, Z., Abbey, C., & Rozelle, S. (2019). Academic achievement and mental health of left-behind children in rural China: A causal study on parental migration. *China Agricultural Economic Review, 11*(4), 569–582.
Wang, Y., Shen, Z., Liu, W., Liu, Y., & Tang, B. (2024). Will the situation of Left-behind children improve when their parents return? Evidence from China. *Children and Youth Services Review, 164*, Article 107856.
Wen, C., Zhao, X., Xu, L., & Yin, H. (2023). Military experience and household stock market participation: Evidence from China. *International Review of Financial Analysis, 86*, Article 102519.
Wen, Y. J., Hou, W. P., Zheng, W., Zhao, X. X., Wang, X. Q., Bo, Q. J., & Wang, C. Y. (2021). The neglect of left-behind children in China: A meta-analysis. *Trauma, Violence, & Abuse, 22*(5), 1326–1338.
White, K. R., & Wu, Q. (2014). Application of the life course perspective in child welfare research. *Children and Youth Services Review, 46*, 146–154.
Wolfolds, S. E., & Siegel, J. (2019). Misaccounting for endogeneity: The peril of relying on the Heckman two-step method without a valid instrument. *Strategic Management Journal, 40*(3), 432–462.
Wu, Q., Lu, D., & Kang, M. (2015). Social capital and the mental health of children in rural China with different experiences of parental migration. *Social Science & Medicine, 132*, 270–277.

Wu, Y., & Guo, Z. (2020). An analysis of the nutritional status of left-behind children in rural China and the impact mechanisms of child malnutrition. *Children and Youth Services Review, 119*, Article 105598.
Xu, P., & Cheng, J. (2021). Individual differences in social distancing and mask-wearing in the pandemic of COVID-19: The role of need for cognition, self-control and risk attitude. *Personality and Individual Differences, 175*, Article 110706.
Yang, B., Xiong, C., & Huang, J. (2021). Parental emotional neglect and left-behind children's externalizing problem behaviors: The mediating role of deviant peer affiliation and the moderating role of beliefs about adversity. *Children and Youth Services Review, 120*, Article 105710.
Yang, H., & Zhang, X. (2023). Persistence of culture: How the entrepreneurial culture of origin contributes to migrant entrepreneurship. *Small Business Economics, 61*(3), 1179–1204.
Yang, S., Wang, Y., Lu, Y., Zhang, H., Wang, F., & Liu, Z. (2023). Long-term effects of the left-behind experience on health and its mechanisms: Empirical evidence from China. *Social Science & Medicine, 338*, Article 116315.
Yi, J., Chu, J., & Png, I. P. (2022). Early-life exposure to hardship increased risk tolerance and entrepreneurship in adulthood with gender differences. *Proceedings of the National Academy of Sciences, 119*(15). e2104033119.
Yu, W., Stephan, U., & Bao, J. (2023). Childhood adversities: Mixed blessings for entrepreneurial entry. *Journal of Business Venturing, 38*(2), Article 106287.
Xiao, Y., Chen, Y., Meng, Q., Tian, X., & Wang, Y. (2019). Suicide ideation and suicide plan in Chinese left-behind children: Prevalence and associated factors. *Journal of Affective Disorders, 257*, 662–668.
Zhang, C. (2020). Clans, entrepreneurship, and development of the private sector in China. *Journal of Comparative Economics, 48*(1), 100–123.
Zhang, C., Yang, X., & Xu, W. (2021). Parenting style and aggression in Chinese undergraduates with left-behind experience: The mediating role of inferiority. *Children and Youth Services Review, 126*, Article 106011.
Zhang, H., Chi, P., Long, H., & Ren, X. (2019). Bullying victimization and depression among left-behind children in rural China: Roles of self-compassion and hope. *Child Abuse & Neglect, 96*, Article 104072.
Zhang, H., Behrman, J. R., Fan, C. S., Wei, X., & Zhang, J. (2014). Does parental absence reduce cognitive achievements? Evidence from rural China. *Journal of Development Economics, 111*, 181–195.
Zhao, C., & Chen, B. (2022). Parental migration and non-cognitive abilities of left-behind children in rural China: Causal effects by an instrumental variable approach. *Child Abuse & Neglect, 123*, Article 105389.
Zhao, C., & Guo, J. (2022). Are veterans happy? Long-term military service and the life satisfaction of elderly individuals in China. *Journal of Happiness Studies, 23*(2), 477–508.
Zhao, C., & Li, X. (2022). Living under the shadow: Adverse childhood experiences and entrepreneurial behaviors in Chinese adults. *Journal of Business Research, 138*, 239–255.
Zhao, C., Wu, Y., & Guo, J. (2022). Mobile payment and Chinese rural household consumption. *China Economic Review, 71*, Article 101719.
Zhao, C., Xu, G., & Cai, W. (2024). Troops are business schools: Military service and entrepreneurial behaviors in China. *Entrepreneurship Theory and Practice*. https://doi.org/10.1177/104225872412499
Zhao, H., O'Connor, G., Wu, J., & Lumpkin, G. T. (2021). Age and entrepreneurial career success: A review and a meta-analysis. *Journal of Business Venturing, 36*(1), Article 106007.
Zhao, Q., Yu, X., Wang, X., & Glauben, T. (2014). The impact of parental migration on children's school performance in rural China. *China Economic Review, 31*, 43–54.
Zheng, X., Zhang, Y., & Jiang, W. (2023). Migrating with parents or left-behind: Associations of internal migration with cognitive and noncognitive outcomes among Chinese children. *Current Psychology, 42*(22), 19208–19229.
Zhou, C., Sylvia, S., Zhang, L., Luo, R., Yi, H., Liu, C., & Rozelle, S. (2015). China's left-behind children: Impact of parental migration on health, nutrition, and educational outcomes. *Health Affairs, 34*(11), 1964–1971.
Zou, H., Chen, X., Lam, L. W. R., & Liu, X. (2016). Psychological capital and conflict management in the entrepreneur-venture capitalist relationship in China: The entrepreneur perspective. *International Small Business Journal, 34*(4), 446–467.

Environ Geochem Health (2023) 45:3003–3024
https://doi.org/10.1007/s10653-022-01386-7

ORIGINAL PAPER

Can national forest city construction mitigate air pollution in China? Evidence from a quasi-natural experiment

Xing Li · Chunkai Zhao

Received: 9 November 2021 / Accepted: 27 August 2022 / Published online: 21 September 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract As air pollution in Chinese cities becomes a growing concern, measures to alleviate air pollution have attracted the attention of all sectors of society. By using the data for 283 prefecture-level cities from 2003 to 2016, we utilized the quasi-natural experiment of the national forest city construction (NFCC) and employed the difference-in-differences approach to examine the effects of NFCC on air pollution. The results show that the NFCC led to a 12.14% and 4.29% reduction in $PM_{2.5}$ concentrations and SO_2 emissions, respectively. A series of robustness tests such as instrumental variable estimates, placebo tests, and eliminating disturbing policies all supported these findings. In addition, we provided evidence that the environment benefits of the NFCC could be explained by increasing green spaces, strengthening environmental regulations, and forming green development models. Furthermore, the results from heterogeneity analysis indicate that the NFCC was more effective in smaller cities, southern cities, and western cities. Our findings are of significance to Chinese cities on the road to sustainable development and provide some insights for other developing countries and emerging markets to control air pollution.

Keywords National forest city construction (NFCC) · Air pollution · Quasi-natural experiment · Difference-in-differences (DID) · China

X. Li
College of Economics and Management, Southwest University, No. 2, Tiansheng Road, Beibei District, Chongqing, China
e-mail: xingli277@126.com

C. Zhao (✉)
College of Economics and Management, South China Agricultural University, No.483, Wushan, Tianhe District, Guangzhou, Guangdong Province, China
e-mail: ckz_zhao@163.com

Introduction

With the accelerated industrialization and urbanization, environmental problems such as air pollution are becoming more and more prominent in China. According to the Global Environmental Performance Index, jointly published by Yale University, Columbia University, and the World Economic Forum, China ranked 120th out of 180 countries and regions in 2018 and even ranked fourth from the bottom in air quality (Wendling et al., 2018). Increasing air pollution brings a series of socioeconomic problems such as endangering public health (Azizullah et al., 2011; Brunekreef & Holgate, 2002; Neidell, 2004), reducing labor productivity (Zivin & Neidell, 2012; Lichter et al., 2017), and inducing social risks (Deschenes et al., 2020; Ostro et al., 2014).

To address the air pollution problem, the Chinese government has implemented a series of policies, such as Air Pollution Prevention and Control Action Plan (APPCAP) (e.g., Cai et al., 2018; Feng et al., 2019; Huang et al., 2018; Ma et al., 2021), clean

energy policies (Chen et al., 2017; Sheehan et al., 2014; Yuan & Zuo, 2011), "Two control Zones" policy (Cai et al., 2016; Hao et al., 2000), and driving restrictions (Han et al., 2020; Zhong et al., 2017). In addition, some previous studies concluded that some policies involving urban development and management could also improve air quality (Fang et al., 2014; Li et al., 2021; Yu, 2014; Yu & Zhang, 2021; Zhu et al., 2019). Different from previous studies, this paper sought to examine whether the national forest city construction (NFCC) helped mitigate air pollution.

The concept of forest cities was first proposed by scholars in the USA and Canada in the 1960s (Xu et al., 2020) and promoted in the USA, Europe, and Japan (Nowak et al., 2006; Tajima, 2003). In China, the national forest city is officially defined as a city whose ecosystem is dominated by forest vegetation and whose ecological construction has achieved integrated urban and rural development (Liao et al., 2021; Zhang et al., 2021). At the China Urban Forestry Forum, the national forestry authorities announce the annual list of national forest cities that meet related standards. Considering that the NFCC is a model centered on urban vegetation and green development in China (Zhang et al., 2021), it may become an important way to mitigate urban air pollution.

Specifically, the NFCC directly increased green spaces and improve vegetation coverage, which reduced the content of pollutants in the air (e.g., Escobedo et al., 2011; Lei et al., 2021; Nowak et al., 2006). In addition, to pass the assessment of NFCC, local governments further strengthened environmental regulation and indirectly improve air quality (Cole et al., 2005; Tanaka, 2015). More importantly, the concept of sustainable development caused by NFCC might give rise to green development models (Pérez-Campuzano et al., 2016; Wang et al., 2013). For example, the NFCC is more conducive to the development of eco-friendly related industries and promotes green innovation. Taken together, the NFCC employed more ecological means to improve the city environment, which is an important direction for sustainable development in Chinese cities.

Since 2004, China has established 192 national forest cities in 26 provinces. As the NFCC is in full swing throughout the country, three studies explored the environmental benefits of NFCC. Xu et al. (2020) took the Beijing–Tianjin–Hebei region, the most serious air pollution in China, as the case, finding that the NFCC effectively reduced $PM_{2.5}$ concentrations. Coincidentally, from the perspective of NFCC and residents' living environment, Zhang et al. (2021) reached a similar conclusion that NFCC had a positive impact on mitigating haze pollution. Yet these studies paid little attention to the direct causal relationship between the NFCC and air pollution mitigation and the indirect socioeconomic mechanisms. Moreover, Liao et al. (2021) concluded that the NFCC could help reduce carbon emissions and promote low-carbon development in Chinese cities.

By regarding the NFCC as a quasi-natural experiment and utilizing the difference-in-differences (DID) approach, we examined the impact of NFCC on air pollution in Chinese cities. This research contributed to the existing literature in the following ways. First, our findings enriched the research field on the environmental and ecological benefits of urban forestry (e.g., Escobedo et al., 2011; Fan et al., 2017; Lei et al., 2021; Margaritis & Kang, 2017; Nowak et al., 2006), and the positive effects of the NFCC on environmental quality improvement in China (Liao et al., 2021; Xiong, 2011; Xu et al., 2020; Zhang et al., 2021). Additionally, unlike previous policies that emphasized administrative and economic instruments (e.g., Cai et al., 2018; Feng et al., 2019; Fu et al., 2021; Li et al., 2021; Sheehan et al., 2014; Zhong et al., 2017; Zhu et al., 2019), we provided evidence that the NFCC, which concentrated on an ecological approach, is also effective in reducing air pollution in China.

Second, this study further supplemented the potential mechanisms that NFCC alleviated air pollution. The previous literature mainly believed that green spaces, forest coverage, urban sanitation and investment in environmental facilities were important ways for the NFCC to play a role (Liao et al., 2021; Xu et al., 2020; Zhang et al., 2021). We have innovatively increased the investigation on channels in terms of environmental regulations and green development models. Compared to afforestation and green spaces, the NFCC raised the government and public attention to environmental protection in cities and strengthening of environmental regulations, and the formation of green development models might bring long-term environmental benefits.

Third, our heterogeneous results showed that the NFCC worked better for less economically developed

western regions and cities with relatively small populations, which provided insight into how some large cities could develop better measures to leverage the role of NFCC in improving air quality. Moreover, we found that compared to more polluted and sparsely vegetated northern cities, the NFCC played a more positive role in southern cities, which provided suggestive evidence that some Chinese northern cities promoted sustainable development by the NFCC.

The NFCC in China

The idea of forest cities in China originated in the 1980s, but it was not until 2004 that the official process of evaluating "national forest cities" was launched. In 2007, the National Greening Committee and the National Forestry Administration of China[1] promulgated the "National Forest Cities Evaluation Indexes" and the "National Forest Cities declaration methods", which further standardized the evaluation criteria and certification of national forest cities. The China Urban Forest Forum held every year announces the list of national forest cities and delves into various theoretical and practical issues of the NFCC. In 2013, the National Forestry Administration released the "Outline for Promoting Ecological Civilization (2013–2020)", which again emphasized the role of the NFCC in urban ecological restoration. In addition, "Thirteenth Five-Year Plan for National Economic and Social Development" also explicitly listed the NFCC as a national strategy for implementing China's ecological civilization.

Since Guiyang became the first national forest city in China in 2004, 192 cities (including municipalities, prefecture-level cities, districts, and counties) have been rated as NFCCs as of 2019. As shown in Fig. 1, national forest cities have spread across 26 provinces in China and are concentrated in the Shandong Peninsula, the Central Plains, the Yangtze River Basin, and the Pearl River Delta. Overall, there are more national forest cities in the southern and east-central provinces. At present, NFCC has become an important city brand, playing an active role in protecting forest resources, promoting green culture, creating livable cities, and improving the urban ecological environment (Liao et al., 2021; Xu et al., 2020).

Selection items of national forest cities include 7 aspects, namely comprehensive index, forest coverage, forest ecological network, forest health, public recreation, ecological culture, and green countryside, and each of them has relatively strict quantitative criteria.[2] In terms of forest coverage, the total forest coverage of the northern and southern cities should reach 40% and 30%, respectively, while the greening rate of built-up areas needs to be over 35% and the total public green space per capita ought to exceed 9 square meters. From the aspect of ecological network, greening rates of waterfront (rivers or lakes) and roads (roads or railroads) are required to reach 80% or more. For ecological culture construction, national forest cities need to hold various ecological science activities more than three times a year, and build more than two ecological science knowledge education bases or places. Furthermore, the public awareness rate of creating a national forest city is supported to reach above 90%, and the support rate should be at least 80%. The State Forestry Administration organizes expert to assess the national forest city designation after 3 years. Cities will be warned if do not meet the above requirements. After two years, these failed cities will be re-examined, and if they still do not reach the standards, the title of national forest city will be canceled.

Mechanism analysis

In this section, we discussed in detail the mechanisms of the NFCC on air pollution. We classified possible mechanisms into the following three categories: green spaces, environmental regulations, and green development models.

Green spaces

Forest cities focus on forest construction and forestry protection to reshape the city's ecosystem and environment by increasing green vegetation (Zhang et al., 2021). With the NFCC, the city became an effective

[1] In 2018, the National Forestry Administration was changed to the National Forestry and Grassland Administration.

[2] Information source: http://www.forestry.gov.cn/zlszz/4249/20190807/091811888773590.html.

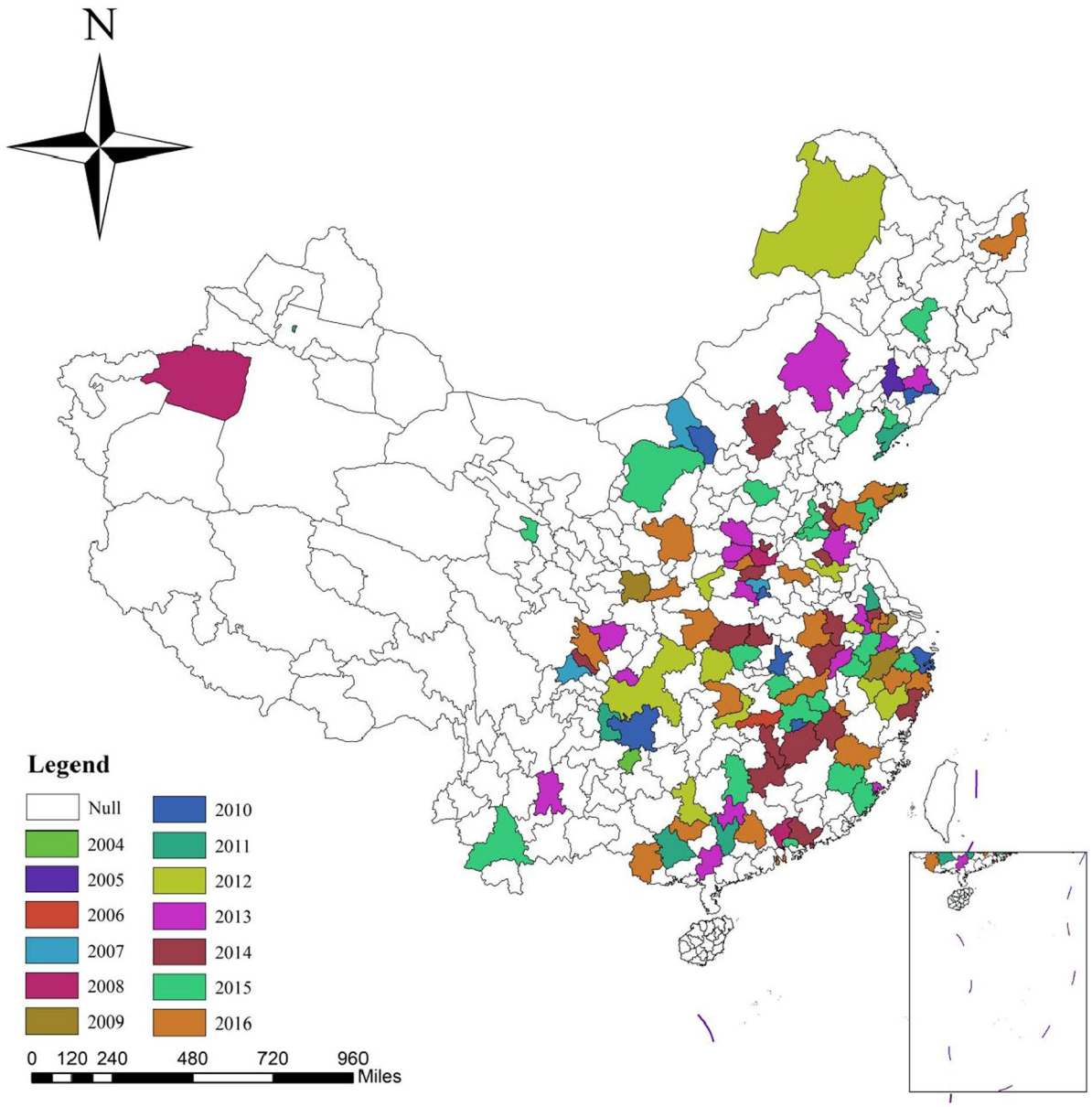


Fig. 1 Geographical distribution of China's National Forest Cities over the years

carrier of forestry resources. Also, plenty of recreational parks sprung up and the green area and vegetation cover increased (Xiong, 2011), bringing a better ecological environment and less pollution (Lei et al., 2021). In China, some cities have increased their green spaces and forest coverage after the NFCC by planting trees and creating urban forest parks (Xu et al., 2020). For example, from 2017 to 2019, the Baoding municipal government established 51 forest demonstration sites, with afforestation area of 933 square kilometers, which promoted the overall forest coverage to 38.1%.[3]

The positive role of urban greening on mitigating air pollution has been supported by many previous studies. On the one hand, some plants can absorb

[3] Data source: https://www.sohu.com/a/340348611_329129.

airborne pollutants such as $PM_{2.5}$, nitrogen oxides, and sulfides (Escobedo et al., 2011; Nowak et al., 2006). On the other hand, some urban greening plants can reduce and lower the airborne dust and particulate matter in large quantities (Lei et al., 2021; Liu et al., 2013; McDonald et al., 2007). For example, on average, a hectare of lawn can absorb more than 30 tons of soot per year in cities (Nowak et al., 2006). Since the trees in the city are cultivated and selected for years, in the context of NFCC, local governments should invest more special funds to improve the survival rate of trees and their positive effect on the environment. What's more, some artificially cultivated urban tree species have more significant purifying effect on urban waste gas compared to natural trees (Churkina et al., 2015).

To sum up, we considered that as a typical integrated urban ecosystem dominated by forest vegetation, the NFCC can be an effective approach to mitigate air pollution by enhancing forest cover and green spaces.

Environmental regulations

In addition to directly increasing forest cover, as one of the ecological and cultural brands of cities, the initial intention of NFCC is to establish an ecosystem with a good habitat and harmonious coexistence between the humans and nature (Xu et al., 2020). Therefore, to achieve these goals, the NFCC is likely to prompt local governments to increase environmental regulation, which contributes to reducing emissions and improving air quality.

As highlighted in "The NFCC in China" section, the NFCC is subject to periodic evaluations. If a city fails in two rounds of assessment, the national forest city designation will be revoked. Thus, under strict assessment criteria, similar to other city-level environmental policies (Wan et al., 2021; Yu & Zhang, 2021), the NFCC may evolve into an environmental regulatory tool for some cities. On the one hand, local governments in forest cities might impose mandatory interventions and constraints on some projects and enterprises that deteriorate air quality (Liao et al., 2021), inducing them to reduce pollutant gas emissions and upgrade pollutant purification technologies (Tanaka, 2015; Zheng et al., 2015). On the other hand, with NFCC, some special funds and fiscal stimulus programs helped some heavy polluters and industries to transform or shift (Sheehan et al., 2014).

Moreover, the NFCC promoted the application of public participation-type environmental regulation tools. Public participation tools rely mainly on individual consciousness, without the need for laws, regulations and monitoring systems (O'rourke & Macey, 2003). As mentioned earlier, the NFCC requires the dissemination of ecological knowledge to the public and the increasing awareness of environmental protection among citizens. As direct perceivers of air pollution information, citizens have a natural information advantage and are able to participate in environmental monitoring as a third-party force (Johnson, 2020; Tu et al., 2019). The public can monitor the policy implementation, provide timely feedback on environmental information, solve the information asymmetry between enterprises and governments (Martens, 2006), and improve the enforcement effectiveness of environmental regulations in national forest cities.

Given that the positive role of environmental regulations in improving air quality in China has been widely confirmed in previous studies (e.g., Cai et al., 2018; Du & Li, 2020; Tanaka, 2015; Xie et al., 2017; Zheng et al., 2015), the NFCC is likely to control air pollution by strengthening environmental regulations.

Green development models

Similar to other city-level energy and environmental policies (e.g., Chen et al., 2017; Pérez-Campuzano et al., 2016; Wang et al., 2013; Yu & Zhang, 2021; Yuan & Zuo, 2011), the NFCC may also construct green development models to enhance air quality and promote sustainable development.

First, the NFCC demands rational and scientific urban planning. On the basis of constructing a good living environment, it is necessary to realize the integration of ecology, culture, and modernization (Xiong, 2011; Zhang et al., 2021). The development philosophy of local governments is likely to change, with more emphasis on low-carbon economy and environmental quality. For example, the NFCC further optimized urban transportation networks and green transportation systems would be established (Li et al., 2005), such as green public transportation systems and introductions of new energy vehicles. Clearly, these measures keep expanding the

environmental benefits of NFCC and alleviating air pollution.

Second, considering the advantages of national forest cities in forest tourism, forest ecology and greening technology, they can further develop industries related to environment and ecological protection based on their advantages (Han et al., 2019). These green industries brought about higher economic value added and less pollutant emissions (Guo et al., 2017; Hart & Ahuja, 1996). By funding green enterprises and developing green industrial parks, green innovation capabilities of national forest cities can be strengthened, which contributes to the creation of an environmentally friendly economic system (Ruiz-Pérez et al., 2001). Additionally, for some resource-based and heavy industrial cities, the NFCC provided them with new development directions. Taking Taiyuan, Benxi, and Xinyu as examples, they turned to develop ecological industries and green industries by NFCC and adjusting industrial structures to reducing pollution and improve the air quality (Xu et al., 2020).

In summary, the NFCC constructed conditions for cities to build green development models. Through the development of green transportation and green industries, the NFCC helped to promote innovative development and control air pollution.

Methodology

Considering the quasi-natural experiment of NFCCs in China, we used the DID model, which is widely used in environmental and ecological economics in recent years (e.g., Cheng et al., 2019; Fu et al., 2021; Tang et al., 2018; Wang & Watanabe, 2019; Yang et al., 2021; Zhong et al., 2017). By comparing the average treatment effects of the treatment and control groups before and after the policy, the DID approach could well exclude the effects of confounding factors to exclude endogeneity. Therefore, we first differenced the treatment and control groups and then differenced the treatment group before and after NFCC to obtain the causal effect. The baseline DID model used in this paper can be written as:

$$AP_{it} = \alpha_0 + \alpha_1 NFCC_{it} + \alpha_2 X_{it} + \delta_i + \lambda_t + \varepsilon_{it} \quad (1)$$

where the subscripts i and t represent the city and the year, respectively. δ_i and λ_t denate the city-fixed effect and year-fixed effect, respectively.[4] AP_{it} refers to the air pollution of city i in year t. $NFCC_{it}$ equals to 1 if the city i has implemented the NFCC policy and 0 otherwise. α_1 of $NFCC_{it}$ is the core coefficient we are most concerned about. X_{it} is a series of control variables for city characteristics and climate characteristics. ε_{it} is the random error term. Considering that the sample in this study is at the city level, we used the city-level clustered standard errors.

As mentioned earlier, the prerequisite for the use of the DID model is that the parallel trend assumption is satisfied. In other words, the air pollution status of the treatment and control groups needs to have the same trend before the NFCC policy implemented. Similar to the previous literature (Fu et al., 2021; Liao et al., 2021; Tanaka, 2015; Tang et al., 2018), we utilized the event study analysis method to test the parallel trends. The model is set as follows.

$$AP_{it} = \beta_0 + \sum \beta_t NFCC_{it} + \beta_2 X_{it} + \delta_i + \lambda_t + \mu_{it} \quad (2)$$

where $NFCC_{it}$ denotes the relative time t of city i from being implemented the NFCC policy. If t equals to 0, it means the year of being selected as a national forest city. Thus, a negative (positive) t indicates before (after) the implementation of the NFCC policy. For example, t equals to -1 is the first year prior to the implementation of the NFCC policy; t equals to 1 involves the first year after being selected as an national forest city. The significance of coefficients β_t indicates whether there is a difference between the treatment group and the control group.

Since the theoretical analysis part of this paper has carried out in-depth analysis of the impact of various mechanism variables on sulfur dioxide and $PM_{2.5}$, we used the classic mechanism analysis method in environmental economics (Acemoglu et al., 2016; Almond et al., 2009; Li & Wang, 2022; Li et al., 2022). Models are constructed as follows.

$$MEC_{it} = \theta_0 + \theta_1 NFCC_{it} + \theta_2 X_{it} + \delta_i + \lambda_t + \sigma_{it} \quad (3)$$

[4] Since city- and year-fixed effects are added in the model, dummy variables measuring the time of NFCC policy and the treatment and control groups do not need to be controlled for again. That is, we can use the reduced form DID models (Duflo, 2001).

$$AP_{it} = \rho_0 + \rho_1 MEC_{it} + \rho_2 X_{it} + \delta_i + \lambda_t + \varphi_{it} \quad (4)$$

where MEC_{it} stands for mechanism variables. In Eqs. (3) and (4), θ_1 and ρ_1 are the core estimated coefficients we are concerned with. If they are significant and the sign is as expected, it indicates that the mechanism is working under are the core estimated coefficients we are concernedwith. If they are significant and the sign is as expected, it indicates that the mechanism is working under the DID framework. It should be noted that because of the excessive missing values for some of mechanism variables, we did not use the mediating effect model. Based on the existing theory and previous literature, analyzing the relationship between the mechanism variables and the explanatory variables can also provide convincing evidence for the mechanism tests (Almond et al., 2009; Chen & Zhao, 2021; Li et al., 2022; Zhao et al., 2022).

Data and variables

The explained variable is air pollution of cities. Haze pollution is one of the types of air pollution that the Chinese government and the public have recently paid close attention to. Consistent with some previous literature (e.g., Cai et al., 2018; Xu et al., 2020; Zhang et al., 2021; Zhao et al., 2021), we chose $PM_{2.5}$ concentrations ($\mu g/m^3$) to measure haze pollution. In addition, SO_2, as another typical air pollutants in cities, has also been the focus of scholars (Guo et al., 2017; He, 2006; Tu et al., 2019). We used SO_2 emissions (100,000 tons) of cities as proxies. The $PM_{2.5}$ data are from the satellite monitoring data by Dalhousie University, and the original data of SO_2 emissions come from the China City Statistical Yearbook.

It is worth mentioning that other pollutants of interest from previous studies, such as CO, PM_{10}, and NOx (e.g., Dong et al., 2019; Ebenstein et al., 2015; Feng et al., 2019; Han et al., 2020; Zhong et al., 2017), were not included in the explained variables. The main reason is that the Ministry of Environmental Protection of China has started to disclose statistical information on these pollutants since 2013 aFor climate-related control variablesnd only some major cities published them before (Ebenstein et al., 2015; Han et al., 2020). Due to the availability of data, we have to choose these two pollutants, $PM_{2.5}$ and SO_2.

As discussed in "Methodology" section, the core explanatory variable is a dummy variable for the NFCC policy. Similar to most policy evaluations based on quasi-natural experiments (Cheng et al., 2019; Duflo, 2001; Fu et al., 2021; Wang & Watanabe, 2019; Zhong et al., 2017), we assigned a city's year after implementation of the NFCC policy to 1, and 0 otherwise. The NFCC data come from the annual selection of the National Forestry and Grassland Administration since 2004.[5]

We further described the distribution of the two pollutants. We first calculated the proportion of the number of cities implementing NFCC in the eastern, central, and western regions to all cities,[6] as well as cities in the south and north.[7] We found that the proportion of cities implementing NFCC in the eastern, central and western regions was 36.44%, 36.11% and 18.63%, respectively; the proportion of cities that implement NFCC is 9.46% and 6.87% in southern and northern cities, respectively. We then observed the joint nuclear density distribution of SO_2 and $PM_{2.5}$ in different regions, as shown in Fig. 2. Clearly, compared with the central and western regions, the SO_2 emissions in the eastern region are larger and more concentrated. In terms of $PM_{2.5}$ concentrations, the western cities are also lower. In addition, compared with southern cities, northern cities have greater heterogeneity in $PM_{2.5}$ and SO_2 emissions. In general, the emission of environmental pollutants in the eastern region is relatively large, so the proportion of cities that implement NFCC may be also large.

The control variables mainly portray city characteristics and climate characteristics. According to related previous studies (e.g., Chen & Xu, 2017; Yang et al., 2021; Yu & Yang, 2021; Zhang et al., 2021), city characteristic variables include industrial structure (*Structure*), economic growth (*Growth*), fixed

[5] Information source: http://www.forestry.gov.cn/zlszz/4249/index.html.

[6] The eastern region includes eight provinces, namely Hebei, Shandong, Liaoning, Jiangsu, Zhejiang, Fujian, Guangdong, and Hainan. There are eight provinces in the central region, and they are Heilongjiang, Jilin, Shanxi, Jiangxi, Anhui, Henan, Hubei, and Hunan. The remaining 11 provinces or autonomous regions belong to the western region, namely Xinjiang, Inner Mongolia, Ningxia, Gansu, Qinghai, Tibet, Yunnan, Guizhou, Sichuan, Guangxi, and Shaanxi.

[7] The south and the north are divided by the Qinling-Huaihe line (Almond et al., 2009; Zhao et al., 2021).

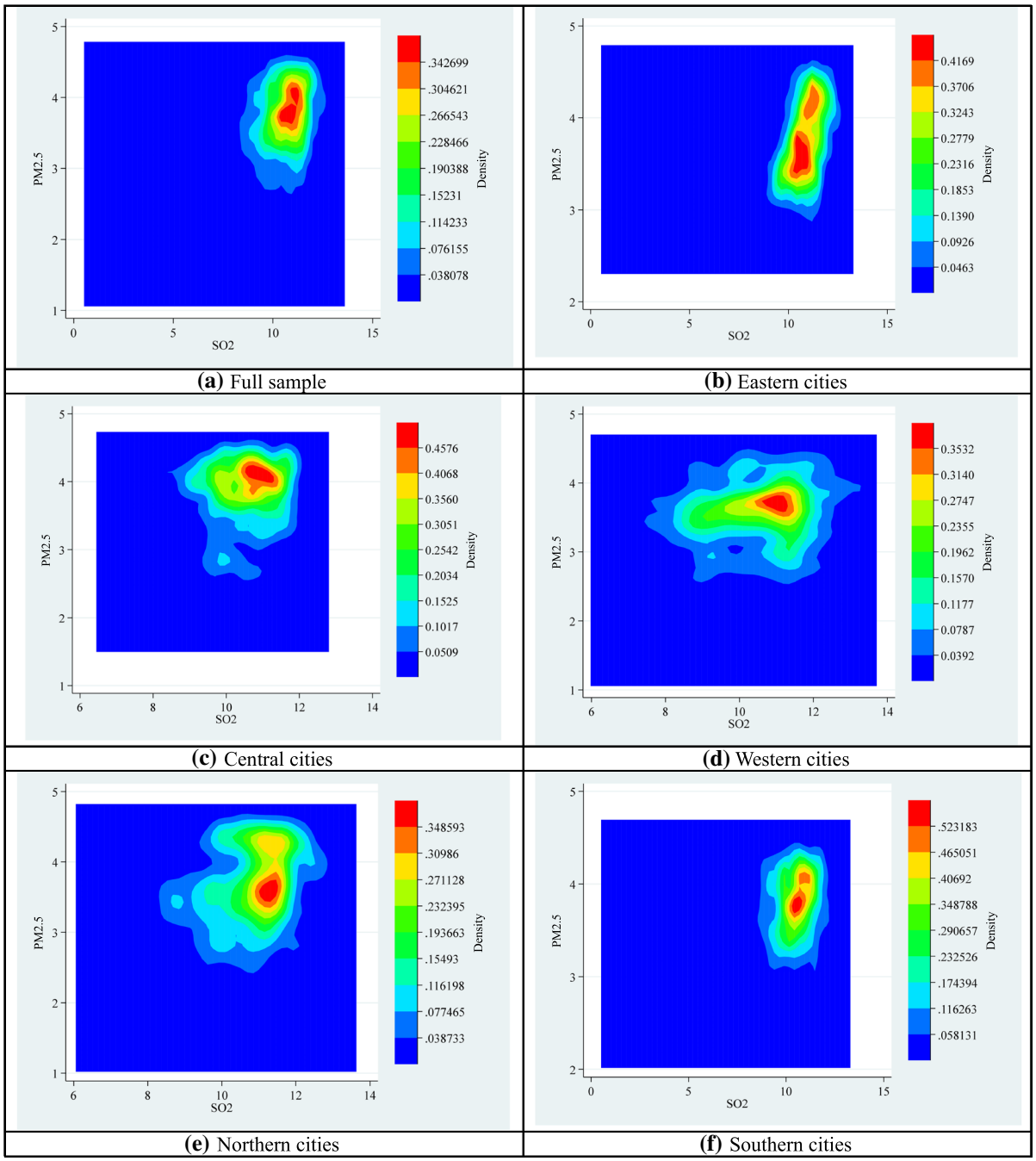


Fig. 2 Joint kernel density function plot of SO_2 and $PM_{2.5}$. *Note*: The kernel density function from a Gaussian product kernel

assets investment (*Investment*), city size (*Size*), energy consumption (*Energy*), industrial agglomeration (*Agglomeration*), and technological innovation (*Innovation*). Among them, *Structure* is measured by the percentage of employees in the secondary industry, and *Growth* refers to the annual growth rate of GDP (%). *Investment* is expressed by total investment in fixed assets (10 billion yuan), and population density (100 persons/km^2) is the proxy for *Size*. *Energy* and *Agglomeration* involve electricity consumption per

Table 1 Descriptive statistics

Variable	N	Mean	SD	Min	Max
$PM_{2.5}$	5180	39.728	20.612	3.5957	94.610
SO_2	4027	0.5741	0.52392	0.0028	3.3224
NFCC	5441	0.0811	0.2729	0	1
Structure	4032	0.4367	0.1416	0.1014	0.8008
Growth	4031	0.1209	0.0401	− 0.0380	0.2580
Investment	4032	9.0740	11.540	0.2825	69.827
Size	4031	4.1991	3.1669	0.1066	22.276
Energy	4024	0.7078	1.1575	0.0101	8.2144
Agglomeration	4031	0.9406	0.3014	0.2182	1.7107
Innovation	4698	7.9338	9.0034	0	11.097
Temperature	4671	14.356	5.5203	0.1000	25.200
Humidity	4671	0.6649	0.1074	0.3400	0.8465
Precipitation	4613	2.5876	1.5112	0.1000	7.4271
Sunshine	4655	5.6036	1.4987	2.2027	8.9414

capita (100 kWh) and the Herfindahl index calculated by the number of employees in the secondary industry, respectively. Lastly, we used the number of patent applications (logarithmic) to measure the technological innovation of cities.

For climate-related control variables, we mainly selected variables such as *Temperature*, *Humidity*, *Precipitation*, and *Sunshine*. Among them, *Temperature* is measured by annual average temperature (°C), *Humidity* refers to *the r*elative humidity (%), *Precipitation* is the average annual rainfall (mm), and Sunshine is expressed by average annual sunshine duration (hour). In terms of control variables, data on patent applications and climate variables come from the State Intellectual Property Office and National Weather Service, respectively, and original data of all other city characteristics are from the China City Statistical Yearbook.

Table 1 shows the descriptive statistics of above variables. The average $PM_{2.5}$ concentrations were 39.73 μg/m^3 and the SO_2 emissions were 5.74 million tons. The standard deviation of $PM_{2.5}$ and SO_2 indicates the large variation of $PM_{2.5}$ and SO_2 emissions in different cities, which provides sufficient heterogeneity for this study to employ the DID model to estimate. In addition, about 8.11% of the cities became NFCCs from 2003 to 2016.

Results

Baseline results

We employed the model (1) to examine the casual effect of the NFCC on air pollution, and the baseline results are shown in Table 2. To test the sensitivity of the estimates to different control variables, we included different kinds of control variables in different columns. Specifically, columns (1) and (4) controlled for city- and year-fixed effects only, columns (2) and (5) further added city characteristic variables, and columns (3) and (6) included all control variables and fixed effects. We found that the coefficients on *NFCC* are significantly negative in all columns, with statistical significance at the 1% level, indicating that the implementation of NFCC policy did mitigate air pollution.

Specifically, in columns (3) and (6), the coefficients on *NFCC* are − 0.1214 and − 0.0429, respectively, suggesting that compared to the control group without NFCC policy, the NFCC reduced $PM_{2.5}$ concentrations by 12.14% and SO_2 emissions by 4.29%. According to the back-of-the-envelope calculation (Deschenes et al., 2020), we found that during the sample period (2003–2016), the total environmental benefits of the NFCC in all cities were a 0.025 g/m^3 reduction in $PM_{2.5}$ concentrations and a 10.08 million tons reduction in SO_2 emissions. It is worth noting that the conclusions of this study are similar to Zhang et al. (2021), which both verify the inhibitory effect of NFCC on $PM_{2.5}$. To judge the economic significance of the estimates, we compared the effect of NFCC with other related policies. We found that the inhibitory effect of NFCC on the emission of environmental pollutants was less than that of the APPCAP, but greater than that of the carbon market pilot policy (Feng et al., 2019; Yu & Zhang, 2021). The possible reason is that the APPCAP are a mandatory policy issued by the Chinese government for areas with serious environmental pollution, and the pollution reduction efforts are quite strong (Huang et al., 2018; Li & Wang, 2022). The core of the NFCC is the increase of forest density, which is a subsidiary impact on pollution reduction (Liao et al., 2021; Zhang et al., 2021). The carbon market has the weakest ability to reduce pollution, which may be related to the imperfection of the market mechanism (Li et al., 2021).

Table 2 Baseline results

	$PM_{2.5}$			SO_2		
	(1)	(2)	(3)	(4)	(5)	(6)
NFCC	− 0.1476***	− 0.1318**	− 0.1214**	− 0.1078***	− 0.0493***	− 0.0429***
	(0.0496)	(0.0527)	(0.0530)	(0.0179)	(0.0176)	(0.0177)
Structure		− 0.0141	− 0.0144		0.0118*	0.0097
		(0.0122)	(0.0117)		(0.0069)	(0.0067)
Growth		0.0232	0.0022		0.0021	0.0055
		(0.0528)	(0.0457)		(0.0256)	(0.0254)
Investment		− 0.0054*	− 0.0047		0.0007	0.0003
		(0.0033)	(0.0032)		(0.0013)	(0.0013)
Size		0.0001**	0.0001*		− 0.0001	− 0.0001
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
Energy		− 0.0007	− 0.0005		− 0.0012**	− 0.0012**
		(0.0005)	(0.0005)		(0.0005)	(0.0005)
Agglomeration		0.6730	0.6576		− 0.3740	− 0.2543
		(0.5676)	(0.5469)		(0.3260)	(0.3180)
Innovation		− 0.0438	− 0.0430		− 0.0179	− 0.0195
		(0.0395)	(0.0388)		(0.0267)	(0.0268)
Temperature			0.0874***			0.0042
			(0.0278)			(0.0104)
Humidity			− 0.0155***			0.0046*
			(0.0037)			(0.0026)
Precipitation			− 0.0013			− 0.0000
			(0.0022)			(0.0003)
Sunshine			− 0.0004***			0.0001
			(0.0001)			(0.0000)
City-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.6065***	3.9616***	4.6757***	0.5356***	0.4559***	− 0.1294
	(0.0226)	(0.2261)	(0.5847)	(0.0131)	(0.1041)	(0.3417)
Adjusted R^2	0.3248	0.3865	0.4035	0.1884	0.2810	0.2894
Observations	5180	3990	3886	4027	4004	3900

The values in parentheses are the standard errors clustered at the city level

*, **, ***Represent the significant levels of 10%, 5%, and 1%, respectively. The following tables are the same

Parallel trend tests

By using the model (2), we tested whether the NFCC policy satisfied the parallel trends. The results are shown in Fig. 3, and we found that the coefficients β_t are all insignificant before being selected as a national forest city for both $PM_{2.5}$ and SO_2. These results indicated that there is no significant difference in the trend of air pollution change between the treatment group and the control group. However, SO_2 did not decrease significantly until the third year, which may be related to the timing of plant growth and technological progress afterward. In addition, these results suggested that the coefficient of two air pollutants decreased after implementing the NFCC policy, implying that NFCC reduced the air pollution in cities. In summary, our results provided supporting evidence that the NFCC policy met the parallel trends and it was valid and appropriate to employ the DID model in this study.

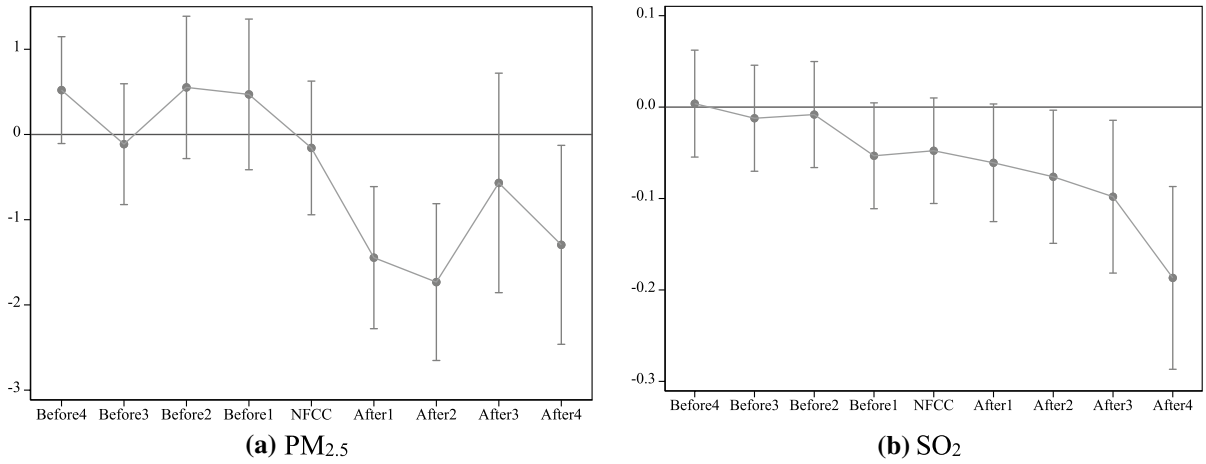


Fig. 3 Parallel trend tests. *Note*: The vertical axis represents the estimated coefficient, while the horizontal axis indicates the relative time node of the NFCC policy implementation

Robustness tests

The baseline results show that the NFCC mitigated air pollution; however, we preformed multiple robustness tests to ensure that these results were convincing.

Instrumental variable approach

Although the DID model is used in the main specifications, the issue of endogeneity deserves further consideration. On the one hand, since the selection of national forest cities was not random, some omitted variables that we cannot observe may affect the estimation results. On the other hand, the baseline estimates may also have the reverse causality problem. A city with severe air pollution will attract the attention of governments, and it is possible to alleviate severe air pollution by planting more trees or even implementing NFCC policies implementing NFCC policy. Therefore, we further employed an instrumental variable (IV) approach for robustness tests.

Specifically, we selected the average slope of each city as the IV (Qian, 2008). From the perspective of relevance, whether a city can become an *NFCC* has a lot to do with the amount and density of local forests, and the density of forests is closely related to the slope. For example, broad-leaved forest often appears on slightly sloped terrain, the distribution of young forests on different slopes is basically the same, and the proportion of over-mature forests increases with the increase of slope (Keim & Skaugset, 2003; Tsui et al., 2004). Some literature in forest science also confirmed that slope is closely related to tree survival and forest formation (Sharma et al., 2010; Stage & Salas, 2007). In terms of the exclusion restriction assumption, the slope of the city is formed by the movement of the earth's crust and unlikely to directly affect air pollution. In addition, the emission of air pollutants mainly depends on economic factors, such as economic development, industrialization, technological innovation, and mineral resources (Cai et al., 2018; Huang et al., 2018; Ma et al., 2021; Sheehan et al., 2014; Yuan & Zuo, 2011; Zhu et al., 2019). There is no direct evidence that slope is associated with air pollution. The slope data of Chinese cities are from You et al. (2018).

Table 3 reports the estimated results by using the IV approach. Although the coefficients on *NFCC* are insignificant in columns (1) and (4) of Panel B, estimated coefficients on *NFCC* are significantly negative at the 1% level in columns (3) and (6), indicating the robustness of our baseline results. Moreover, in Panel A, it is clear that the first-stage F values are well above the Stock-Yogo critical value for a weak IV (Stock & Yogo, 2005; Zhao et al., 2022) and other statistics (e.g., Kleibergen-Paap rk LM statistic) also suggest that this IV is strong.

Placebo tests

Similar to previous studies using the DID model (Cheng et al., 2019; Fu et al., 2021; Wang &

Table 3 Robustness checks: IV estimations

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: First-stage results	NFCC					
Slope	− 0.0125***	− 0.0029***	− 0.0472**	− 0.0133*	− 0.0027***	− 0.0455***
	(0.0001)	(0.0011)	(0.0203)	(0.0077)	(0.0008)	(0.0176)
City characteristics	No	Yes	Yes	No	Yes	Yes
Climate variable	No	No	Yes	No	No	Yes
City-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
F value of the first stage	16.38	26.69	15.22	13.73	27.82	30.84
Observations	4538	3990	3886	4013	4004	3900
Panel B: Second-stage results	$PM_{2.5}$			SO_2		
NFCC	− 1.6519	− 0.0814***	− 0.2369***	− 0.0072	− 0.0139***	− 0.0167***
	(1.8491)	(0.0261)	(0.0627)	(0.0075)	(0.0057)	(0.0062)
City characteristics	No	Yes	Yes	No	Yes	Yes
Climate variable	No	No	Yes	No	No	Yes
City-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen–Paap rk LM statistic	1.845	33.850	19.852	17.426	34.132	20.031
Cragg–Donald Wald F statistic	1.348	26.761	15.223	14.841	27.029	15.397
Anderson–Rubin Wald test	759.46	35.26	160.27	0.98	10.73	13.33
Observations	4538	3990	3886	4013	4004	3900

The values in parentheses are the standard errors clustered at the city level. Panel A reports the first-stage results, where the explanatory variable is NFCC. Panel B presents the second-stage results, of which the explanatory variable in the first three columns is $PM_{2.5}$ and the explanatory variable in the last three columns is SO_2

Table 4 Robustness checks: placebo tests

	1 year in advance (1)	2 years in advance (2)	3 years in advance (3)
Panel A: $PM_{2.5}$			
NFCC	0.2188	0.2707	− 0.5011
	(0.6531)	(0.6629)	(0.5378)
Baseline controls	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes
Observations	3886	3886	3886
Panel B: SO_2			
NFCC	− 0.3246	0.0743	− 0.0455
	(0.6312)	(0.4558)	(0.5019)
Baseline controls	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes
Observations	3900	3900	3900

The values in parentheses are the standard errors clustered at the city level

Table 5 Robustness checks: eliminating disturbing policies

	APPCAP	Low-carbon cities	Carbon market pilots	Smart cities	Driving restriction policies	NLGC	"Two control Zones"
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: $PM_{2.5}$							
NFCC	− 0.9300**	− 0.1811***	− 0.1347***	− 0.1039**	− 0.1003**	− 0.0924**	− 0.0653***
	(0.3757)	(0.0527)	(0.0411)	(0.0420)	(0.0479)	(0.0425)	(0.0227)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3526	2560	3368	2271	2566	2016	1327
Panel B: SO_2							
NFCC	− 0.0353**	− 0.0509***	− 0.0489***	− 0.0322*	− 0.0402**	− 0.0512**	− 0.0377***
	(0.0172)	(0.0192)	(0.0199)	(0.0168)	(0.0185)	(0.0221)	(0.0106)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3540	2574	3382	2285	2580	2030	1341

The values in parentheses are the standard errors clustered at the city level

Watanabe, 2019), we further adopted some placebo tests. Specifically, we advanced the timing of NFCC policy implementation (treatment group) by one, two, and three years, respectively. Judging from the results reported in Table 4, the inhibitory effect of the NFCC on air pollution is no longer significant, indicating that the simulated treatment effect we constructed did not hold, and the main results have great credibility.

Eliminating disturbing policies

In further robustness checks, we considered some other disruptive policies that may have an impact on air pollution of cities. First, the APPCAP is the most important policy implemented by the Chinese government to control air pollution (Feng et al., 2019; Huang et al., 2018), especially in "2+26" cities in the North China Plain, which were implemented in the first batch. If this environmental policy in cities with severe air pollution was not ruled out, it should lead to an overestimation. In column (1) of Table 5, we excluded the "2+26" cities[8] and found that the coefficients on *NFCC* remained significantly negative in both two panels.

Second, as a result of China's 2030 carbon peak and 2060 carbon neutrality targets, China's low-carbon city pilots and carbon market pilots have achieved carbon emission reductions while curbing air pollutant emissions (Fu et al., 2021; Yu & Zhang, 2021). Considering that most of the CO_2 and air pollutants are emitted from similar sources (Cheng et al., 2020; Lei et al., 2011), we eliminated these two city-level low-carbon policies. The results in columns (2) and (3) of Table 5 suggest that after controlling the effects of these disturbing policies, the NFCC still reduced air pollution.

Third, in recent years, the use of big data and artificial intelligence for monitoring and controlling pollution has become increasingly widespread (Honarvar & Sami, 2019), and the positive impact of smart cities with these information technologies at their core on air pollution has received attention from scholars (Liu et al., 2021; Zhu et al., 2019). China has been setting up smart city pilots since 2012, and there were more than 200 smart cities at all levels (including counties and towns) set up by 2016. We therefore excluded smart city pilot policy, and the results are presented in column (4) of Table 5. The estimates in two panels are qualitatively similar to the main estimates in Table 2, although somewhat smaller in magnitude.

[8] Data source: http://www.gov.cn/xinwen/2019-08/23/content_5423736.htm.

Table 6 Robustness checks: using alternative specifications

	Deleting NFCC before 2007 and after 2017	Per capita indicators	PSM-DID (NNM)	PSM-DID (KM)
	(1)	(2)	(3)	(4)
Panel A: $PM_{2.5}$				
NFCC	− 0.1354***	− 0.0239**	− 0.1356***	− 0.1321***
	(0.0397)	(0.0108)	(0.0392)	(0.0438)
Baseline controls	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes
Observations	3595	3886	3568	3580
Panel B: SO_2				
NFCC	− 0.0457***	− 0.0175***	− 0.0424***	− 0.0446***
	(0.0166)	(0.0059)	(0.0182)	(0.0177)
Baseline controls	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes
Observations	3595	3900	3582	3594

The values in parentheses are the standard errors clustered at the city level

Fourth, to alleviate traffic congestion and air pollution problems, many cities in China have implemented driving restriction policies. Previous literature provided evidence that driving restrictions did help mitigate air pollution in China (Han et al., 2020; Zhong et al., 2017). Thus, we tried to exclude this driving restriction policies to conduct a robustness test. In column (5) in Table 5, we found that the results are virtually unchanged.

Fifth, we considered another program similar to the NFCC, National Landscape Garden City (NLGC). Since the Chinese Ministry of Housing and Urban-Rural Development declared Beijing the first NLGC in 1992, over 200 cities have become NLGCs. Clearly, NLGCs are also centered on urban ecology and may have an impact on urban air pollution. Therefore, in column (6), we removed the sample of NLGCs, fingding that the positive effect of NFCC on mitigating air pollution remains significant while considering the effect of NLGC.

Last, to reduce the formation of acid rain, the Law on the Prevention and Control of Atmospheric Pollution in China stipulates that, according to natural conditions such as weather, topography, soil, etc., areas that have produced or may produce acid rain or other areas with serious sulfur dioxide pollution may be designated as acid rain control areas or sulfur dioxide pollution, namely "Two control Zones" (Cai et al., 2016; Hao et al., 2000). The establishment of the "Two control Zones" can also significantly reduce SO_2 emissions, thus overestimating the effect of NFCC. After excluding the influence of "Two control Zones" in column (7) of Table 5, we found that NFCC can still significantly reduce SO_2 and $PM_{2.5}$.

Using alternative specifications

To exclude the bias of variable selection and model setting on the results, we tested the robustness by using alternative specifications. First, the specific selection criteria for NFCC were issued in 2007, so we excluded national forest cities before 2007. Moreover, we removed cities that were selected as national forest cities after 2017, as these candidates are different from the other cities in the control group. The results in column (1) of Table 6 indicate that this change has little effect on the results.

Second, considering the size differences, we utilized the per capita indicators as proxies for the explanatory variables. In column (2) of Table 6, the estimates of NFCC are quite insensitive to the variable measurement method adopted, suggesting that the main specification is satisfactory for capturing size differences.

Table 7 Mechanism tests: the NFCC and green spaces

	Green coverage area per capita in built-up areas		Green spaces per capita	
	(1)	(2)	(3)	(4)
NFCC	0.2965***	0.3404***	0.1473***	0.0355***
	(0.1092)	(0.1189)	(0.0156)	(0.0149)
Baseline controls	No	Yes	No	Yes
Year-fixed effect	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes
Observations	3999	3875	4000	3877

The values in parentheses are the standard errors clustered at the city level

Third, similar to some previous studies (Fu et al., 2021; Xie et al., 2017), we employed the DID approach on the basis of the propensity score matching (PSM) method. We used two matching methods, nearest-neighbor matching (NNM) and kernel matching (KM), and the estimation results from PSM-DID approach are reported in columns (3) and (4) of Table 6, respectively. Clearly, in two panels, the estimates are similar to the main results in Table 2.

Mechanism tests

Given the main finding that NFCC significantly curbed air pollution, we further examined the multiple potential mechanisms highlighted in "Mechanism analysis" section. These mechanisms included green spaces, environmental regulations, and green development models.

Green spaces

As a new urban ecosystem centered on forest vegetation and greening, the improvement of forest coverage and urban green spaces is the most significant feature of forest cities. Two variables are used to measure the green spaces in Chinese cities, green coverage area per capita in built-up areas and green spaces per capita, whose raw data come from the China City Statistical Yearbook. The results are reported in Table 7.

The DID estimates indicate that the coefficients of *NFCC* are significantly positive in all columns regardless of the inclusion of baseline control variables. These results are similar to findings in some previous studies (Xu et al., 2020; Zhang et al., 2021), suggesting that the NFCC increased green spaces in Chinese cities. Specifically, the NFCC increased green coverage area per capita in built-up areas by 34.04% and green spaces per capita by 3.55%. Considering that the positive role of urban green spaces in enhancing air quality has been widely confirmed in previous studies (e.g., Escobedo et al., 2011; Lei et al., 2021; Nowak et al., 2006), we found that urban green space is indeed negatively correlated with air pollutants in "Table 11 of Appendix". Thus, our estimates provided evidence that the NFCC could mitigate air pollution by increasing green spaces in China.

Environmental regulations

Next, we considered whether the NFCC affected environmental regulations. Two indicators, industrial soot removal rate and comprehensive solid waste utilization rate, were constructed to serve as proxies for environmental regulations. Previous studies on the measurement of environmental regulations were not uniform. One category used environmental policies and laws to analyze the role of environmental regulations (Cai et al., 2018; Tanaka, 2015; Wan et al., 2021; Xie et al., 2017), while the other focused on the measurement domain in terms of pollutant utilization and removal (Du & Li, 2020; Zhao et al., 2021). Since the NFCC has been considered as a

Table 8 Mechanism tests: the NFCC and environmental regulations

	Soot removal rate		Comprehensive solid waste utilization rate	
	(1)	(2)	(3)	(4)
NFCC	0.1391	0.3012	0.5538**	0.4870**
	(0.4627)	(0.3552)	(0.2251)	(0.2357)
Baseline controls	No	Yes	No	Yes
Year-fixed effect	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes
Observations	1689	1652	4008	3887

The values in parentheses are the standard errors clustered at the city level

Table 9 Mechanism tests: the NFCC and green development models

	Environmental employment	Green transportation	Green Inventions	Innovation index
	(1)	(2)	(3)	(4)
NFCC	0.0712**	0.0001	0.1190***	0.1370***
	(0.0337)	(0.0002)	(0.0466)	(0.0271)
Baseline controls	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes
Observations	3869	3834	2974	3895

The values in parentheses are the standard errors clustered at the city level

quasi-natural experiment, and it is difficult for us to construct environmental regulation variables through other policies or laws, the latter option is more suitable for mechanism testing in this study. The original data for these two indicators are also obtained from the Chinese City Statistical Yearbook.

The results of the mechanism analysis are reported in Table 8. We found that the coefficients on *NFCC* are positive but insignificant in the first two columns, which may be influenced by too many missing observations. Compared to the baseline results in Table 2, the number of observations is less than half. Additionally, the coefficients on *NFCC* are positive and significant at the 5% statistical level in columns (3) and (4), indicating that there is a positive association between the NFCC and the enhancement of environmental regulations. As highlighted in previous research that environmental regulations are one of core instruments for controlling air pollution in China (e.g., Cai et al., 2018; Du & Li, 2020; Tanaka, 2015; Xie et al., 2017; Zheng et al., 2015), the results in "Table 12 of Appendix" also show that environmental regulation variables are both negatively associated with air pollution in our sample. In sum, these results are suggestive of support for the environmental regulation-related mechanisms.

Green development models

With the deepening of the forest city concept, the NFCC gradually formed green development models to mitigate air pollution under the guidance of the government. We examined this mechanism from several perspectives, and the results are represented in Table 9.

In column (1), we adopted the number of employment related to environmental protection to measure city's green industry (Raff & Earnhart, 2019) and found that the coefficient on *NFCC* is statistically significant at the 5% level. These results imply that the NFCC is likely to promote the development of green industries. In column (2), we explored the government's guidance of green living for residents as characterized by per capita public transportation ownership (Mulalic & Rouwendal, 2020). Similarly, we found that the NFCC can significantly contribute to the development of green transportation in cities.

Moreover, given the positive role of technological innovation and green innovation in enhancing environmental quality (Popp, 2006; Zhao et al., 2021; Zhu et al., 2019), we further tested whether the NFCC inhibited air pollution through these mechanisms. Green innovation is expressed as the logarithm of green patent applications and technological innovation is measured by the city innovation index.[9] In columns (3) and (4), we found that the NFCC significantly contributed to the city's green innovation and technological innovation capacity, which are in line with our previous theoretical prediction.

In line with previous analysis in this paper, based on Eq. (4), we found that all four mechanism variables involving the green development model are significantly and negatively associated with air pollution in "Table 13 in Appendix". In summary, our findings confirmed that the NFCC can lead to green development models to reduce air pollution in Chinese cities.

[9] The original data of city innovation index comes from Kou and Liu (2017) and the data related to green patents are compiled in accordance with the "Green List of International Patent Classification" issued by the World Intellectual Property Office.

Table 10 Heterogeneity effects by cities

	Smaller cities (1)	Larger Cities (2)	Northern cities (3)	Southern cities (4)	Eastern cities (5)	Central cities (6)	Western cities (7)
Panel A: $PM_{2.5}$							
NFCC	− 0.9252**	− 0.1982***	− 0.8737	− 0.0698*	− 0.3618	0.3145	− 0.1562**
	(0.4644)	(0.0691)	(0.6244)	(0.0405)	(0.6256)	(0.8578)	(0.0649)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2971	915	1765	2121	1179	1085	1622
Panel B: SO_2							
NFCC	− 0.0614***	− 0.0423	− 0.0396	− 0.0676***	− 0.0412	− 0.0268	− 0.0664**
	(0.0195)	(3.6298)	(0.0684)	(0.0202)	(0.0411)	(0.0356)	(0.0322)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2985	915	1779	2121	1193	1085	1622

The values in parentheses are the standard errors clustered at the city level. Baseline control variables and fixed effects are added in all regressions

Heterogeneity analysis

The role of NFCC in mitigating air pollution may be heterogeneous. First, according to the State Council's criteria for classifying the size of cities,[10] we set cities with a resident population of more than 1 million as large cities, and the others as small cities. As reported in columns (1) and (2) of Table 10, the suppression effects of NFCC on $PM_{2.5}$ are both significant in different city sizes, while stronger in smaller city subsamples. Also, the NFCC is more beneficial to alleviate SO_2 emissions in smaller cities. To sum up, the NFCC is more effective in smaller cities and possible reason is that smaller cities could have more land for greening and forestry and better utilize the environmental benefits of NFCC.

Second, as mentioned above, southern and northern cities are not evaluated on the same criteria on the NFCC. For instance, the forest coverage rate of southern cities needs to be at least 40%, while northern cities only need to reach 30% or more. In addition, according to Figure 2, there are also some differences in air pollution between northern and southern cities in China. In columns (3) and (4) of Table 10, we compared and analyzed this difference and found that for both $PM_{2.5}$ and SO_2, the NFCC had greater inhibitory effects on air pollution in southern cities. However, the coefficients on NFCC are insignificant in northern cities.

Third, we analyzed the heterogeneity effects in the eastern, central, and western regions. Figure 2 tells us that SO_2 emissions and $PM_{2.5}$ concentrations in the different regions also show greater heterogeneity. In the last three columns of Table 10, the coefficients on *NFCC* are only significantly positive in column (7) for two panels, indicating that the NFCC had more effective in reducing air pollution in the western regions. In general, the western regions of China are dominated by plateaus and mountains, which are more suitable for the development of urban forest ecosystems. For example, Guiyang is the first national forest city in China. In addition, compared with central and eastern cities, the western cities are less industrialized and has relatively better air quality (Zhao et al., 2021; Zhou et al., 2019). The role of

[10] Data source: http://www.gov.cn/zhengce/content/2014-11/20/content_9225.htm.

NFCC with ecological means as the core is valid for reducing air pollution.

Conclusions

Air pollution has become a major challenge for most emerging market countries, and China has made great efforts to control air pollution. The NFCC, with the goal of ecological construction, is a key step for the Chinese government to create livable environment and sustainable cities. However, few studies have completely evaluated the effects of NFCC on urban environmental quality (Liao et al., 2021; Xu et al., 2020; Zhang et al., 2021), and the underlying mechanisms are poorly understood. By using data from 283 Chinese prefecture-level cities from 2003 to 2016 and regarding the NFCC as a quasi-natural experiment, we employed the DID model to examine the effects of NFCC on air pollution.

We found that the NFCC significantly reduced air pollution, which is supported by a series of robustness checks, including the IV approach, placebo tests, eliminating disturbing policies, and using alternative specifications. Specifically, the implementation of the NFCC policy resulted in an average reduction of 12.14% and 4.29% in $PM_{2.5}$ and SO_2, respectively. For all cities, these effects were equivalent to reducing the total $PM_{2.5}$ concentrations by 0.025 g/m^3 and the total SO_2 emissions by 10.08 million tons over the sample period (2003–2016). Moreover, we analyzed the potential mechanisms and found that the environmental benefits could be explained by increasing green spaces, strengthening environmental regulations, and forming green development models. Furthermore, the results from heterogeneity effects indicated that the NFCC was more effective in mitigating air pollution in smaller cities, southern cities, and western cities.

Our findings contributed to the literature on assessing environmental policies at the city level with a quasi-natural experiment of the NFCC. We found that the NFCC focused on ecological goal, helped to improve air quality as well, which distinguished it from previous policies on administrative and economic methods in China (e.g., Cai et al., 2018; Chen et al., 2007; Cheng et al., 2019; Fu et al., 2021; Huang et al., 2018; Ma et al., 2021; Tanaka, 2015; Zhong et al., 2017). Thus, we offered a new method of urban air pollution control from the perspective of the NFCC. Unlike traditional administrative-style policy measures that might have negative impacts (Lo, 2014; Tang et al., 2018), the use of softer ecological tools could also help achieve improvements in air quality.

More importantly, the NFCC provided directions for the formation of green development patterns in cities, not only in air pollution reductions, but also for the possible stable and long-term effects on promoting sustainable urban development. Our study had some insights not only for China but also for other high pollution countries, especially emerging market countries and developing countries. Specifically, some cities could utilize the NFCC opportunity to promote a green development model and reduce the cost of air pollution control. In addition, the NFCC played a lesser role in large cities, northern cities, and eastern. Therefore, local governments in these cities need to pay attention to the environmental benefits brought by NFCC, provide good public services, scientifically coordinate and plan green spaces and construction lands, and restore and build new urban forest elements according to local conditions.

Author contributions XL done conceptualization, supervision, formal analysis, supervision, and writing—original draft, writing—review and editing draft; CZ performed formal analysis, methodology, visualization, writing—original draft, and writing—review and editing draft.

Funding This work was supported by the Fundamental Research Funds for the Central Universities (No. QCDC-2020-21).

Data availability All data included in this study are available upon request by contact with the corresponding author.

Declarations

Conflict of interest The author declared that there is no conflict of interest.

Consent for publication All the authors consent to participate and consent to publish before reference in the manuscript.

Appendix

See Tables 11, 12 and 13.

Table 11 Green spaces and SO_2/$PM_{2.5}$

	SO_2		$PM_{2.5}$	
	(1)	(2)	(3)	(4)
Green coverage area per capita in built-up areas	− 0.4621***		− 0.3626***	
	(0.0595)		(0.0523)	
Green spaces per capita		− 0.1436***		− 0.1061***
		(0.0148)		(0.0153)
Baseline controls	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes
Observations	3875	3877	3875	3877

The values in parentheses are the standard errors clustered at the city level

Table 12 Environmental regulations and SO_2/$PM_{2.5}$

	SO_2		$PM_{2.5}$	
	(1)	(2)	(3)	(4)
Soot removal rate	− 0.5371***		− 0.8009***	
	(0.0532)		(0.0608)	
Comprehensive solid waste utilization rate		− 0.0701***		− 0.0676***
		(0.0159)		(0.0190)
Baseline controls	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes
Observations	1652	3887	1652	3887

The values in parentheses are the standard errors clustered at the city level

Table 13 Green development models and SO_2/$PM_{2.5}$

	SO_2				$PM_{2.5}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Environmental employment	− 0.9084***				− 0.1297***			
	(0.0810)				(0.0241)			
Green transportation		− 0.0750***				− 0.5703***		
		(0.0205)				(0.0718)		
Green Inventions			− 0.1540***				− 0.1734***	
			(0.0219)				(0.0234)	
Innovation index				− 0.1485***				− 0.1733***
				(0.0247)				(0.0269)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City-fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3869	3834	2974	3895	3869	3834	2974	3895

The values in parentheses are the standard errors clustered at the city level

References

Acemoglu, D., Moscona, J., & Robinson, J. A. (2016). State capacity and American technology: Evidence from the Nineteenth Century. *American Economic Review, 106*, 61–67.

Almond, D., Chen, Y., Greenstone, M., & Li, H. (2009). Winter heating or clean air? Unintended impacts of China's Huai River Policy. *American Economic Review, 99*, 184–190.

Azizullah, A., Khattak, M., Richter, P., & Hader, D. P. (2011). Water pollution in Pakistan and its impact on public health: A review. *Environment International, 37*(2), 479–497.

Brunekreef, B., & Holgate, S. T. (2002). Air pollution and health. *The Lancet, 360*(9341), 1233–1242.

Cai, S., Ma, Q., Wang, S., Zhao, B., Brauer, M., Cohen, A., & Burnett, R. T. (2018). Impact of air pollution control policies on future $PM_{2.5}$ concentrations and their source contributions in China. *Journal of Environmental Management, 227*, 124–133.

Cai, X., Lu, Y., Wu, M., & Yu, L. (2016). Does environmental regulation drive away inbound foreign direct investment? Evidence from a quasi-natural experiment in China. *Journal of Development Economics, 123*, 73–85.

Chen, B., & Zhao, C. (2021). Poverty reduction in rural China: Does the digital finance matter? *PLoS ONE, 16*(12), e0261214.

Chen, C., Chen, B., Wang, B., Huang, C., Zhao, J., Dai, Y., & Kan, H. (2007). Low-carbon energy policy and ambient air pollution in Shanghai, China: A health-based economic assessment. *Science of the Total Environment, 373*(1), 13–21.

Chen, N., & Xu, L. (2017). Relationship between air quality and economic development in the provincial capital cities of China. *Environmental Science and Pollution Research, 24*(3), 2928–2935.

Cheng, J., Yi, J., Dai, S., & Xiong, Y. (2019). Can low-carbon city construction facilitate green growth? Evidence from China's pilot low-carbon city initiative. *Journal of Cleaner Production, 231*, 1158–1170.

Cheng, K., Hao, W., Wang, Y., Yi, P., Zhang, J., & Ji, W. (2020). Understanding the emission pattern and source contribution of hazardous air pollutants from open burning of municipal solid waste in China. *Environmental Pollution, 263*, 114417.

Churkina, G., Grote, R., Butler, T. M., & Lawrence, M. (2015). Natural selection? Picking the right trees for urban greening. *Environmental Science & Policy, 47*, 12–17.

Cole, M. A., Elliott, R. J., & Shimamoto, K. (2005). Industrial characteristics, environmental regulations and air pollution: An analysis of the UK manufacturing sector. *Journal of Environmental Economics and Management, 50*(1), 121–143.

Deschenes, O., Wang, H., Wang, S., & Zhang, P. (2020). The effect of air pollution on body weight and obesity: Evidence from China. *Journal of Development Economics, 145*, 102461.

Dong, K., Hochman, G., Kong, X., Sun, R., & Wang, Z. (2019). Spatial econometric analysis of China's PM_{10} pollution and its influential factors: Evidence from the provincial level. *Ecological Indicators, 96*, 317–328.

Du, W., & Li, M. (2020). Assessing the impact of environmental regulation on pollution abatement and collaborative emissions reduction: Micro-evidence from Chinese industrial enterprises. *Environmental Impact Assessment Review, 82*, 106382.

Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review, 91*(4), 795–813.

Ebenstein, A., Fan, M., Greenstone, M., He, G., Yin, P., & Zhou, M. (2015). Growth, pollution, and life expectancy: China from 1991–2012. *American Economic Review, 105*(5), 226–231.

Escobedo, F. J., Kroeger, T., & Wagner, J. E. (2011). Urban forests and pollution mitigation: Analyzing ecosystem services and disservices. *Environmental Pollution, 159*(8–9), 2078–2087.

Fan, P., Xu, L., Yue, W., & Chen, J. (2017). Accessibility of public urban green space in an urban periphery: The case of Shanghai. *Landscape and Urban Planning, 165*, 177–192.

Fang, C., Ma, H., Wang, Z., & Li, G. (2014). The sustainable development of innovative cities in China: Comprehensive assessment and future configuration. *Journal of Geographical Sciences, 24*(6), 1095–1114.

Feng, Y., Ning, M., Lei, Y., Sun, Y., Liu, W., & Wang, J. (2019). Defending blue sky in China: Effectiveness of the "Air Pollution Prevention and Control Action Plan" on air quality improvements from 2013 to 2017. *Journal of Environmental Management, 252*, 109603.

Fu, Y., He, C., & Luo, L. (2021). Does the low-carbon city policy make a difference? Empirical evidence of the pilot scheme in China with DEA and PSM-DID. *Ecological Indicators, 122*, 107238.

Guo, Y., Zeng, Z., Tian, J., Xu, F., Chen, L., & Zhou, A. (2017). Uncovering the strategies of green development in a Chinese province driven by reallocating the emission caps of multiple pollutants among industries. *Science of the Total Environment, 607*, 1487–1496.

Han, J., Tang, B., & Hou, S. (2019). Spatial pattern characteristics and influencing factors of national forest cities in China. *Journal of Landscape Research, 11*(5), 35–40.

Han, Q., Liu, Y., & Lu, Z. (2020). Temporary driving restrictions, air pollution, and contemporaneous health: Evidence from China. *Regional Science and Urban Economics, 84*, 103572.

Hao, J., Wang, S., Liu, B., & He, K. (2000). Designation of acid rain and SO_2 control zones and control policies in China. *Journal of Environmental Science and Health, 35*, 1901–1914.

Hart, S. L., & Ahuja, G. (1996). Does it pay to be green? An empirical examination of the relationship between emission reduction and firm performance. *Business Strategy and the Environment, 5*(1), 30–37.

He, J. (2006). Pollution haven hypothesis and environmental impacts of foreign direct investment: The case of industrial emission of sulfur dioxide (SO_2) in Chinese provinces. *Ecological Economics, 60*(1), 228–245.

Honarvar, A. R., & Sami, A. (2019). Towards sustainable smart city by particulate matter prediction using urban big data, excluding expensive air pollution infrastructures. *Big Data Research, 17*, 56–65.

Huang, J., Pan, X., Guo, X., & Li, G. (2018). Health impact of China's Air Pollution Prevention and Control Action Plan: An analysis of national air quality monitoring and mortality data. *The Lancet Planetary Health, 2*(7), e313–e323.

Johnson, T. (2020). Public participation in China's EIA process and the regulation of environmental disputes. *Environmental Impact Assessment Review, 81*, 106359.

Keim, R. F., & Skaugset, A. E. (2003). Modelling effects of forest canopies on slope stability. *Hydrological Processes, 17*, 1457–1467.

Kou, Z., & Liu, X. (2017). *FIND report on city and industrial innovation in China (2017)*. Fudan Institute of Industrial Development, School of Economics, Fudan University.

Lei, Y., Davies, G. M., Jin, H., Tian, G., & Kim, G. (2021). Scale-dependent effects of urban greenspace on particulate matter air pollution. *Urban Forestry & Urban Greening, 61*, 127089.

Lei, Y., Zhang, Q., Nielsen, C., & He, K. (2011). An inventory of primary air pollutants and CO_2 emissions from cement production in China, 1990–2020. *Atmospheric Environment, 45*(1), 147–154.

Li, F., Wang, R., Liu, X., & Zhang, X. (2005). Urban forest in China: Development patterns, influencing factors and research prospects. *International Journal of Sustainable Development & World Ecology, 12*(2), 197–204.

Li, L., Sun, J., Jiang, J., & Wang, J. (2021). The effect of environmental regulation competition on haze pollution: Evidence from China's province-level data. *Environmental Geochemistry and Health*. https://doi.org/10.1007/s10653-021-00854-w

Li, X., Hu, Z., Cao, J., & Xu, X. (2022). The impact of environmental accountability on air pollution: A public attention perspective. *Energy Policy, 161*, 112733.

Li, X., & Wang, D. (2022). Does transfer payments promote low-carbon development of resource-exhausted cities in China? *Earth's Future, 10*, e2021EF002339.

Liao, L., Zhao, C., Li, X., & Qin, J. (2021). Towards low carbon development: The role of forest city constructions in China. *Ecological Indicators, 131*, 108199.

Lichter, A., Pestel, N., & Sommer, E. (2017). Productivity effects of air pollution: Evidence from professional soccer. *Labour Economics, 48*, 54–66.

Liu, L., Guan, D., Peart, M. R., Wang, G., Zhang, H., & Li, Z. (2013). The dust retention capacities of urban vegetation: A case study of Guangzhou, South China. *Environmental Science and Pollution Research, 20*(9), 6601–6610.

Liu, W., Xu, Y., Fan, D., Li, Y., Shao, X. F., & Zheng, J. (2021). Alleviating corporate environmental pollution threats toward public health and safety: The role of smart city and artificial intelligence. *Safety Science, 143*, 105433.

Lo, K. (2014). China's low-carbon city initiatives: The implementation gap and the limits of the target responsibility system. *Habitat International, 42*, 236–244.

McDonald, A. G., Bealey, W. J., Fowler, D., Dragosits, U., Skiba, U., Smith, R. I., & Nemitz, E. (2007). Quantifying the effect of urban tree planting on concentrations and depositions of PM_{10} in two UK conurbations. *Atmospheric Environment, 41*(38), 8455–8467.

Ma, H., Di, D., Li, L., Zhang, W., & Wang, J. (2021). Environmental decentralization, environmental public service, and public health: Evidence from 289 cities in China. *Environmental Geochemistry and Health*. https://doi.org/10.1007/s10653-021-01112-9

Martens, S. (2006). Public participation with Chinese characteristics: Citizen consumers in China's environmental management. *Environmental Politics, 15*(2), 211–230.

Margaritis, E., & Kang, J. (2017). Relationship between green space-related morphology and noise pollution. *Ecological Indicators, 72*, 921–933.

Mulalic, I., & Rouwendal, J. (2020). Does improving public transport decrease car ownership? Evidence from a residential sorting model for the Copenhagen metropolitan area. *Regional Science and Urban Economics, 83*, 103543.

Neidell, M. J. (2004). Air pollution, health, and socio-economic status: The effect of outdoor air quality on childhood asthma. *Journal of Health Economics, 23*(6), 1209–1236.

Nowak, D. J., Crane, D. E., & Stevens, J. C. (2006). Air pollution removal by urban trees and shrubs in the United States. *Urban Forestry & Urban Greening, 4*(3–4), 115–123.

O'rourke, D., & Macey, G. P. (2003). Community environmental policing: Assessing new strategies of public participation in environmental regulation. *Journal of Policy Analysis and Management, 22*(3), 383–414.

Ostro, B., Malig, B., Broadwin, R., Basu, R., Gold, E. B., Bromberger, J. T., & Green, R. (2014). Chronic $PM_{2.5}$ exposure and inflammation: Determining sensitive subgroups in mid-life women. *Environmental Research, 132*, 168–175.

Pérez-Campuzano, E., Ávila-Foucat, V. S., & Perevochtchikova, M. (2016). Environmental policies in the Peri-urban area of Mexico City: The perceived effects of three environmental programs. *Cities, 50*, 129–136.

Popp, D. (2006). International innovation and diffusion of air pollution control technologies: The effects of NOX and SO2 regulation in the US, Japan, and Germany. *Journal of Environmental Economics and Management, 51*(1), 46–71.

Qian, N. (2008). Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance. *Quarterly Journal of Economics, 123*(3), 1251–1285.

Raff, Z., & Earnhart, D. (2019). The effects of Clean Water Act enforcement on environmental employment. *Resource and Energy Economics, 57*, 1–17.

Ruiz-Pérez, M., Maoyi, F., Xiaosheng, Y., & Belcher, B. (2001). Bamboo forestry in China: Toward environmentally friendly expansion. *Journal of Forestry, 99*(7), 14–20.

Sharma, C. M., Baduni, N. P., Gairola, S., Ghildiyal, S. K., & Suyal, S. (2010). Effects of slope aspects on forest compositions, community structures and soil properties in natural temperate forests of Garhwal Himalaya. *Journal of Forestry Research, 21*(3), 331–337.

Sheehan, P., Cheng, E., English, A., & Sun, F. (2014). China's response to the air pollution shock. *Nature Climate Change, 4*(5), 306–309.

Stage, A. R., & Salas, C. (2007). Interactions of elevation, aspect, and slope in models of forest species composition and productivity. *Forest Science, 53*(4), 486–492.

Stock, J. H., & Yogo, M. (2005). *Testing for weak instruments in linear IV regression, in identification and inference for econometric models: Essay in honor of Thomas Rothenberg*. Cambridge University Press.

Tanaka, S. (2015). Environmental regulations on air pollution in China and their impact on infant mortality. *Journal of Health Economics, 42*, 90–103.

Tang, P., Yang, S., Shen, J., & Fu, S. (2018). Does China's low-carbon pilot programme really take off? Evidence from land transfer of energy-intensive industry. *Energy Policy, 114*, 482–491.

Tajima, K. (2003). New estimates of the demand for urban green space: Implications for valuing the environmental benefits of Boston's big dig project. *Journal of Urban Affairs, 25*(5), 641–655.

Tsui, C.-C., Chen, Z.-S., & Hsieh, C.-F. (2004). Relationships between soil properties and slope position in a lowland rain forest of southern Taiwan. *Geoderma, 123*, 131–142.

Tu, Z., Hu, T., & Shen, R. (2019). Evaluating public participation impact on environmental protection and ecological efficiency in China: Evidence from PITI disclosure. *China Economic Review, 55*, 111–123.

Wan, L., Wang, C., Wang, S., Zang, J., & Li, J. (2021). How can government environmental enforcement and corporate environmental responsibility consensus reduce environmental emergencies? *Environmental Geochemistry and Health*. https://doi.org/10.1007/s10653-021-00916-z

Wang, C., Lin, J., Cai, W., & Zhang, Z. (2013). Policies and practices of low carbon city development in China. *Energy & Environment, 24*(7–8), 1347–1372.

Wang, L., & Watanabe, T. (2019). Effects of environmental policy on public risk perceptions of haze in Tianjin City: A difference-in-differences analysis. *Renewable and Sustainable Energy Reviews, 109*, 199–212.

Wendling, Z., Esty, D., Emerson, J., Levy, M., de Sherbinin, A., et al. (2018). *The 2018 environmental performance index report*. Yale Center for Environmental Law and Policy.

Xie, R. H., Yuan, Y. J., & Huang, J. J. (2017). Different types of environmental regulations and heterogeneous influence on "green" productivity: Evidence from China. *Ecological Economics, 132*, 104–112.

Xiong, Y. (2011). Uncertainty evaluation of the coordinated development of urban human settlement environment and economy in Changsha city. *Journal of Geographical Sciences, 21*(6), 1123–1137.

Xu, C., Dong, L., Yu, C., Zhang, Y., & Cheng, B. (2020). Can forest city construction affect urban air quality? The evidence from the Beijing-Tianjin-Hebei urban agglomeration of China. *Journal of Cleaner Production, 264*, 121607.

Yang, Y., Lu, Y., Yang, L., Gou, Z., & Liu, Y. (2021). Urban greenery cushions the decrease in leisure-time physical activity during the COVID-19 pandemic: A natural experimental study. *Urban Forestry & Urban Greening, 62*, 127136.

You, Z., Feng, Z., & Yang, Y. (2018). Relief degree of land surface dataset of China (1 km). *Journal of Global Change Data & Discovery, 2*(2), 151–155.

Yu, L. (2014). Low carbon eco-city: New approach for Chinese urbanisation. *Habitat International, 44*, 102–110.

Yu, Y., & Zhang, N. (2021). Low-carbon city pilot and carbon emission efficiency: Quasi-experimental evidence from China. *Energy Economics, 96*(2), 105125.

Yuan, X., & Zuo, J. (2011). Transition to low carbon energy policies in China: From the Five-Year Plan perspective. *Energy Policy, 39*(6), 3855–3859.

Zhang, Y., Zhang, T., Zeng, Y., Cheng, B., & Li, H. (2021). Designating national forest cities in China: Does the policy improve the urban living environment? *Forest Policy and Economics, 125*, 102400.

Zhao, C., Deng, M., & Cao, X. (2021). Does haze pollution damage urban innovation? Empirical evidence from China. *Environmental Science and Pollution Research, 28*(13), 16334–16349.

Zheng, S., Yi, H., & Li, H. (2015). The impacts of provincial energy and environmental policies on air pollution control in China. *Renewable and Sustainable Energy Reviews, 49*, 386–394.

Zhong, N., Cao, J., & Wang, Y. (2017). Traffic congestion, ambient air pollution, and health: Evidence from driving restrictions in Beijing. *Journal of the Association of Environmental and Resource Economists, 4*(3), 821–856.

Zhou, L., Tian, L., Gao, Y., Ling, Y., Fan, C., Hou, D., & Zhou, W. (2019). How did industrial land supply respond to transitions in state strategy? An analysis of prefecture-level cities in China from 2007 to 2016. *Land Use Policy, 87*, 104009.

Zhu, S., Li, D., & Feng, H. (2019). Is smart city resilient? Evidence from China. *Sustainable Cities and Society, 50*, 101636.

Zhao, C., Wu, Y., & Guo, J. (2022). Mobile payment and Chinese rural household consumption. *China Economic Review, 71*, 101719.

Zivin, G. J., & Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review, 102*(7), 3652–3673.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Environment, Development and Sustainability (2024) 26:6315–6340
https://doi.org/10.1007/s10668-023-02964-0

Did innovative city constructions reduce carbon emissions? A quasi-natural experiment in China

Xing Li[1] · **Chunkai Zhao**[2]

Received: 8 July 2021 / Accepted: 16 January 2023 / Published online: 28 January 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract
Using panel data for Chinese prefecture-level and above cities from 2003 to 2017, we employed the difference-in-differences (DID) method to explore whether innovative city constructions reduced carbon emissions. The results show that innovative city constructions significantly curb carbon emissions in Chinese cities, which is supported by a series of robustness tests. Specifically, we found that innovative city constructions could reduce carbon emissions by 5.21%, which is equivalent to a total of about 5.46 million tons. In addition, we provided evidence that innovative city constructions curb carbon emissions by promoting technological innovation, industrial upgrading and industrial green development, ecological improvement, and government intervention. Furthermore, heterogeneous results indicate that the carbon reduction effect of innovative city constructions is mainly found in eastern cities and cities with better industrial development and human capital base. Our findings further enrich the study on the environmental benefits of national innovation incentives and public policies and provide important insights in carbon neutrality and achieving sustainable development, especially for developing and emerging market countries.

Keywords Innovative city constructions · Carbon emissions · Carbon neutrality · Quasi-natural experiment · China

1 Introduction

Cities are the main boosters of high-quality economic development and the main sources of excessive carbon emissions, as they absorb large amounts of people and production resources (Liu et al., 2021). Recent estimates show that China's 338 prefecture-level and

✉ Chunkai Zhao
ckz_zhao@163.com

Xing Li
xingli277@126.com

[1] College of Economics and Management, Southwest University, No. 2, Tiansheng Road, Beibei District, Chongqing, China

[2] College of Economics and Management, South China Agricultural University, No.483, Wushan, Tianhe District, Guangzhou, Guangdong Province, China

above cities consume 85% of natural resources and discharge more than 75% of CO_2 (Liao et al., 2021). Therefore, in order to achieve China's goals of achieving carbon peaking by 2030 and carbon neutrality by 2060 (Gallagher et al., 2019; Peng & Deng, 2021), it is necessary to implement some emission reduction policies at the city level, like the low-carbon city pilot policy (Peng & Deng, 2021; Qiu et al., 2021; Yu & Zhang, 2021) and the carbon market pilot policy (Hu et al., 2020). Different from the previous studies evaluating the effects of environmental or climate policies on carbon emission reduction, we tried to explore the causal relationship between innovative city constructions based on national innovation incentives and carbon emissions.

As international competition has become increasingly fierce since the twenty-first century, many countries have made the construction of national innovation system a national strategy. China is no exception. At the Fifth Plenary Session of the 16th CPC Central Committee in 2005, the Chinese government proposed the idea of "constructing an innovative country." As an important support point for the national strategy of the innovative country construction, China started the pilot program of national innovative cities from Shenzhen in 2008. By 2022, there have been 106 cities established as national innovative cities. The goal of innovative city constructions is to improve the capacity of independent innovation, adjust the economic structure and achieve sustainable development (Yu et al., 2022; Zhou & Li, 2021). Innovative city constructions, which take into account the dual objectives of city innovation capacity enhancement and sustainable development, have the potential to become an important driver of carbon emission reduction in China.

Specifically, the inhibitory effect of innovative city constructions on carbon emissions may be summarized in the following aspects. First, the policy implementation of innovative city constructions may effectively enhance the innovation level of cities and promote technological innovation in industrial production (Yu et al., 2022), which in turn reduces carbon emissions (Ganda, 2019; Lv et al., 2021; Zhang et al., 2017). Second, innovative city construction may promote industrial upgrading and advanced industrial structure (Zhou & Li, 2021), transforming or transferring traditional industries with high consumption and high emission. Third, one of goals for innovative city constructions is to "promote coordinated economic and social development and sustainable development," which encourages ecological improvement and green development patterns in cities.

To explore the carbon reduction effect of innovative cities, we used the innovative city construction pilot as a quasi-natural experiment and adopted the difference-in-differences (DID) model to mitigate the potential endogeneity. Our estimates support the view that innovative cities had a carbon emission reduction effect in China. Moreover, our findings provide evidence that by enhancing innovation, promoting industrial upgrading and industrial green development, improving the ecological environment, and increasing government intervention, innovative city constructions played a role in carbon reduction.

The contributions of this study can be summarized as follows. First, although the role of public policies in reducing carbon emissions was well recognized in the previous literature (e.g., Hu et al., 2020; Liao et al., 2021; Yu & Zhang, 2021), little addressed to innovative cities in China. Therefore, this study enriched the research field on innovation incentives or public policies in reducing carbon emissions. In addition, this paper complemented the research area on the benefits of innovative city constructions from the perspective of environmental governance (e.g., Fang et al., 2014; Li et al., 2021; Yu et al., 2022; Zhou & Li, 2021).

Second, previous research estimated carbon emissions mainly from the perspective of energy consumption (Fu et al., 2021; Liu et al., 2021; Qiu et al., 2021), and this traditional calculation method may have some limitations (Fu et al., 2021; Liao et al., 2021). This

study drew on the practice of Chen (2020) and adopts nighttime light data as weights to modify the traditional estimation results of carbon emissions at the city level, which laid the foundation for more accurate calculation of carbon emissions and the evaluation of the effects of policies related to carbon reduction.

Third, we provided a comprehensive study of the mechanisms by which innovative city constructions constrain carbon emissions. Thus, these mechanisms offered insights for subsequent studies to further explore the environmental control effects of innovative city constructions and related policies in depth. Also, our heterogeneity analysis also provides references for some cities with weak innovation strength and human capital base.

2 Literature review and institutional background

2.1 Brief literature review

There are two strands of literature related to this paper. The first concerns measures to reduce carbon emissions, which have received close attention from numerous scholars. These mainly include energy structure (Isiksal et al., 2019; Pata & Caglar, 2021), international trade and investment (Tawiah et al., 2021; Thio et al., 2022), human capital (Huang et al., 2021; Joof & Isiksal, 2021; Sheraz et al., 2021), industrial structure (Liao et al., 2021; Wang et al., 2019; Zeng & Wang, 2022), finance (Qin et al., 2022; Samour et al., 2021), innovation capacity (Ganda, 2019; Lv et al., 2021), public policies (Hu et al., 2020; Isiksal, 2021; Lv et al., 2022; Xu & Li, 2022; Yu & Zhang, 2021), and so forth. In addition, a large number of studies were done on the relationship between GDP and carbon emissions (Isiksal et al., 2019; Thio et al., 2022), indicating that improving the quality of economic development is also one of important ways to reduce carbon emissions under the theoretical framework of Environmental Kuznets Curve (Isiksal, 2022; Zeng & Wang, 2022).

The second involves research on innovative cities. Previous studies described the characteristics and types of innovative cities (e.g., Caragliu & Del Bo, 2019; Chen & Karwan, 2008; Heinelt & Terizakis, 2021), and analyzed the innovation output effects (Fang et al., 2014; Zhou & Li, 2021). Particularly, some scholars have recently focused on the effects of innovative cities on the environment in China. For instance, Li et al. (2021) found that innovative city construction contributed to eco-efficiency in China, and that this effect is greater in cities with higher levels of technological innovation and human capital. Also, Yu et al. (2022) concluded that innovative city constructions improved the quality of innovation in Chinese cities, further contributing to energy productivity and an environmentally friendly development model.

2.2 Institutional background: innovative city constructions in China

Since the reform and opening up, China has made remarkable achievements in economic and social development, but the constraints of insufficient innovation capacity have gradually emerged. Worse, China is also facing enormous pressure from the technological dominance of developed countries. Therefore, at the Fifth Plenary Session of the 16th CPC Central Committee in 2005, China put forward the strategy of independent innovation and intention to "construct an innovative country."

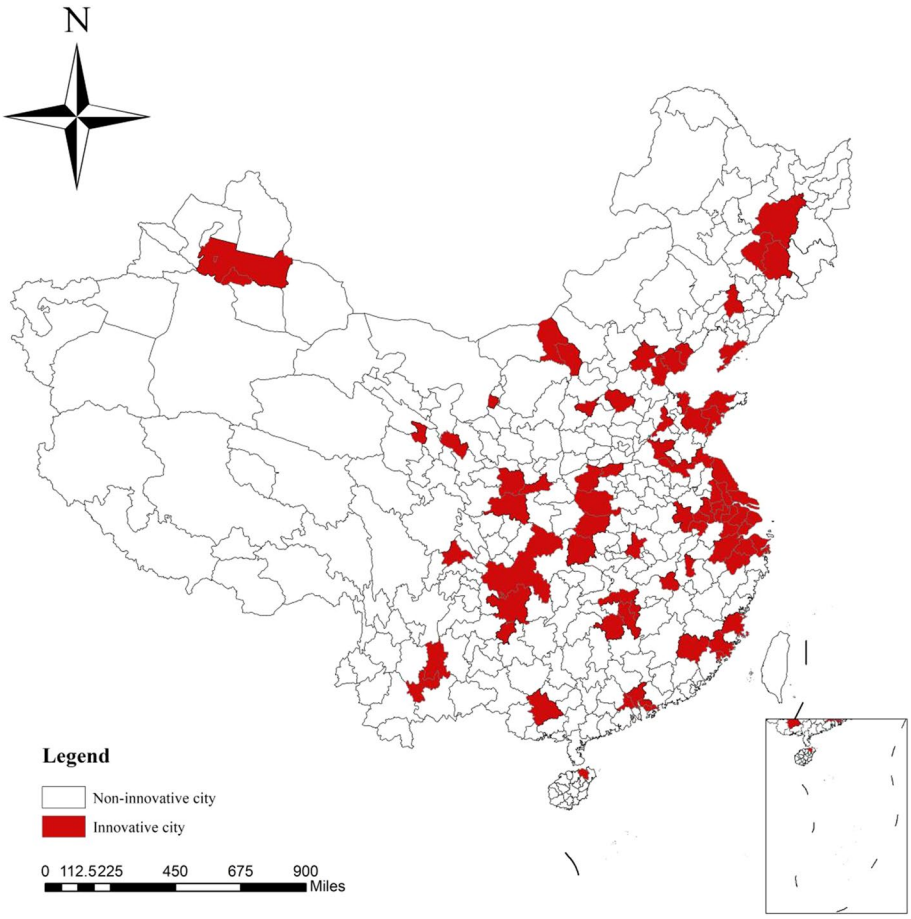


Fig. 1 Distribution of innovative cities in China

To achieve this ambitious goal, the Chinese government started to implement the policy of innovative city pilot in 2008, and Shenzhen became the first national innovative city in China. Subsequently, the National Development and Reform Commission established 60 more innovative cities (districts) between 2010 and 2013. In April 2018, 17 cities were approved to become new national innovative cities. Thus, the number of national innovative cities in China has increased to 78.[1] In Fig. 1, innovative cities have been distributed in 30 provinces (autonomous regions and municipalities) in China until 2016, with the largest number of 10 in Jiangsu province. Overall, innovative cities are concentrated in the Yangtze River Delta, Beijing-Tianjin-Hebei, and the middle and lower reaches of Yangtze River city clusters.

[1] In January 2022, the Ministry of Science and Technology and the National Development and Reform Commission announced a new group of pilot innovative cities, including 25 prefecture-level cities in 13 provinces. Since these cities are not in the time interval of this paper, they are not described in detail and are not shown in Fig. 1.

It should be emphasized that in April 2010, China's Ministry of Science and Technology promulgated the formulation of the "Guidance on Further Promoting the Pilot Work of Innovative Cities" and the "Indicators for Monitoring and Evaluating the Construction of Innovative Cities," which further standardized the selection, construction, and assessment of innovative cities in China. The evaluation indexes of national innovative city constructions mainly include innovation investment, enterprise innovation, achievement transformation, high-tech industry, technology that benefits people, and innovation environment, which strengthen the main position of enterprises and scientific research units. More importantly, in the section of high-tech industry, the emission standard and energy consumption of enterprises are required; in the part involving technology that benefits people, cities are especially encouraged to develop technologies that are conducive to improve the ecological environment, such as air pollution control and water pollution control.

3 Theoretical framework

In this section, we discussed in detail the mechanisms of the national innovative city constructions on carbon emissions. We classified possible channels into the following four categories: technological innovation, industrial upgrading and industrial green development, ecological improvement, and government intervention.

3.1 Technological innovation

The Theory of Endogenous Growth considers technological innovation as a source of promoting economic growth (Grossman & Helpman, 1994; Romer, 1994). However, innovation activities are characterized by high input, high risk, long cycle time, and technology spillover effects (Bloom et al., 2013; Liu & Buck, 2007; Tan, 2001), therefore, many previous studies advocated the promotion of technological innovation through some encouragement and nurturing policies (Acemoglu et al., 2016; Demir et al., 2020; Dolfsma & Seo, 2013). The national innovation system and the city innovation system theory are representative of them (Caragliu & Del Bo, 2019; Freeman, 1987), and they believe that innovation and application of knowledge and technology can be promoted through national and city institutional arrangements and coordination, which reflects the role of national innovation incentives in technological innovation.

As a new model of innovation development, innovative city constructions contribute to the formation of an organic innovation system, which becomes a spatial carrier and platform support for innovative subjects such as enterprises, universities, and research institutions (Ganda, 2019). The innovation system, preferential policies, infrastructures, and innovation atmosphere provided by innovative city constructions create favorable conditions for various technological innovation activities (Hospers, 2008; Zhang et al., 2017). Moreover, innovative city constructions optimize the allocation of resources and create a talent highland through national incentive mechanisms (Caragliu & Del Bo, 2019), which helps innovation agents to reduce innovation costs, optimize innovation mechanisms, share innovation risks, and transform results quickly (Chen & Karwan, 2008). As an important way to reduce carbon emissions, the role of technological innovation has been recognized by many scholars (e.g., Crawford & French, 2008; Irandoust, 2016; Lv et al., 2021; Shahbaz et al., 2020). Thus, we formulated the first hypothesis:

Hypothesis 1 Innovative city constructions may alleviate carbon emissions by facilitating technological innovation.

3.2 Industrial upgrading and industrial green development

The innovation incentives adopted by innovative cities in China include the establishment of high-tech development zones, science and technology parks, innovation incubation platforms, scientific research labs, industry-university-research bases, and so forth. These measures act on the city's target industries and related industries and may effectively promote industrial upgrading. On the one hand, a series of policies for the construction of innovative cities help to improve the inter-industry connection. With the incubation and settlement of related enterprises and the gradual improvement of the industrial chain (Alder et al., 2013), the value-added of the industry will be enhanced to realize the advanced leap of the industry under the effect of scale and agglomeration (Ellison & Glaeser, 1997; Mariotti et al., 2010; Zheng et al., 2017).

On the other hand, innovative city constructions promote the investment of traditional elements such as labor, capital, land, and R&D (Hospers, 2008). Under the effect of knowledge spillover and learning effects, emerging industries are continuously developed and traditional industries are transformed (Bloom et al., 2013; Jaffe et al., 1993), and some high technologies are applied to new fields to steadily increase the share of tertiary industries and facilitate industrial upgrading (Bernstein & Nadiri, 1989; Javorcik, 2004).

More importantly, innovative cities may promote the green development of industries. First, the policy orientation of innovative cities is to develop high-technology-based core industries, and these enterprises tend to be less dependent on fossil fuels (Du et al., 2019; Shahbaz et al., 2020; Xu & Lin, 2017). Under the policy dividend, the core industries and their industrial links are continuously improved and the scale effect is more prominent (Gourdon et al., 2016; Zheng et al., 2016), which leads to the survival space of traditional-type industries with high consumption and high emission being continuously compressed (Qiu et al., 2021; Rexhäuser & Rammer, 2014) and promotes the green and low-carbon development of the policy.

Second, innovative city construction helps accumulate human capital and diffuse knowledge through some talent introduction programs and industry-academia-research cooperation platforms, which further promotes technological improvements in traditional high-consumption and high-emission industries (Adams et al., 2002; Lawson & Lorenz, 1999). In addition, innovative city constructions contribute to the concentration of core industries, and the physical and spatial proximity facilitates the transformation of traditional industries and their participation in the industrial value chain division of labor (Jaffe et al., 1993; Mariotti et al., 2010; Zheng et al., 2017). With the transformation and restructuring of traditional industries with high consumption and high emissions, industries are gradually moving towards a green and sustainable development path. In summary, we derived the second hypothesis:

Hypothesis 2 Innovative city constructions are likely to reduce carbon emissions by promoting industrial upgrading and industrial green development.

3.3 Ecological improvement

In addition to technological innovation as well as industrial upgrading and industrial green development, innovative city construction may also benefit ecological improvement. Innovative cities constructions in China not only require the creation of a relaxed innovation environment, but also demand the construction of ecological civilization in cities. As an important type of innovative cities, cultural innovative cities emphasize the organic integration of city culture, history, ecological environment, and modern technology (Fang et al., 2014; Heinelt & Terizakis, 2021), which would be the development direction of some innovative cities in China.

Moreover, innovative cities can fully utilize their innovative capabilities to improve the ecological environment (Fang et al., 2014). The ecological landscape cultivation and planning of modern cities need to rely on plenty of scientific and technological resources (Ogunseitan, 2010; Van Hoek, 2002), and innovative city constructions can also make use of local ecological resources to develop green industries. For example, the development of industries such as urban greening system, ecological protection, pollution control, and garden cultivation not only helps to promote scientific and technological innovation in ecological industries (Abler & Shortle, 1995; Craig et al., 2016), but also may create an atmosphere of sustainable green development in the city (Bilgen and Sarıkaya, 2015; Fang et al., 2014).

The role of ecological improvement in curbing carbon emissions has been confirmed by numerous studies (e.g., Liao et al., 2021; Ortiz et al., 2014), especially the more powerful carbon sequestration capacity of artificially cultivated urban forests (Nowak et al., 2006; Qiu et al., 2021; Yu et al., 2019). Therefore, we proposed the third hypothesis:

Hypothesis 3 Innovative city constructions may reduce carbon emissions through ecological improvement.

3.4 Government intervention

In addition to these approaches, government intervention is also necessary. With the establishment of innovative city pilots, local governments may take a series of measures, such as more aggressive fiscal policies for environmental innovation and governance, improved green infrastructure, and stronger environmental regulations, which helps reduce carbon emissions.

Specifically, first, through common approaches such as financial support, loan subsidies, direct incentives, and project subsidies, local governments are able to incentivize enterprises' environmental technology research and development as well as green and smart production, stimulating their vitality and motivation to reduce carbon emissions. Second, local governments are likely to promote green infrastructure constructions, such as information network, urban greening, and public transportation networks, which greatly control carbon emissions. Finally, local governments can actively improve carbon accounting information system and sound environmental supervision system. By strengthening environmental regulation, cities achieve carbon reduction and sustainable development goals.

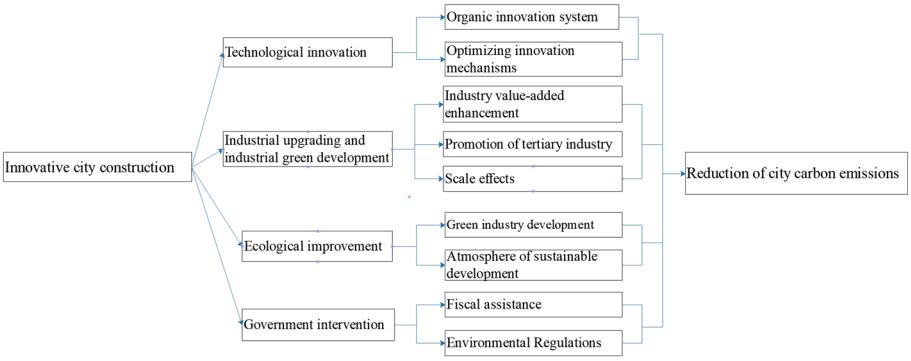


Fig. 2 The mechanisms of innovative city construction on carbon emissions

However, the impact of government intervention on reducing carbon emissions is uncertain (Li & Xu, 2020; Wang et al., 2019), and excessive intervention may lead to more serious environmental problems (Wang et al., 2021). Nevertheless, we proposed the fourth hypothesis:

Hypothesis 4 Innovative city constructions are likely to reduce carbon emissions by strengthening the government intervention.

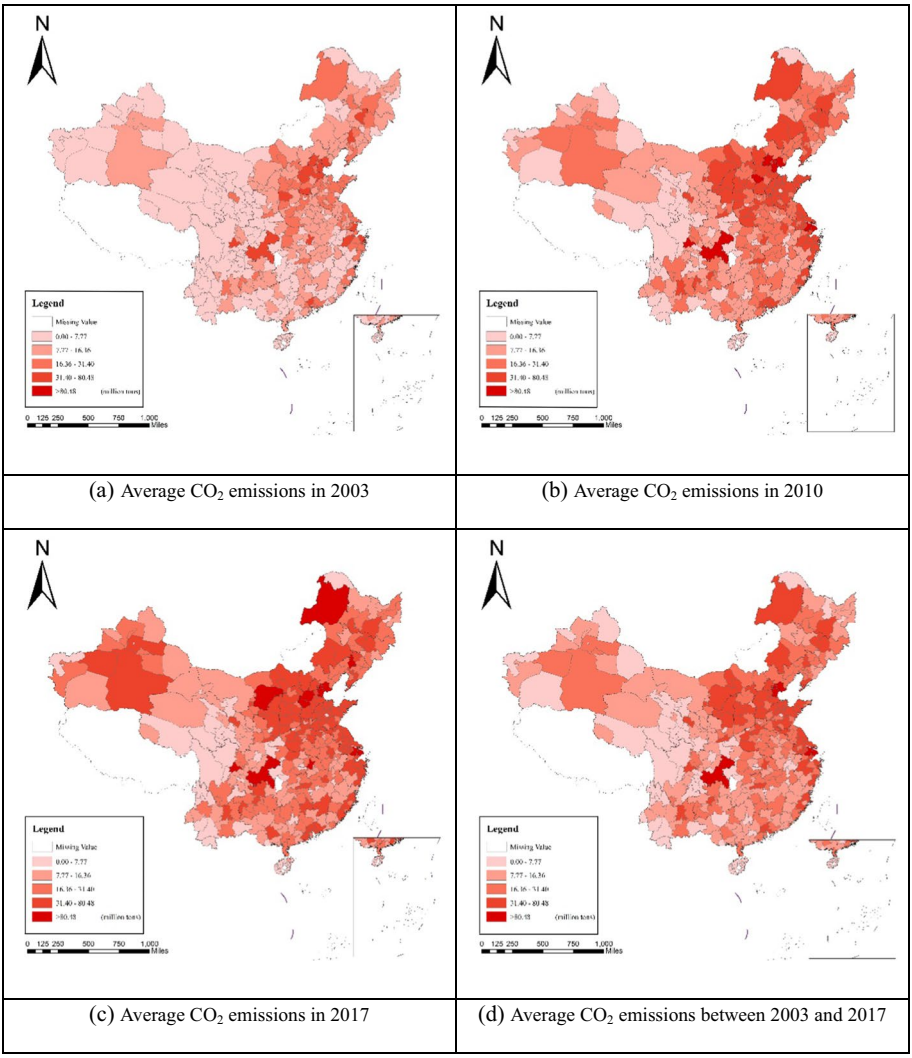
Given the multiplicity of mechanisms through which innovative city constructions reduce carbon emissions, we provided Fig. 2 to more clearly articulate these mechanisms. In addition, it is important to note that these mechanisms are not independent of each other, and it is very difficult to truly identify the role of each mechanism. Below using city panel data and the DID estimation method, we attempted to empirically examine which hypotheses are more consistent with our expectations and provide some suggestive evidence for the role of these mechanisms in explaining the carbon reduction effects of innovative city constructions (Fig. 3).

4 Data and methodology

4.1 Empirical strategy

Considering that the innovative city constructions in China are implemented in batches, it can be considered as a quasi-natural experiment and estimated by using the time-varying DID model. The main advantage of adopting the DID approach is that it can address the endogeneity problem well and is widely used in environmental economics (e.g., Li et al., 2021; Liao et al., 2021; Qiu et al., 2021; Yu & Zhang, 2021). When assessing the effects of public policy, comparing the full group before and after policies or cross-sectionally comparing treatment and control groups may lead to biased estimates. By capturing the relative differences between the treatment and control groups before and after policy implementations, the DID model can eliminate the estimation bias introduced by intervention factors (Li et al., 2022; Liao et al., 2021; Yu & Zhang, 2021). More specifically, the estimated model can be written as follows:

$$Carbon_{it} = \beta_0 + \beta_1 Innovative_{it} + \beta_2 X_{it} + \delta_i + \gamma_t + \varepsilon_{it} \quad (1)$$



(a) Average CO_2 emissions in 2003

(b) Average CO_2 emissions in 2010

(c) Average CO_2 emissions in 2017

(d) Average CO_2 emissions between 2003 and 2017

Fig. 3 Average CO_2 emissions over the years

where the subscript i and t represent the city and year, respectively. $Carbon_{it}$ denotes the carbon emissions of city i in year t. $Innovative_{it}$ is a dummy variable, indicating whether city i is designated as an innovative city in year t. β_1 is the coefficient we were most concerned about, which refers to the impact of the treatment group selected as an innovative city on carbon emissions compared with the control group of non-innovative city. X_{it} is a series of control variables at the city level, including economic variables and climate variables. δ_i indicates the city fixed effect and γ_t is the year fixed effect. ε_{it} denotes the random error term.

The prerequisite for the application of the DID model is that the treatment group and the control group have the same parallel trends before being rated as national innovative

cities (Liao et al., 2021; Zheng et al., 2016; Zhou & Li, 2021). To check the applicability of DID approach, we constructed formula (2) to perform parallel trend testing.

$$Carbon_{it} = \alpha_0 + \alpha_1 Innovative_{it} + \sum \lambda_d T_{id} + \alpha_2 X_{it} + \delta_i + \gamma_t + \varepsilon_{it} \quad (2)$$

In formula (2), subscript d accounts for the time interval from the implementation of the innovative city pilot policy. For example, when d equals −3, representing the 3 years before the implementation of the innovative city pilot policy. The coefficient λ_d indicates the difference between the control group and the control group on carbon emissions before and after the implementation of the policy. Accordingly, it is hypothesized that λ_d should be insignificant when d is negative (before policy implementation).

4.2 Data and variables

4.2.1 Carbon emissions

The core explained variable is carbon emissions at the city level, of which original data are from Chen et al. (2020). Specifically, based on the particle swarm optimization-back-propagation (PSO-BP) algorithm, Chen et al. (2020) corrected the satellite images and then constructed the relationship between nighttime light data and CO_2 emission data of fossil energy consumption in each Chinese province/municipality. Using the total light brightness of each district/county as the weight, Chen et al. (2020) inverted the CO_2 emissions. Lastly, by summing up, we can obtain the city-level carbon emissions from 2003 to 2017.[2]

Figure 3 shows the carbon emissions of each city in 2003, 2010, 2017, and the average carbon emissions during the sample period. Compared with 2003, the carbon emissions of each city in 2010 have increased significantly. In 2017, the carbon emissions of prefecture-level cities also increased compared with 2010, but the increasing trend of carbon emissions was significantly weakened. Additionally, average carbon emissions of prefecture-level cities from 2003 to 2017 is similar, or even lower than 2010. Overall, northern and eastern cities emit more CO_2.

4.2.2 Innovative city

Our explanatory variable is a dummy variable for whether it is selected as an innovative city. The data are from the selection of all cities by the Ministry of Science and Technology of the People´s Republic of China since 2008. As noted above, we marked the cities that became national innovative cities in the current year and later years as 1, and 0 otherwise. As mentioned in Sect. 2, by 2017, the number of innovative city pilots nationwide has reached 79.

4.2.3 City characteristics and climate variables

With reference to the existing literature, we controlled for some variables at the city level for more accurate and unbiased estimation results. At the economic level, considering the secondary industry development, economic development, and fixed asset

[2] We chose this time range mainly because there are many missing values in the data of prefecture-level cities before 2003, and the latest carbon emission data until 2017.

Table 1 Descriptive statistics

Variables	Units	Observations	Mean	SD	Min	Max
Carbon	Million tons	5,430	21.278	19.571	0.9081	153.75
Innovative	Dummy, 0/1	5,764	0.0845	0.2781	0	1
Secondary industry	%	5,632	41.967	15.492	0.5900	98.820
Economic growth	%	5,022	11.278	4.6830	−39.800	37.690
Fixed assets Investment	%	5,043	7.3940	11.565	0.0498	172.46
Population density	1000 people/ square kilometer	4,755	0.4182	0.3493	0.0047	11.564
Temperature	℃	4,092	14.610	5.3824	−2.2000	27.880
Humidity	%	4,092	66.526	10.762	29.000	88.000
Precipitation	mm	4,038	2.5876	1.5112	0.1000	7.4271
Sunshine	Hour	4,076	2036.0	544.94	247.00	3672.9

investment may have an impact on carbon emissions (Liu et al., 2021; Lozano & Gutiérrez, 2008; Yu & Zhang, 2021; Yu et al., 2022), we controlled for these variables. More specifically, *Secondary industry* (%) refers to the proportion of employees in the secondary industry and *Economic growth* (%) is measured by the growth rate of city's GDP. *Fixed assets investment* (%) is defined by the ratio of fixed asset investment to GDP is expressed. Moreover, *Population density* is also included in the control variable set, considering that more densely populated cities are also likely to have more indirect carbon emissions from consumption and production activities (Liao et al., 2021). The unit of *Population density* is 1000 people per square kilometer. These variables are derived from China City Statistical Yearbook.

In addition, since climate factors such as temperature can also be an important factor in carbon emissions (Kaufmann et al., 2013; Liao et al., 2021), we controlled for *Temperature* (℃), *Humidity* (%), *Precipitation* (average daily, mm), and *Sunshine* (average daily, hours). The first two variables are measured by the average annual temperature and humidity of each city, and the last two refer to the total precipitation and total sunshine hours of cities. These climate variable data come from China Meteorological Administration.

4.3 Descriptive statistics

Table 1 reports the descriptive statistics of the sample data, and all continuous variables (e.g., *Secondary industry*) are winsorized by replacing samples among the top 1% and the bottom 1%. China's innovative city pilots account for 8.45% of the total number of sample cities and the average carbon emissions of 21.278 million tons. In addition, the minimum carbon emission of different cities is 0.9081 and the maximum is 153.75, which implies a large variation in carbon emissions across Chinese cities.

Table 2 Baseline results

	(1)	(2)	(3)	(4)
Innovative	−0.0242***	−0.0181*	−0.0194**	−0.0521**
	(0.0079)	(0.0099)	(0.0078)	(0.0213)
Secondary industry		0.0011***		0.0127***
		(0.0002)		(0.0004)
Economic growth		−0.1158**		−1.2619***
		(0.0466)		(0.1078)
Fixed assets Investment		0.0002		0.0167***
		(0.0003)		(0.0007)
Population density		0.0000		0.0003***
		(0.0000)		(0.0001)
Temperature			0.0059	−0.0720***
			(0.0046)	(0.0077)
Humidity			−0.0017***	−0.0116***
			(0.0006)	(0.0013)
Precipitation			0.0159***	0.0253**
			(0.0056)	(0.0126)
Sunshine			−1.7653	−16.3482***
			(1.1206)	(2.4527)
Year fixed effect	Yes	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes	Yes
Constant	1.9748***	1.9772***	3.2063***	4.4221***
	(0.0067)	(0.0148)	(0.1025)	(0.1806)
Adjusted R^2	0.9333	0.9321	0.9076	0.9035
Observations	5430	4426	3468	3202
Cities	291	289	251	251

***, **, * indicate significant at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are in parentheses

5 Empirical results

5.1 Baseline results

To verify the causal relationship between innovative city constructions and carbon emissions, we used formula (1) to perform baseline regression, and Table 2 reports the estimated results. In Table 2, column (1) only controls the city and year fixed effects, while columns (2) and (3) further control the city characteristics and climate variables, respectively. In column (4), we added all control variables and fixed effects, and the results show that innovative city constructions can significantly reduce carbon emissions. On average, innovative city constructions reduce carbon emissions by 5.21%, that is, it can reduce 5.46 million tons of carbon emissions for all cities during the sample period. By using DID estimates, our results confirm that innovative city constructions can be effective in reducing carbon emissions in Chinese cities.

For the control variables, secondary industry, fixed asset investment, and population density are all positively correlated with carbon emissions, which is because manufacturing

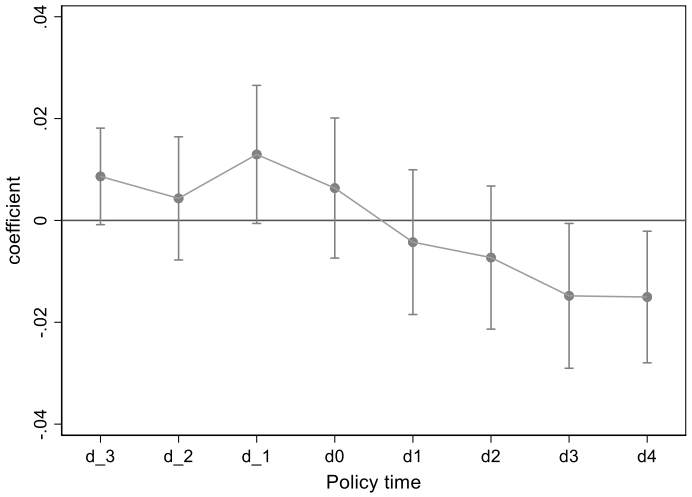


Fig. 4 Parallel trend tests

and fixed assets investment are important sources for carbon emissions in China (Hu et al., 2020; Xu & Li, 2022; Zeng & Wang, 2022). Economic growth is negatively correlated with carbon emissions, suggesting that as the economic quality improves, it can contribute to carbon reduction in Chinese cities (Isiksal, 2022; Zhang et al., 2017). The estimated results for climate variables are similar to the discussions in previous literature (Li et al., 2022; Liao et al., 2021).

5.2 Parallel trend tests

The results are presented in Fig. 4. We found that that coefficients on *Innovative* are all insignificant before the implementation of the policy (before the d0 period), indicating that before being rated as an innovative city, there was no significant difference in carbon emissions between the treatment group and control group. Thus, these results support the parallel trend assumption, indicating that adopting the DID method to assess innovative city constructions is reasonable and convincing. Additionally, Fig. 4 implies that after the policy is implemented, the curve continues to trend downwards, and the significance of the estimated coefficients increase, indicating that innovative city constructions do inhibit carbon emissions and realize the sustainable development of cities.

5.3 Mechanism tests

5.3.1 Technological innovation

To test the Hypothesis 1, we constructed four indicators to measure technological innovation of cities from three aspects: city innovation capacity, innovation input, and innovation output. Among them, city innovation capacity is derived from the comprehensive innovation index of Chinese cities constructed by Kou and Liu (2017). Innovation input is measured by the percentage of government spending on science and technology. Innovation

Table 3 Innovative city constructions, technological innovation, and carbon emissions

	(1)	(2)	(3)	(4)
Innovative*Innovation input	−0.8551***			
	(0.0084)			
Innovative*Patent applications		−0.7524***		
		(0.0086)		
Innovative*Green invention patents			−0.5813***	
			(0.0085)	
Innovative*City innovation capacity				−0.4634***
				(0.0084)
Control variables	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes	Yes
Adjusted R^2	0.9522	0.9316	0.9076	0.9004
Observations	3202	3202	3202	3202
Cities	251	251	251	251

***, **, * indicate significant at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are in parentheses

output is consistent with previous literature (Lv et al., 2021; Zhang et al., 2017), and the number of patent applications and green patent grants are used as proxy indicators.

We used the construction of interaction terms to verify these mechanisms, and the results are shown in Table 3. It is clear that the coefficients of these interaction terms are all negative and significant at the 1% level in each column, which fully indicates that innovative city construction reduces carbon emissions by promoting technological innovation in cities. More importantly, we found that both innovation inputs and innovation outputs as well as the comprehensive innovation capacity of cities play a role in carbon reduction. In sum, our results support the Hypothesis 1 that innovative city constructions can reduce carbon emissions by promoting technological innovation.

5.3.2 Industrial upgrading and industrial green development

In the Hypothesis 2, we considered that innovative city constructions reduce carbon emissions by promoting industrial upgrading and industrial green development. We used the industrial structure upgrading index to measure industrial upgrading (Zhao et al., 2021).[3] In industrial green development, we constructed two mechanism variables, soot removal rate and environment-related employment. Similarly, we constructed the interaction terms between innovative cities and these mechanism variables separately, and the results are shown in Table 4. In column (1), we found that the coefficient on the interaction term between *Innovative* and *Industrial structure* is significantly negative, indicating that the innovative city constructions may reduce carbon emissions by improving the industrial structure of cities. Moreover, the coefficient on the interaction term is significant in column (2), suggesting that innovative city constructions may improve the pollution control technology of enterprises to reduce

[3] The calculation method is: the primary industry*1 + the secondary industry*2 + the tertiary industry*3.

Table 4 Innovative city constructions, industrial upgrading and industrial green development, and carbon emissions

	(1)	(2)	(3)
Innovative*Industrial structure	-0.3996***		
	(0.0084)		
Innovative*Soot removal rate		-0.3309***	
		(0.0084)	
Innovative*Environment-rated employment			-0.1546***
			(0.0082)
Control variables	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes
Adjusted R^2	0.9271	0.9211	0.9003
Observations	3202	3202	3202
Cities	251	251	251

***, **, * indicate significant at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are in parentheses

Table 5 Innovative city constructions, ecological improvement, and carbon emissions

	(1)	(2)	(3)
Innovative*Green space	-0.1268***		
	(0.0096)		
Innovative*Greening of built-up areas		-0.2863***	
		(0.0097)	
Innovative*Vegetation carbon sequestration			-0.3899***
			(0.0101)
Control variables	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes
Adjusted R^2	0.9261	0.9388	0.9125
Observations	3202	3202	3202
Cities	251	251	251

***, **, * indicate significant at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are in parentheses

carbon emissions. In column (3), the coefficient on the interaction term is also negative and significant at the 1% level, which implies that the innovative city constructions can reduce carbon emissions by promoting environment-related employment. All in all, these results provide supportive evidence for the Hypothesis 2.

Table 6 Innovative city constructions, government intervention, and carbon emissions

	(1)	(2)	(3)
Innovative*Government budget expenditure	-0.1627***		
	(0.0607)		
Innovative*Industrial SO_2 Removal		-0.0083***	
		(0.0018)	
Innovative*Utilization rate			-0.2444***
			(0.0724)
Control variables	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes
Adjusted R^2	0.9132	0.9055	0.9271
Observations	3202	3202	3202
Cities	251	251	251

***, **, * indicate significant at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are in parentheses

5.3.3 Ecological improvement

As discussed in the Hypothesis 3, the task of innovative cities is to promote coordinated and sustainable economic and social development and the construction of innovative ecological civilization may promote ecological improvement, thus indirectly reducing carbon emissions of cities. We selected three variables, namely green space, greening of built-up area, and vegetation carbon sequestration, to reflect the ecological status of cities. The first two variables are obtained from the China City Statistical Yearbook, and the data of the last variable are from Chen et al. (2020). Table 5 reports the estimation results with ecological improvement as the mechanism, and we found that interaction terms are significantly negative in the first two columns, implying that innovative city constructions can achieve low-carbon development by improving the greening level of cities. Moreover, in column (3), although the coefficient of the interaction term is negative, it is not statistically significant. Therefore, our results support the Hypothesis 3 that innovative city constructions achieve carbon reduction by improving the ecological environment.

5.3.4 Government intervention

In addition to local government public budget expenditures, the existing literature generally regards the removal of industrial sulfur dioxide and the comprehensive utilization rate of general industrial solid waste as proxy variables for the intensity of government intervention (Li et al., 2022) This paper also multiplies innovative cities with this three variables that characterize the intensity of government intervention to study the impact of government intervention on carbon emissions in innovative cities.

Column (1) of Table 6 reports the impact of the interaction term of innovative cities and government public budget expenditure on CO_2 emissions, column (2) shows the impression of the interaction term of innovative cities and industrial SO_2 removal on CO_2 emissions, and column (3) reports the impact of the interaction term between innovative cities

Table 7 Heterogeneity effects by human capital base and industrial development level

	Human capital base		Industrial development level	
	Higher	Lower	Higher	Lower
	(1)	(2)	(3)	(4)
Innovative	−0.6730***	−0.5589	−0.8068***	−0.7396
	(0.0100)	(0.5100)	(0.0100)	(0.7102)
Control variables	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes	Yes
Adjusted R^2	0.9024	0.9011	0.9021	0.9006
Observations	1867	1335	1602	1600
Cities	148	103	126	125

***, **, * indicate significant at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are in parentheses

and general industrial solid waste comprehensive utilization on CO_2 emissions. We found that after multiplying innovative cities with variables that represent government intervention, carbon emissions can be significantly suppressed, indicating that innovative cities can suppress carbon emissions by promoting government intervention. In summary, our results support the Hypothesis 4 that innovative city constructions achieve carbon reduction by government intervention.

5.4 Heterogeneity analysis

Considering that the carbon reduction benefits of innovative city constructions may vary depending on some city characteristics, we conducted heterogeneity analysis in terms of human capital base, industrial development level, and geographic location.

In Table 7, we grouped cities according to the number of general universities and the quantile of gross industrial output in terms of their human capital base and industrial development level. We found that the inhibitory effect of innovative city constructions on carbon emissions is more effective in cities with higher levels of industrial innovation and technological innovation. It is possible that cities with better industrial development and human capital base can take advantage of the policy dividends of innovative city constructions and provide better industrial base and human capital for innovation construction in cities.

Furthermore, we divided all prefecture-level cities in China into eastern, central, and western regions.[4] The results in Table 8 show that innovative city constructions in the eastern cities are more able to reduce carbon emissions, while in the central and western regions, innovative city constructions have little effects. The possible reason is that due to the developed economy in the eastern region, the human and material capital required for

[4] The eastern region includes eight provinces, namely Hebei, Shandong, Liaoning, Jiangsu, Zhejiang, Fujian, Guangdong, and Hainan. There are eight provinces in the central region, and they are Heilongjiang, Jilin, Shanxi, Jiangxi, Anhui, Henan, Hubei, and Hunan. The remaining 11 provinces or autonomous regions belong to the western region, namely Xinjiang, Inner Mongolia, Ningxia, Gansu, Qinghai, Tibet, Yunnan, Guizhou, Sichuan, Guangxi, and Shaanxi.

Table 8 Heterogeneity effects by locations

	Eastern cities (1)	Central cities (2)	Western cities (3)
Innovative	−0.1181***	−0.1548	−0.0924
	(0.0102)	(0.0905)	(0.1107)
Control variables	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes
Adjusted R^2	0.9044	0.9217	0.9005
Observations	825	959	1418
Cities	77	69	105

***, **, * indicate significant at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are in parentheses

innovation flows to the eastern cities, therefore, the efficiency of innovation in the eastern cities is higher, and the inhibitory effect on carbon emissions is more significant.

5.5 Robustness checks

We have done a series of robustness checks in this section in order to make the conclusions more convincing. First, the reduction in carbon emissions may not be due to innovative city policies, but due to the impact of other policies in the same year. Therefore, we advanced the year of all selected innovative cities by 1 year, 2 years, and 3 years to conduct the placebo tests. If the results are still significantly negative, it means that the reduction in carbon emissions is not caused by innovative cities. The results in Table 9 show that after advancing the selection time of all innovative cities by 1, 2, or 3 years, the inhibitory effect of innovative city policies on carbon emissions is not significant. These estimates indicate that the reduction in carbon emissions is indeed caused by innovative cities.

Second, we eliminated interference from other policies. With increasing emphasis on carbon emissions, China has also formulated a number of climate policies that target carbon emissions reductions, such as carbon market pilots and low-carbon city pilots (e.g., Peng & Deng, 2021; Yang et al., 2022; Yu & Zhang, 2021). These climate policies will cause overestimation of the effectiveness of innovative cities to curb carbon emissions. In columns (1) and (2) of Table 10, we removed all low-carbon city pilots and carbon-trading city pilots, respectively, finding that our results remain unchanged.

In addition to climate policies, the implementation of air pollution control policies may also lead to overestimation of the carbon emission reduction effects of innovative cities. Since China is a coal-based energy consumer, carbon emissions will inevitably be accompanied by a large amount of air pollutants (Hu et al., 2020; Yang et al., 2022). Therefore, air pollution control policies will also curb carbon emissions. According to the Air Pollution Prevention and Control Action Plan, we deleted the "2 + 26" cities with the most serious air pollution (Li et al., 2022).[5] In column (3) of Table 10, we found that excluding these policies has little impact on our main results.

[5] These cities include Beijing, Tianjin, Shijiazhuang, Tangshan, Langfang, Baoding, Cangzhou, Hengshui, Xingtai, Handan, Taiyuan, Yangquan, Changzhi, Jincheng, Jinan, Zibo, Jining, Dezhou, Liaocheng, Binzhou, Heze, Zhengzhou, Kaifeng, Anyang, Hebi, Xinxiang, Jiaozuo, and Puyang.

Table 9 Placebo tests

	(1)	(2)	(3)
Innovative	−0.7366	−0.5951	−0.3299
	(0.7951)	(0.8586)	(0.9583)
Control variables	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes
Adjusted R^2	0.9322	0.9061	0.9137
Observations	3202	3202	3202
Cities	251	251	251

***, **, * indicate significant at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are in parentheses

Table 10 Excluding other policy interference

	Low-carbon cities	Carbon market trading cities	"2 + 26" cities	Smart cities
	(1)	(2)	(3)	(4)
Innovative	−0.0886***	−0.3130***	−0.0213***	−0.0004***
	(0.0250)	(0.0265)	(0.0007)	(0.0001)
Control variables	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes	Yes
Adjusted R^2	0.9137	0.9265	0.9199	0.9214
Observations	2734	2926	3101	1421
Cities	201	212	228	103

***, **, * indicate significant at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are in parentheses

We also ruled out the interference of smart city policies. Similar to innovative cities, smart cities have sustainable innovation characteristics of user innovation, open innovation, mass innovation, and collaborative innovation (Haarstad & Wathne, 2019). Smart cities emphasize the realization of comprehensive sustainable economic, social, and environmental development through value creation and people-oriented (Jiang et al., 2021; Yigitcanlar & Kamruzzaman, 2018). In column (4) of Table 10, after excluding the sample of smart cities, the estimates are qualitatively similar to our main results in Table 2.

Third, to eliminate the potential bias in the model setting, we tested the robustness of the results by replacing key variables and models. In column (1) of Table 11, we adopted carbon emissions per capita as the explanatory variable, while employ the PSM-DID approach to remove self-selection effects in column (2). We found that these results are virtually unchanged in Table 11.

Fourth, we excluded some special observations. As mentioned above, to obtain the city brand of an innovative city, the government will also guarantee it by increasing expenditures on science and education. Since the provincial capital is the political center of a province and represents the image of a province, the provincial government will give priority to ensuring the operation of provincial capital city policies by concentrating a large number

Table 11 Using alternative specifications and excluding some special observations

	Carbon emissions per capita	PSM-DID	Excluding innovative cities established before April 2010	Excluding municipalities, provincial capitals, and sub-provincial cities	Excluding super innovative cities
	(1)	(2)	(3)	(4)	(5)
Innovative	−0.0213***	−0.0004***	−0.0732***	−0.0211***	−0.0297**
	(0.0007)	(0.0001)	(0.0091)	(0.0015)	(0.0148)
Control variables	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.8996	0.9137	0.9061	0.9211	0.9036
Observations	3202	2706	2976	2981	3348
Cities	251	246	240	216	248

***, **, * indicate significant at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are in parentheses

of policy resources. This kind of government behavior will also overestimate the results. Column (3) of Table 11 reports the results after deleting municipalities, provincial capitals, and sub-provincial cities. We found that after excluding government influence, innovative city constructions can still significantly suppress carbon emissions.

In addition, as explained in Sect. 2, some cities had been established as innovative cities before the Ministry of Science and Technology of China promulgated the developed Monitoring and Evaluation Indicators for Innovative City Construction. Therefore, in the robustness test, we removed these 17 cities that were explicitly designated as innovative cities before April 2010. In column (4) of Table 11, we found that the impact of this change on our main specification is minimal.

Furthermore, the "China City Innovation Report (2013)" ranked the innovation capabilities of cities. Three provincial capitals and above[6] with the strongest innovation capacity include Beijing, Shanghai, and Shenzhen, and among the ordinary prefecture-level cities, Suzhou, Dongguan, and Foshan occupy the top three. These cities may become super innovative cities due to a series of inherent advantages. In column (5) of Table 11, we found that after excluding these six super innovative cities, the coefficient on *Innovative* still keeps significantly negative at the 10% level, suggesting that innovative city constructions still achieve carbon emission reductions.

6 Conclusions and discussion

6.1 Conclusions

As China's economy enters the stage of high-quality development, how to achieve the dual goals of economic growth and carbon emission reduction has attracted widespread attention. The innovative city constructions based on national innovation incentives may provide a new solution. By utilizing the quasi-natural experiment of innovative city constructions, we used annual data from 281 prefecture-level cities in China from 2003 to 2017 and employ the DID model to identify the causal relationship between innovative city constructions and carbon emissions. The results show that innovative city constructions can significantly curb carbon emissions, which is supported by some robustness checks, including placebo tests, eliminating other policy interference, and replacing key variables and estimation methods. Specifically, we found that innovative city constructions reduce carbon emissions by 5.21%, which is approximately equivalent to 5.46 million tons for all cities. In addition, we explored several potential mechanisms to study of how innovative city constructions affect carbon emissions. We found that the positive effect of innovative construction on reducing carbon emissions is explained by the fact that innovative cities can enhance urban innovation, promote industrial upgrading and industrial green development, and benefit ecological improvement. Furthermore, heterogeneity analysis results indicate that the impact of the innovative city constructions on carbon emissions is valid for only eastern cities and cities with better industrial development and human capital base.

[6] These cities include 27 provincial capitals, 4 municipalities directly under the central government and 15 sub-provincial cities. In terms of administrative level, these three types of cities are higher than the average prefecture-level city.

6.2 Discussion

6.2.1 Theoretical contributions

This paper mainly has the following three contributions. First, previous studies confirmed that some environmental or climate policies can be effective in curbing carbon emissions at the city level; however, strict environmental policies may be detrimental to economic growth. Based on a quasi-natural experiment of innovative city constructions, we found that with the help of innovation drivers, innovative cities contribute to achieving the dual goals of economic growth and carbon emission reduction. In addition, although two studies assessed the environmental effects of innovative city constructions in China (Li et al., 2022; Yu et al., 2022), they did not focus on carbon emissions. Therefore, this paper further enriches the literature related to the realization of carbon neutrality in the policy context of carbon neutrality.

Second, due to the lack of monitoring data in China, traditional literature only uses input–output methods such as energy consumption to estimate carbon emissions (Fu et al., 2021; Liu et al., 2021; Qiu et al., 2021). We drew on the practice of Chen (2020) and use nighttime light data as weights to modify the traditional estimation results of carbon emissions at the prefecture-level city level, which provided insights for more accurate calculation of carbon emissions when exploring paths to reduce carbon emissions.

Third, we analyzed and checked in detail the role of four mechanisms by which innovative city constructions reduce carbon emissions: technological innovation, industrial upgrading and industrial green development, ecological improvement, and government intervention. Although some mechanisms, such as technological innovation, have been identified in the previous literature (Zhou & Li, 2021), we further extended the effects of these mechanisms to the environmental level. Thus, our study provides a reference for understanding the mechanisms of the role of public policies such as innovative city constructions in influencing carbon emissions.

6.2.2 Policy implications

The implications of this paper are quite obvious. First, we confirmed that government incentives for innovation work. Cities can find a balance between economic growth and environmental protection through paths such as technological innovation, industrial upgrading, and green development. Therefore, it is necessary to take some effective measures to exploit the environmental benefits of innovative city constructions. Possible measures include actively introducing green industries, providing more subsidies to environmentally friendly enterprises, and improving low-carbon technologies. Second, the results of heterogeneity analysis suggest that the carbon reduction effect of innovative city constructions is greater in cities with better industrial and human capital bases. Therefore, for some cities with weak industrial or human capital bases, the policy dividends of innovative city constructions should be fully utilized to achieve economic growth and sustainable development.

These conclusions and insights are not only applicable to China, but can also be used as references for other countries. Carbon reduction and the promotion of sustainable development have become a global consensus, and the dual goals of economic development and carbon reduction through national innovation incentives are worth promoting to other countries. China's experience proves that innovative city constructions can be more effective in

improving technological innovation and promoting green development in industries. In a word, our study confirms the environmental benefits of innovative cities, which provides a template for other countries, especially in some developing and emerging market countries with high carbon emissions.

6.2.3 Limitations and future studies

There are some limitations in this paper. First, although we used China's prefecture-level cities as a sample and obtain robust results, it lacks micro evidence. In fact, it is quite possible that innovative city constructions influence low-carbon and green behaviors of local residents and enterprises. Therefore, the follow-up studies may consider collecting some data at the enterprise and individual level to provide more detail evidence about innovative cities. Second, our measurement of some mechanisms may not be comprehensive enough, such as industrial green development and ecological improvement. As more city- or county-level data become publicly available, perhaps further studies could test the effectiveness of these mechanisms from a wider range of perspectives.

Funding This work was supported by the Fundamental Research Funds for the Central Universities (No. QCDC-2020–21).

Data availability All data included in this study are available upon request by contact with the corresponding author.

Declarations

Conflict of interest We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

Abler, D. G., & Shortle, J. S. (1995). Technology as an agricultural pollution control policy. *American Journal of Agricultural Economics, 77*(1), 20–32.

Acemoglu, D., Moscona, J., & Robinson, J. A. (2016). State capacity and American technology: Evidence from the nineteenth century. *American Economic Review, 106*(5), 61–67.

Adams, J. D. (2002). Comparative localization of academic and industrial spillovers. *Journal of Economic Geography, 2*(3), 253–278.

Alder, S., Shao, L., & Zilibotti, F. (2013). Economic reforms and industrial policy in a panel of Chinese cities. *Journal of Economic Growth, 21*(4), 305–349.

Bernstein, J. I., & Nadiri, M. I. (1989). Research and development and intra-industry spillovers: An empirical application of dynamic duality. *Review of Economic Studies, 56*(2), 249–267.

Bilgen, S., & Sarıkaya, İ. (2015). Exergy for environment, ecology and sustainable development. *Renewable and Sustainable Energy Reviews, 51*, 1115–1131.

Bloom, N., Schankerman, M., & Reenen, J. V. (2013). Identifying technology spillovers and product market rivalry. *Econometrica, 81*(4), 1347–1393.

Caragliu, A., & Del Bo, C. F. (2019). Smart innovative cities: The impact of smart city policies on urban innovation. *Technological Forecasting and Social Change, 142*, 373–383.

Chen, J., Gao, M., Cheng, S., Hou, W., Song, M., Liu, X., Liu, Y., & Shan, Y. (2020). County-level CO_2 emissions and sequestration in China during 1997–2017. *Scientific Data, 7*, 391.

Chen, S., & Karwan, K. (2008). Innovative cities in China: Lessons from Pudong New District, Zhangjiang high-tech park and SMIC village. *Innovation, 10*(2–3), 247–256.

Craig, M., Mcdonald-Buller, E., & Webster, M. (2016). Technology adoption under time-differentiated market-based instruments for pollution control. *Energy Economics, 60*, 23–34.
Crawford, J., & French, W. (2008). A low-carbon future: Spatial planning's role in enhancing technological innovation in the built environment. *Energy Policy, 36*(12), 4575–4579.
Demir, C., Cergibozan, R., & Ari, A. (2020). Environmental dimension of innovation: Time series evidence from Turkey. *Environment, Development and Sustainability, 22*(3), 2497–2516.
Dolfsma, W., & Seo, D. B. (2013). Government policy and technological innovation: A suggested typology. *Technovation, 33*(6–7), 173–179.
Du, K., Li, P., & Yan, Z. (2019). Do green technology innovations contribute to carbon dioxide emission reduction? Empirical evidence from patent data. *Technological Forecasting and Social Change, 146*, 297–303.
Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in US manufacturing industries: A dartboard approach. *Journal of Political Economy, 105*(5), 889–927.
Fang, C., Ma, H., Wang, Z., & Li, G. (2014). The sustainable development of innovative cities in China: Comprehensive assessment and future configuration. *Journal of Geographical Sciences, 24*, 1095–1114.
Freeman, C. (1987). *Technology policy and economic performance: Lessons from Japan*. Pinter.
Fu, Y., He, C., & Luo, L. (2021). Does the low-carbon city policy make a difference? Empirical evidence of the pilot scheme in China with DEA and PSM-DID. *Ecological Indicators, 122*, 107238.
Gallagher, K. S., Zhang, F., Orvis, R., Rissman, J., & Liu, Q. (2019). Assessing the policy gaps for achieving China's climate targets in the Paris Agreement. *Nature Communications, 10*(1), 1–10.
Ganda, F. (2019). The impact of innovation and technology investments on carbon emissions in selected organisation for economic Co-operation and development countries. *Journal of Cleaner Production, 217*, 469–483.
Gourdon, J., Monjon, S., & Poncet, S. (2016). Trade policy and industrial policy in China: What motivates public authorities to apply restrictions on exports? *China Economic Review, 40*, 105–120.
Grossman, G. M., & Helpman, E. (1994). Endogenous innovation in the theory of growth. *Journal of Economic Perspectives, 8*(1), 23–44.
Haarstad, H., & Wathne, M. W. (2019). Are smart city projects catalyzing urban energy sustainability? *Energy Policy, 129*, 918–925.
Heinelt, H., & Terizakis, G. (2021). Innovative cities: How to explain differences between cities? An answer based on an interpretive approach. *Urban Research and Practice, 14*(5), 487–501.
Hospers, G. J. (2008). Governance in innovative cities and the importance of branding. *Innovation, 10*(2–3), 224–234.
Hu, J., Pan, X., & Huang, Q. (2020). Quantity or quality? The impacts of environmental regulation on firms' innovation-Quasi-natural experiment based on China's carbon emissions trading pilot. *Technological Forecasting and Social Change, 158*, 120122.
Huang, C., Zhang, X., & Liu, K. (2021). Effects of human capital structural evolution on carbon emissions intensity in China: A dual perspective of spatial heterogeneity and nonlinear linkages. *Renewable and Sustainable Energy Reviews, 135*, 110258.
Irandoust, M. (2016). The renewable energy-growth nexus with carbon emissions and technological innovation: Evidence from the Nordic countries. *Ecological Indicators, 69*, 118–125.
Isiksal, A. Z. (2021). Testing the effect of sustainable energy and military expenses on environmental degradation: Evidence from the states with the highest military expenses. *Environmental Science and Pollution Research, 28*(16), 20487–20498.
Isiksal, A. Z. (2022). The decline in carbon intensity: The role of financial expansion and hydro-energy. *Environmental Science and Pollution Research, 29*, 16460–16471.
Isiksal, A. Z., Samour, A., & Resatoglu, N. G. (2019). Testing the impact of real interest rate, income, and energy consumption on Turkey's CO_2 emissions. *Environmental Science and Pollution Research, 26*(20), 20219–20231.
Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics, 108*(3), 577–598.
Javorcik, S. B. (2004). Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. *American Economic Review, 94*(3), 605–627.
Jiang, H., Jiang, P., Wang, D., & Wu, J. (2021). Can smart city construction facilitate green total factor productivity? A quasi-natural experiment based on China's pilot smart city. *Sustainable Cities and Society, 69*, 102809.
Joof, F., & Isiksal, A. Z. (2021). Do human capital and export diversification decline or augment CO_2 emissions? Empirical evidence from the MINT countries. *Journal of Environmental Accounting and Management, 9*(2), 111–125.

Kaufmann, R. K., Gopal, S., Tang, X., Raciti, S. M., Lyons, P. E., Geron, N., & Craig, F. (2013). Revisiting the weather effect on energy consumption: Implications for the impact of climate change. *Energy Policy, 62*, 1377–1384.

Kou, Z. L., & Liu, X. Y. (2017). *FIND Report on City and Industrial Innovation in China*. Fudan Institute of Industrial Development.

Lawson, C., & Lorenz, E. (1999). Collective learning, tacit knowledge and regional innovative capacity. *Regional Studies, 33*(4), 305–317.

Li, X., & Xu, H. (2020). Effect of local government decision-making competition on carbon emissions: Evidence from China's three urban agglomerations. *Business Strategy and the Environment, 29*(6), 2418–2431.

Li, X., Zhao, C., & Huang, M. (2022). Reassessing the effect of low-carbon city policy in china: new evidence from the nighttime light data. *Frontiers in Energy Research, 9*, 798448.

Li, Y., Zhang, J., Yang, X., Wang, W., Wu, H., Ran, Q., & Luo, R. (2021). The impact of innovative city construction on ecological efficiency: A quasi-natural experiment from China. *Sustainable Production and Consumption, 28*, 1724–1735.

Liao, L., Zhao, C., Li, X., & Qin, J. (2021). Towards low carbon development: The role of forest city constructions in China. *Ecological Indicators, 131*, 108199.

Liu, J., Li, M., & Ding, Y. (2021). Econometric analysis of the impact of the urban population size on carbon dioxide (CO_2) emissions in China. *Environment, Development and Sustainability, 23*(12), 18186–18203.

Liu, X., & Buck, T. (2007). Innovation performance and channels for international technology spillovers: Evidence from Chinese high-tech industries. *Research Policy, 36*(3), 355–366.

Lozano, S., & Gutiérrez, E. (2008). Non-parametric frontier approach to modelling the relationships among population, GDP, energy consumption and CO_2 emissions. *Ecological Economics, 66*, 687–699.

Lv, K., Cheng, Y., & Wang, Y. (2021). Does regional innovation system efficiency facilitate energy-related carbon dioxide intensity reduction in China? *Environment, Development and Sustainability, 23*(1), 789–813.

Lv, T., Hu, H., Xie, H., Zhang, X., Wang, L., & Shen, X. (2022). An empirical relationship between urbanization and carbon emissions in an ecological civilization demonstration area of China based on the STIRPAT model. *Environment, Development and Sustainability,*. https://doi.org/10.1007/s10668-022-02144-6

Mariotti, S., Piscitello, L., & Elia, S. (2010). Spatial agglomeration of multinational enterprises: The role of information externalities and knowledge spillovers. *Journal of Economic Geography, 10*(4), 519–538.

Nowak, D. J., Crane, D. E., & Stevens, J. C. (2006). Air pollution removal by urban trees and shrubs in the United States. *Urban Forestry and Urban Greening, 4*(3–4), 115–123.

Ogunseitan, O. A. (2010). Green bridge over troubled media greening through IT: Information technology for environmental sustainability by Bill Tomlinson. *Journal of Industrial Ecology, 14*(5), 859–862.

Ortiz, J. C., Bozec, Y. M., Wolff, N. H., Doropoulos, C., & Mumby, P. J. (2014). Global disparity in the ecological benefits of reducing carbon emissions for coral reefs. *Nature Climate Change, 4*(12), 1090–1094.

Pata, U. K., & Caglar, A. E. (2021). Investigating the EKC hypothesis with renewable energy consumption, human capital, globalization and trade openness for China: Evidence from augmented ARDL approach with a structural break. *Energy, 216*, 119220.

Peng, T., & Deng, H. (2021). Research on the sustainable development process of low-carbon pilot cities: The case study of Guiyang, a low-carbon pilot city in south-west China. *Environment, Development and Sustainability, 23*(2), 2382–2403.

Qin, X., Wu, H., & Li, R. (2022). Digital finance and household carbon emissions in China. *China Economic Review, 76*, 101872.

Qiu, S., Wang, Z., & Liu, S. (2021). The policy outcomes of low-carbon city construction on urban green development: Evidence from a quasi-natural experiment conducted in China. *Sustainable Cities and Society, 66*, 102699.

Rexhäuser, S., & Rammer, C. (2014). Environmental innovations and firm profitability: Unmasking the Porter hypothesis. *Environmental and Resource Economics, 57*(1), 145–167.

Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic Perspectives, 8*(1), 3–22.

Samour, A., Isiksal, A. Z., & Türsoy, T. (2021). The effect of the transmission impact of the US interest rate on Turkey's CO_2 emissions: Evidence from bootstrap ARDL. *Indonesian Journal of Sustainability Accounting and Management*. https://doi.org/10.32456/ijsam.v5i2.302

Shahbaz, M., Raghutla, C., Song, M., Zameer, H., & Jiao, Z. (2020). Public-private partnerships investment in energy as new determinant of CO_2 emissions: The role of technological innovations in China. *Energy Economics, 86*, 104664.

Sheraz, M., Deyi, X., Ahmed, J., Ullah, S., & Ullah, A. (2021). Moderating the effect of globalization on financial development, energy consumption, human capital, and carbon emissions: Evidence from G20 countries. *Environmental Science and Pollution Research, 28*(26), 35126–35144.

Tan, J. (2001). Innovation and risk-taking in a transitional economy. *Journal of Business Venturing, 16*(4), 359–376.

Tawiah, V. K., Zakari, A., & Khan, I. (2021). The environmental footprint of China-Africa engagement: An analysis of the effect of China-Africa partnership on carbon emissions. *Science of the Total Environment, 756*, 143603.

Thio, E., Tan, M., Li, L., Salman, M., Long, X., Sun, H., & Zhu, B. (2022). The estimation of influencing factors for carbon emissions based on EKC hypothesis and STIRPAT model: Evidence from top 10 countries. *Environment, Development and Sustainability, 24*(9), 11226–11259.

Van Hoek, R. I. (2002). Case studies of greening the automotive supply chain through technology and operations. *International Journal of Technology Management, 23*(1–3), 89–89.

Wang, K., Wu, M., Sun, Y., Shi, X., Sun, A., & Zhang, P. (2019). Resource abundance, industrial structure, and regional carbon emissions efficiency in China. *Resources Policy, 60*, 203–214.

Wang, K., Zhao, B., Ding, L. L., & Miao, Z. (2021). Government intervention, market development, and pollution emission efficiency: Evidence from China. *Science of the Total Environment, 757*, 143738.

Xu, B., & Lin, B. (2017). Does the high-tech industry consistently reduce CO_2 emissions? Results from nonparametric additive regression model. *Environmental Impact Assessment Review, 63*, 44–58.

Xu, H., & Li, X. (2022). Effect mechanism of Chinese-style decentralization on regional carbon emissions and policy improvement: Evidence from China's 12 urban agglomerations. *Environment, Development and Sustainability,*. https://doi.org/10.1007/s10668-021-02063-y

Yang, H., Gan, T., Liang, W., & Liao, X. (2022). Can policies aimed at reducing carbon dioxide emissions help mitigate haze pollution? An empirical analysis of the emissions trading system. *Environment, Development and Sustainability, 24*, 1959–1980.

Yigitcanlar, T., & Kamruzzaman, M. (2018). Does smart city policy lead to sustainability of cities? *Land Use Policy, 73*, 49–58.

Yu, Y., Chen, X., & Zhang, N. (2022). Innovation and energy productivity: An empirical study of the innovative city pilot policy in China. *Technological Forecasting and Social Change, 176*, 121430.

Yu, Y., & Zhang, N. (2021). Low-carbon city pilot and carbon emission efficiency: Quasi-experimental evidence from China. *Energy Economics, 96*, 105125.

Yu, Z., Liu, S., Wang, J., Wei, X., Schuler, J., Sun, P., Harper, R., & Zegre, N. (2019). Natural forests exhibit higher carbon sequestration and lower water consumption than planted forests in China. *Global Change Biology, 25*(1), 68–77.

Zeng, S., & Wang, M. (2022). Theoretical and empirical analyses on the factors affecting carbon emissions: case of Zhejiang Province, China. *Environment, Development and Sustainability,*. https://doi.org/10.1007/s10668-022-02148-2

Zhang, Y. J., Peng, Y. L., Ma, C. Q., & Shen, B. (2017). Can environmental innovation facilitate carbon emissions reduction? Evidence from China. *Energy Policy, 100*, 18–28.

Zhao, C., Deng, M., & Cao, X. (2021). Does haze pollution damage urban innovation? Empirical evidence from China. *Environmental Science and Pollution Research, 28*(13), 16334–16349.

Zheng, G., Barbieri, E., Tommaso, M. R. D., & Zhang, L. (2016). Economic and technology development zones, SEZ and other special economic enclaves: Zooming in on the Chinese case. *China Economic Review, 38*(2), 238–249.

Zheng, S., Sun, W., Wu, J., & Kahn, M. E. (2017). The birth of edge cities in China: Measuring the effects of industrial parks policy. *Journal of Urban Economics, 100*, 80–103.

Zhou, Y., & Li, S. (2021). Can the innovative-city-pilot policy promote urban innovation? An empirical analysis from China. *Journal of Urban Affairs*. https://doi.org/10.1080/07352166.2021.1969243

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Check for updates

Received: 8 October 2023 | Revised: 12 December 2023 | Accepted: 19 February 2024

DOI: 10.1111/rode.13096

REGULAR ARTICLE

WILEY

Beyond the scenery: Understanding the impact of rural tourism development on household consumption in China

Binhui Wei[1] | **Chunkai Zhao**[1] | **Mingzhong Luo**[1] | **Zhimin Jia**[2] | **Hailong Xia**[3]

[1]College of Economics and Management, South China Agricultural University, Guangzhou, Guangdong, China

[2]School of Finance, Shanghai University of Finance and Economics, Shanghai, China

[3]Research Center For Rural Economy, Ministry of Agriculture and Rural Affairs, Beijing, China

Correspondence
Chunkai Zhao, College of Economics and Management, South China Agricultural University, No. 483, Wushan Street, Tianhe District, Guangzhou, Guangdong, China.
Email: ckz_zhao@163.com

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 72304102, 72373043; Philosophy and Social Science Program Discipline Co-construction Project of Guangdong Province, Grant/Award Number: GD23XYJ68; National Social Science Foundation of China, Grant/Award Number: 23&ZD112

Abstract

The positive effects of rural tourism are widely recognized, yet its influence on local household consumption has been insufficiently explored. Our study fills this gap by investigating the relationship between rural tourism development and household consumption, using the Leisure Agriculture and Rural Tourism Demonstration County (LARTDC) policy as a quasi-natural experiment framework. Employing a staggered difference-in-differences (DID) method, we found that rural tourism development contributes to a 4.91% increase in per capita household consumption. The robustness of our main results is confirmed using advanced techniques such as Bacon decomposition and DID_M estimations. We attributed this positive impact to the rise in nonfarm employment and infrastructure development. Our heterogeneity analysis further reveals the inclusive nature of rural tourism, showing its significant role in improving household consumption patterns, diminishing consumption inequality, and enhancing welfare, especially in disadvantaged rural households. Further analysis confirms that greater LARTDC policy exposure may lead to higher levels of household consumption. Moreover, we revealed that rural tourism increased the

freedom of employment, especially for the elderly and women; however, it did not squeeze out leisure time or hurt the quality of life and welfare. Theoretically, this study contributes to the discourse on tourism's role in fostering economic growth and social equity, addressing the patterns of high savings and low consumption in rural China. Practically, our findings suggest that post-COVID-19 recovery strategies in developing countries may emphasize rural tourism development, aiming to stimulate household spending and achieve balanced growth.

KEYWORDS

household consumption, infrastructure construction, nonfarm employment, rural tourism development, staggered difference-in-differences (DID)

JEL CLASSIFICATION

D12, O21, R58, Z32

1 | INTRODUCTION

In the late 19th century, rural tourism grew in Europe and North America as people sought to escape the pollution and fast-paced life of industrial cities. Rural tourism is now described as a major new force and is currently the fastest-growing sector of the tourism industry in terms of providing jobs, raising incomes, and increasing economic diversity (Banerjee et al., 2018; C. Liu et al., 2020; Zhang et al., 2022). More importantly, in some countries, rural tourism has been considered to be an important tool for economic growth and fiscal independence (Boukas & Ziakas, 2016). Thus, it is necessary to take some measures to promote the development of rural tourism. The European Union (EU), for example, is considered to be the most successful region in developing rural tourism and has given great support to European rural tourism in terms of funding, policy, education, and training through a customized EU Fifth Framework Agreement (Saxena et al., 2007).

China is vast in territory and abundant in natural landscapes, agricultural resources, and traditional folklore, which offers great potential for rural tourism. With the promotion of China's Rural Revitalization Strategy, rural tourism is seen as an effective path to promote rural economic development and enhance the welfare of rural residents (Deng et al., 2022; X. S. Yang & Xu, 2022). Since 2010, the Ministry of Agriculture and Rural Affairs of China and the National Tourism Administration have been implementing the Leisure Agriculture and Rural Tourism Demonstration County (LARTDC) policy to develop special rural tourism across the country. There were many successful cases of revitalizing traditional villages through rural tourism, attracting global attention (P. Chen, Clarke, et al., 2022). Some were even listed as world cultural heritage sites, such as Hongcun Village and Kaiping Diaolou.

The widespread interest in rural tourism, known for its beneficial impacts like economic growth, improved infrastructure, and enhanced quality of life (e.g., Deng et al., 2022; Hwang &

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Lee, 2015; X. S. Yang & Xu, 2022), has notably overlooked its influence on local consumption. This gap in research is crucial, especially when considering the direct causal relationship between the growth of rural tourism and household consumption in these areas. Moreover, although existing studies explored the impact of tourism development on local consumption, the conclusions are divergent. On the one hand, the "tourism multiplier effect" suggests a positive influence on local economies. Tourism development boosts residents' earnings due to increased tourist expenditure, which increases consumption levels and promotes diversification of consumption behaviors (Deng et al., 2022; Schubert & Schamel, 2021). On the other hand, another strand of literature pointed out that tourism development may lead to an increase in local prices, which is detrimental to local consumption growth (Gunter et al., 2022; Rivera et al., 2016; Rosalina et al., 2021).

This conflict may stem from differences in research perspectives and potential endogeneity issues. We try to reconcile the conflict in two ways. On one hand, from the research perspective, we chose the particular rural tourism development in China. At the practical level, Chinese rural households tend to save rather than consume due to a range of factors such as precautionary motives, liquidity constraints, cultural backgrounds, and consumption habits (e.g., Curtis et al., 2015; Vatsa et al., 2023; C. Zhao et al., 2022). The nonfarm employment opportunities brought about by tourism development are likely to ease the mobility constraints of rural households and stimulate consumption. Moreover, rural tourism development may contribute to infrastructure construction in rural areas (Huang et al., 2023; Hwang & Lee, 2015), providing greater consumer convenience. At the theoretical level, consumption is an important reflection of household welfare and provides an accurate measure of household living standards (H. Chen, Ding, et al., 2022; Meyer & Sullivan, 2023). Also, consumption is one of the main drivers of economic growth. Thus, scholars worked to explore a range of reasons that have contributed to the growth of household consumption, especially in developing countries (Emran & Hou, 2013). Thus, it is important to actively raise the level of rural consumption and promote the consumption welfare of rural households.

On the other hand, to mitigate endogenous issues and explore the causal relationship between rural tourism and household consumption, we considered LARTDC policy implementations as a quasi-natural experiment and regarded it as a proxy for rural tourism (Huang et al., 2023). Using a large representative Chinese micro database, the National Fixed-sited Rural Survey (NFRS), we employed the staggered DID method to effectively identify the causal effect of rural tourism development on household consumption growth in China (Aaronson & Mazumder, 2011; Khandker, Bakht, & Koolwal, 2009). Our estimates indicated that LARTDC policy implementations significantly raised household consumption levels. Also, the results showed that rural tourism had inclusive characteristics, contributing more to the consumption welfare of disadvantaged groups and improving the consumption structure and consumption inequality by boosting household enjoyment consumption.

In comparison to previous research, the contributions of this study were evident in the following aspects. First, we aimed to explore the impact of rural tourism on household consumption, further complementing the literature on both the benefits of rural tourism development (Curtis et al., 2015; Daugstad, 2008), and the measures to promote rural consumption in China (Curtis et al., 2015; Emran & Hou, 2013; Vatsa et al., 2023; C. Zhao et al., 2022). Also, in the debate about tourism development and local household consumption, we were more inclined to support the positive effects of tourism development on rural residents' consumption by using the staggered DID as our identification method. Additionally, we further considered spillover effects based on the fundamental parallel trend and some frontier validation methods in

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

staggered DID estimates, such as the Bacon decomposition and DID_M estimations, were also employed, ensuring the reliability of our results.

Second, few scholars have paid attention to the potential mechanism of rural tourism development on household consumption. Based on the Chinese settings, we demonstrated and tested two mechanisms: nonfarm employment and infrastructure construction. These findings helped us understand more precisely how rural tourism could play a role in enhancing rural household welfare. Therefore, we had several policy implications regarding rural tourism constructions and encouraged other developing countries to actively promote rural tourism to improve the consumption welfare of rural residents.

Third, as an additional understanding, we provided evidence of the positive impact of rural tourism on the consumption of socially disadvantaged groups to verify its inclusive characteristics. Results of the heterogeneity analysis showed a positive effect of rural tourism development both in improving consumption structure and alleviating consumption inequality, as well as in promoting the consumption of some typically disadvantaged rural households. Additionally, we found that rural tourism development did not significantly reduce residents' leisure leading to a decline in welfare. Thus, this study complemented related studies on rural tourism to enhance social welfare and promote social equity (Daugstad, 2008). Moreover, our research provided insights into how localities can take advantage of their strengths to develop rural tourism for rural revitalization and social equity.

2 | LITERATURE REVIEW

2.1 | Public policies and rural tourism development

Rural tourism plays a pivotal role in stimulating economic growth and diversification in rural areas (Deng et al., 2022). It supports the provision of public goods, including infrastructure development, environmental conservation, and cultural heritage preservation (Mahadevan et al., 2017). Instead, the private sector often lacks the necessary commitment to these public endeavors (Bramwell & Lane, 2014). However, the impacts of rural tourism can be both positive, such as bolstering local economies and culture, and negative, like causing environmental degradation (Fleischer & Tchetchik, 2005; Streimikiene et al., 2021). Appropriate government interventions can balance these effects, amplifying benefits and curbing detriments (Hwang & Lee, 2015). More importantly, the interplay between tourism, agriculture, and diversification indicates the need for policy-driven incentives to tackle challenges like information asymmetry, market failures, and public goods provision (Chi & Han, 2021; J. Yang et al., 2021).

Rural tourism policies primarily aim to both preserve and harness the natural and cultural riches of rural regions while diversifying local economies. These policies prioritize infrastructure investment, environmental and cultural protection, community engagement, and fostering public–private partnerships (Andergassen & Candela, 2013; Mwesiumo et al., 2022). Effective policy execution demands collaboration between the government and other stakeholders (Bramwell & Lane, 2014). Therefore, the rural tourism policy is a comprehensive industrial policy centered on rural tourism development, which not only includes the creation of rural tourism landscapes but also covers the development and extension of the entire industrial chain (Hwang & Lee, 2015). Thus, it is essential to rely on public policies to develop rural tourism.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.2 | The benefits of rural tourism development

In an era marked by rapid globalization and increasing emphasis on environmental, cultural, and social values, rural tourism has ascended in importance in both academic discussions and practical applications. First, rural tourism contributes to the improvement of rural infrastructure, natural environment, and landscape, which helps to enhance village cohesion, promote agricultural transformation, and local cultural promotion (Deng et al., 2022; Mahadevan et al., 2017). With the development of rural tourism, improvements in roads, telecommunications, and other public amenities benefit not just visitors but enhance the quality of life for locals (Ma et al., 2023; Zhong et al., 2020). Moreover, rural tourism not only absorbs the inflow of capital but also triggers a domino effect, bringing economic opportunities to all sectors. This effect rejuvenates adjacent domains, from local handicrafts to transportation and gastronomy (Rosalina et al., 2021). Thus, given these macro-economic benefits, previous literature considered that rural tourism had the potential to promote economic growth, reduce poverty, and narrow the rural–urban divide (P. Chen, Clarke, et al., 2022; X. S. Yang & Xu, 2022).

Second, scholars also focused on the micro economic benefits of rural tourism. Rural tourism development creates more employment opportunities and enhances local farmers' income (e.g., Dang et al., 2024; Fleischer & Tchetchik, 2005; Hwang & Lee, 2015; L. Zhao & Xia, 2020). For instance, Fleischer and Tchetchik (2005) found that local farmers and budding entrepreneurs gained additional and stable sources of income through rural tourism. Such an income effect may further translate into improved living conditions, human capital accumulation, and increased social welfare (Dang et al., 2024; L. Zhao & Xia, 2020). Moreover, rural tourism has the potential to transform the lifestyles of rural residents, expand social networks, change living environments, and increase happiness (Wang et al., 2021; X. S. Yang & Xu, 2022). For example, rural tourism serves as a spotlight for local craftsmen, artists, and cultural aficionados, promoting social interaction and strengthening social network (Bramwell & Lane, 2014; Pérez-Olmos & Aguilar-Rivera, 2021).

Third, as a component of rural tourism, the multiple benefits of agritourism have also received attention from scholars. Agritourism enables local communities to diversify their income sources by engaging in activities like farm stays, agricultural workshops, and direct sales of farm products (Arroyo et al., 2013). This not only boosts the local economy but also encourages sustainable farming practices and cultural exchange. They are fostering a connection between urban and rural communities (Qu et al., 2022). Furthermore, agritourism has a significant impact on preserving local traditions and crafts. By participating in agritourism, tourists may experience traditional farming techniques, local cuisine, and artisanal crafts, which helps to preserve these cultural practices (Pérez-Olmos & Aguilar-Rivera, 2021). Additionally, agritourism also promotes environmental stewardship. Farms engaged in agritourism are incentivized to maintain picturesque landscapes and biodiversity, which not only attract tourists but also contribute to ecological preservation (Abadi & Khakzand, 2022; Hatan et al., 2021).

2.3 | Tourism development and consumption

Regarding the relationship between tourism and consumption, previous literature focused on tourist consumption behavior, demonstrating intricate relationships (Han, 2021; Soltani et al., 2021). For example, the Tourism Area Life Cycle model outlines the varying stages of a

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

tourist destination's life cycle, emphasizing how tourist spending shifts with each stage (Aldao et al., 2021). As destinations evolve, advancements in infrastructure and amenities often parallel increases in tourist outlay. Delving into the psyche of tourists, they may differ in motivations, information-gathering tendencies, and cultural orientations, leading to varied expenditure patterns (C. Zhao et al., 2022). Furthermore, subsequent literature has further emphasized the impact of several economic factors on tourist spending, such as infrastructure development, exchange rates, pricing strategies, and market segmentation (Katircioglu, 2014).

Another strand of literature explored the effect of tourism development on local household consumption, but there is no consensus. On one hand, tourism may modulate residents' consumption tendencies (Schubert & Schamel, 2021). This interplay unfolds across various fronts, encompassing economic, sociocultural, and environmental realms. In terms of the economic effect, the "tourism multiplier effect" illustrates that tourist expenditures may translate into locals' earnings, which further promotes household consumption (Deng et al., 2022). From a sociocultural perspective, tourism development may reshape local consumption. Exposure to different cultures induces a convergence of consumer behaviors, especially in dining, clothing, and entertainment (Cole & Davies, 2014; Durbarry, 2008), which creates more opportunities for local households to spend money. At the environmental level, some prior research was interested in analyzing the energy consumption effects of tourism development (Bano et al., 2022; Katircioglu, 2014), further confirming the role of tourism development in improving environmental quality.

On the other hand, tourism development may have adverse effects on local consumption, especially in rural areas. Several studies highlighted that tourism development often leads to inflationary pressures, affecting the consumption patterns of residents (Gunter et al., 2022). For instance, increased demand for goods and services in tourist areas often results in higher prices, thereby reducing the purchasing power of local consumers (Brida et al., 2011). In addition, the influx of tourists often shifts the focus of local businesses toward tourist-oriented products and services, which may not align with the needs of the local population, thereby limiting their consumption choices (Telfer & Sharpley, 2015). This shift can also lead to a dependence on tourism, making local economies vulnerable to fluctuations in tourist arrivals and spending (Rosalina et al., 2021), which may also have negative impacts on local consumption.

3 | THEORETICAL MECHANISMS AND HYPOTHESIS DEVELOPMENT

In this section, we discussed the potential mechanisms by which rural tourism development affected household consumption from a theoretical perspective. We considered that these two channels included nonfarm employment and infrastructure construction.

3.1 | Non-farm employment

Tourism is a typical labor-intensive service industry. To meet the various needs of tourists in tourism activities, such as food, accommodation, transportation, entertainment, and shopping, the tourism sector requires a large number of service workers (Alrwajfah et al., 2020). Even though digital tools and robots are becoming more widely used in tourist attractions, some specific tourism jobs still require a large workforce (Seyitoğlu & Ivanov, 2021). In addition to the

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

jobs created directly by the tourism sector, tourism economics have spillover effects on employment (Nunkoo & Ramkissoon, 2011). For example, the industrial chain formed around a scenic spot may lead to the development of surrounding industries and agriculture (Mahadevan et al., 2017), which contributes to the employment rate. According to information published by the World Tourism Organization, for every direct job added to the tourism sector, society gains 5–7 indirect jobs.[1]

Rural tourism, with its strong links to agriculture and low employment threshold, can create more service industry jobs that match rural human capital, thus absorbing surplus rural labor (Hwang & Lee, 2015). Generally speaking, rural tourism requires a lower overall quality of staff and has a greater demand for female labor (T. Li et al., 2021). The employment needs of some disadvantaged groups also can be met accordingly in tourism sectors or travel companies (Wilson et al., 2001). Even women and elderly people left behind who are less competitive for employment can, through certain labor skills training, take on the work of tour guides, receptions, and cleaning required for rural tourism (T. Li et al., 2021; Su, 2011). Thus, as the implementation of the LARTDC policy brought about rural tourism development in China, it might create more off-farm employment opportunities for local rural laborers to increase their income and consumption.

For rural households, non-farm employment is a strategy with a comparative advantage that is more conducive to enhancing household consumption welfare (Ma et al., 2018). As a typical weak industry, agriculture is vulnerable to external factors such as climate, technology, and markets (Brooks & Donovan, 2020), leading to uncertainty in agricultural income and determining rural households to have more precautionary savings and lower consumption levels (Lee & Sawada, 2010). Instead, non-farm employment helps overcome these difficulties to boost household consumption. On one hand, non-farm employment can increase the diversity and stability of rural household income and reduce the uncertainty of future income (Šegota et al., 2017). On the other hand, when rural residents view such stable income expectations as lasting, the crowding-out effect of the precautionary saving motive on consumption will be greatly weakened (De Brauw & Rozelle, 2008). Therefore, we formulated the first hypothesis:

> **Hypothesis 1.** Rural tourism development may improve rural household consumption by promoting non-farm employment.

3.2 | Infrastructure construction

Infrastructure such as transportation, telecommunications, and electricity facilities are the basic elements of rural tourism activities (Khadaroo & Seetanah, 2007; C. Liu et al., 2020). Tourists not only want to enjoy the beautiful rural scenery but also expect to have perfect tourism facilities and services (Buckley et al., 2021; Su, 2011). The condition of rural tourism infrastructure has a direct impact on tourism quality and visitor perception. However, the infrastructure in rural areas is relatively poor, especially in developing countries (T. Li et al., 2021; Zhang et al., 2022). In particular, some special rural tourism sites rely on special geographical conditions such as deserts, mountains, and islands, where the original infrastructure is deficient.

In the process of promoting rural tourism construction, actively improving infrastructure construction is a top priority (Nunkoo & Ramkissoon, 2011), which requires joint support from the government, financial institutions, and social funds (Huang et al., 2023; C. Liu et al., 2020). The LARTDC constructions in China are very typical cases. In addition to direct financial

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

allocations to build infrastructure, the government has also liberalized market access for private capital and actively attracted a variety of capital into related agricultural and service industries to complement infrastructure improvements (Huang et al., 2023). Similarly, according to the data released by the Department for Promotion of Industry and Internal Trade in India, from 2000 to 2019, rural tourism and related industries attracted around US$ 14.42 billion of foreign investments, leading to further improvements in the infrastructure of rural tourists' attractions (Pandya, 2020). With the construction of infrastructure and the integration of rural tourism resources, rural areas may achieve spatial optimization. Rural production and living spaces are effectively differentiated, which will further enhance the habitat and consumption convenience of rural residents.

There is no doubt that well-developed infrastructure is one of the keys to boosting household consumption growth (Khandker, Koolwal, & Samad, 2009). Consumption activities are accompanied by the transfer of physical goods or services from production to consumers, during which there are transaction costs created by barriers in time and space. Infrastructure construction helps to improve the consumption environment and relax consumption constraints; also, it can indirectly increase household consumption by raising income through capital accumulation and economic growth (Aker & Mbiti, 2010). Especially in developing countries and backward areas, relatively poor infrastructure may constrain consumption levels (Khandker, Bakht, & Koolwal, 2009). If this constraint is improved, rural households may be able to shift their consumption behavior and have a higher propensity to consume (Donaldson, 2018). Rural tourism development has further improved the rural infrastructure and created the preconditions for promoting the growth of rural household consumption. Taken together, we developed the second hypothesis:

> **Hypothesis 2.** The second channel for rural tourism to stimulate household consumption is infrastructure construction.

4 | RESEARCH DESIGN

4.1 | Policy background

In recent years, the Chinese government has attached great importance to rural tourism development and introduced a series of policies and measures one after another. Among them, the LARTDC was one of the most extensive and strongest support policies implemented by the government to accelerate the development of rural tourism. In 2010, the Ministry of Agriculture and Rural Affairs of China and the National Tourism Administration jointly issued the Opinions on the Creation of National Leisure Agriculture and Rural Tourism Demonstration Counties, aiming to promote the expansion of agricultural functions, the restructuring of the rural economy, and the employment of farmers. Through the development of rural tourism, LARTDCs are expected to boost the tourism economy and promote consumption, becoming a new growth point for rural areas. From 2010 to 2017,[2] 389 LARTDCs were approved to be established in eight batches. In Figures 1 and 2, we depicted their temporal and spatial distribution. We could see that these LARTDCs were distributed in 31 provincial-level administrative regions in mainland China, and there was not much difference between the distribution in the east, central, and west. Moreover, we found that they were mainly concentrated in the Yangtze River basin, Beijing-Tianjin-Hebei region, three northeastern provinces, and Xinjiang and Gansu in the northwest.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

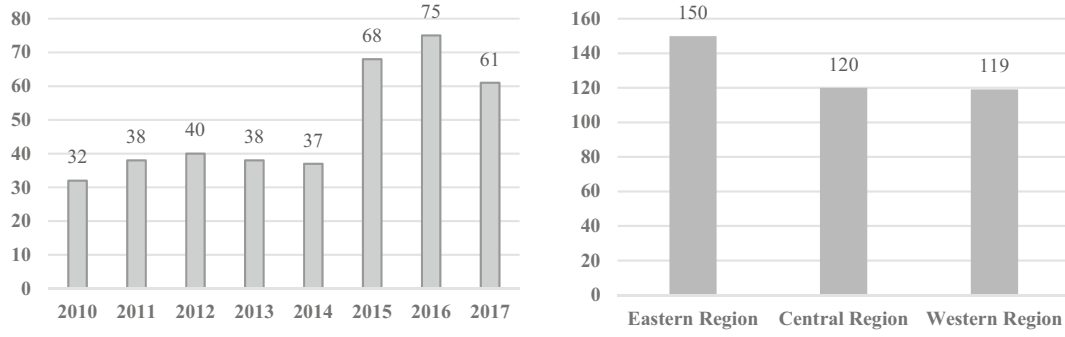


FIGURE 1 Temporal and regional distribution of the LARTDC. LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.

The creation of LARTDC was an important initiative to promote the development of rural tourism in China. The Chinese government hopes that the approved LARTDCs will take this opportunity to further strengthen standardized construction and drive the rapid development of rural tourism. Based on the LARTDC, rural tourism is expected to become one of the local pillar industries, contributing to rural revitalization and common prosperity. Moreover, a series of measures were implemented in the LARTDCs, such as strengthening planning guidance, providing more financial and fiscal support by expanding access to financing, and enhancing promotion and training to create local special tourism.

The establishment of LARTDCs is subject to specific criteria. First, rural tourism spots must be spread across over 30% of the county's townships and should have already developed significant rural tourism industries or clusters. Second, these sites should showcase unique regional, folk, and cultural attributes, offering attractive experience programs, dining, and services. They should also integrate local agricultural practices, animal husbandry, and agri-food processing to develop leisure agriculture and rural tourism products. Third, rural tourism needs to be one of the dominant industries in the county, annually attracting over 1 million visitors, with a workforce comprising over 60% local farmers. In addition, more than 30% of the employees should hold tourism-related vocational qualifications, or over 60% should have received specialized training.[3] These criteria ensure that the rural tourism industry in LARTDCs represents a leading standard.

The development of rural tourism faces several challenges, including land scarcity, limited funding, a shortage of skilled workers, underdeveloped infrastructure, and insufficient public services. To overcome these obstacles, local governments in LARTDCs have adopted a range of supportive policies, including land reforms, financing policies, talent development measures, and marketing strategies. Previous studies have confirmed that these measures have significantly enhanced rural tourism land use, facilitated the financing of tourism enterprises, and promoted the inflow of tourism talent as well as the creation of well-known tourism brands (Huang et al., 2023; Zhong et al., 2020). Thus, using the LARTDC policy as a benchmark for assessing rural tourism development is reasonable and scientifically valid.

4.2 | Data

In this study, we used the rural household-level data sourced from the NFRS, a comprehensive annual survey conducted by China's Ministry of Agriculture and Rural Affairs since 1986. Each

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

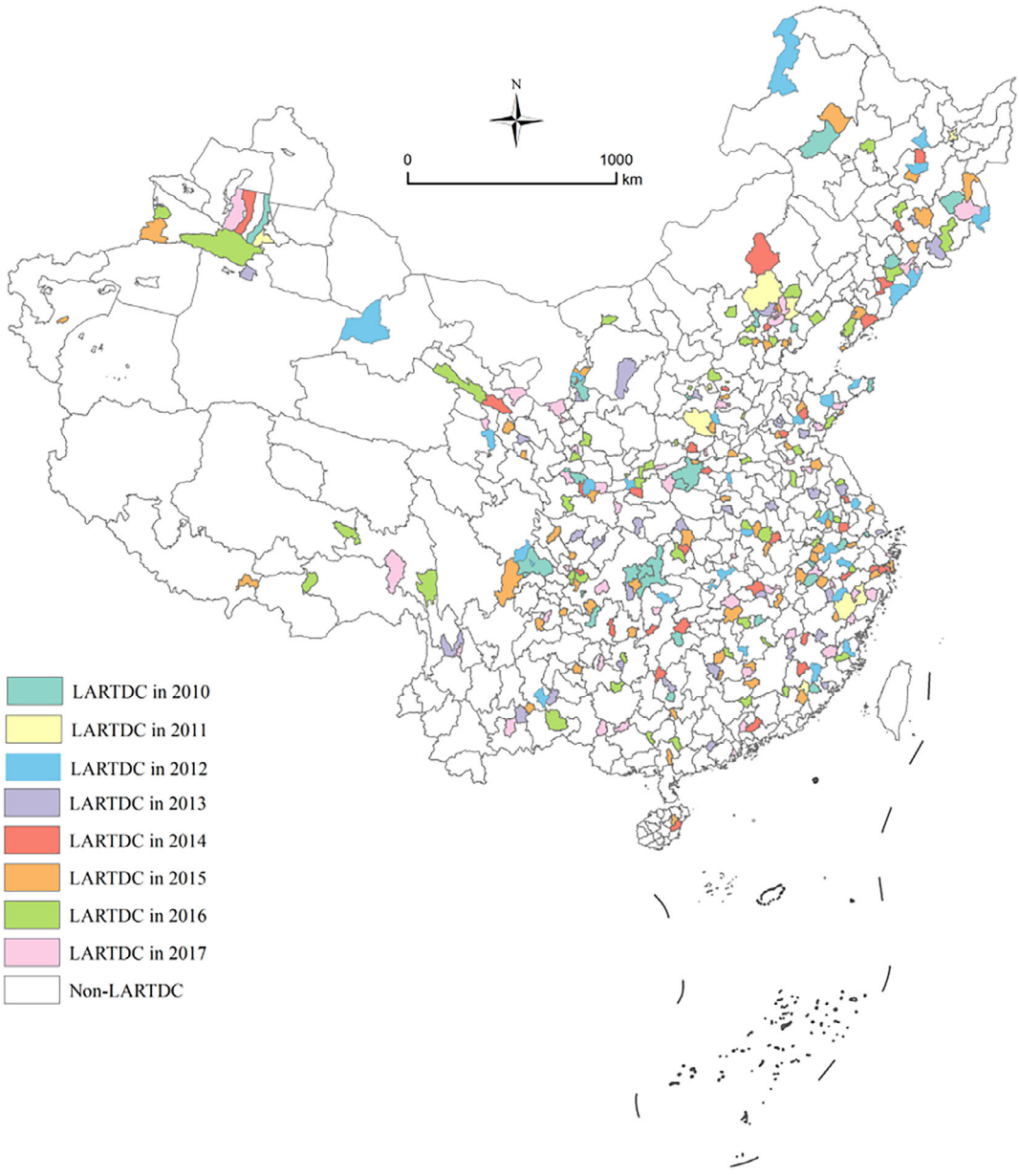


FIGURE 2 The spatial distribution of the LARTDC in batches. LARTDC, Leisure Agriculture and Rural Tourism Demonstration County. [Colour figure can be viewed at wileyonlinelibrary.com]

wave of the NFRS database encompasses over 20,000 rural households across more than 300 administrative villages in 31 provinces, municipalities, and autonomous regions of mainland China, offering a nationally representative and authoritative perspective on rural life. The survey employs a meticulous stratified sampling method, selecting villages based on distinct characteristics like topography, climate, and geographic location, thereby ensuring a diverse and representative sample. Subsequently, a random selection of households within these villages is conducted for detailed surveying. The NFRS is rich in detail, offering comprehensive

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

insights into the production, consumption, employment, and living conditions of rural households, thereby providing a nationally representative and authoritative dataset (Huang, 2010). In terms of household consumption, the NFRS delivers detailed accounts spanning food, housing, clothing, cultural services, travel, and durable goods.

In constructing a comprehensive multidimensional tracking dataset, we have undertaken a detailed process of collating and integrating data from various levels: householder, household, and village. To ensure the highest level of fidelity and consistency, our data handling process involved meticulous cleaning and integration procedures. First, we merged data from householders, households, and villages and excluded households lacking complete information about the head of household or village, allowing for a more comprehensive view of each entity's economic situation over time. Second, we excluded samples with missing key variables. For example, 2713 observations that did not report information on rural household consumption were removed.

In addition, the list of LARTDCs was from the official documents of the Ministry of Agriculture and Rural Affairs of China.[4] Based on the specific pilot county names, we could match them with the NFRS data. We selected nine waves of survey data from 2009 to 2017, as the first batch of LARTDCs was approved in 2010. Moreover, some macro data at the county level were obtained from the China County Statistical Yearbook. Taken together, these data provided a reliable source to effectively identify the effect of rural tourism on household consumption in China.

4.3 | Variables

Our core explained variable is *Household consumption*. As mentioned above, the NFRS provides detailed information on rural household consumption. Consistent with some previous literature (Emran & Hou, 2013), we used household consumption per capita to exclude the effect of household size on consumption. Meanwhile, to reduce the heteroskedasticity, we employed a logarithmic form. Moreover, different types of consumption were defined to fully examine the impact of rural tourism on household consumption (X. Li & Luo, 2021). Specifically, they were subsistence consumption (expenditures on food, clothing, and housing), developmental consumption (expenditures on living services, cultural services, and transportation and communication), and enjoyment consumption (expenditures on travel and durable goods).

The explanatory variable in this paper is rural tourism, using the LARTCD policy as a proxy. More specifically, *LARTCD* is a dummy variable reflecting whether the LARTDC policy was implemented in county c in year t. In the staggered DID framework, it was assigned a value of 1 if the policy was implemented and 0 otherwise.

To accurately assess the effect of rural tourism on household consumption, following some existing literature (e.g., Gradín & Wu, 2020; Lugauer et al., 2019; C. Zhao et al., 2022), we selected four levels of control variables. First, at the householder level, considering the potential nonlinear effects, we chose *Age* and *Age squared* (C. Zhao et al., 2022). The remaining control variables include *Gender*, *Hukou*, *Education*, *Self-rated health*, and *Party*, as we considered the positive effects of the female householder, urban *hukou*,[5] good education and health, and party membership on household consumption (Lugauer et al., 2019). Second, in terms of household characteristics, we controlled for two important economic variables, *Total income* and *Current deposit*, because of the possible influence of household income and deposit on household consumption (Gradín & Wu, 2020; C. Zhao et al., 2022). Third, village characteristics include

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Medical care ratio, *Village income*, *Village Terrain*, *Village officials*, and *Village economic regions*, as these village-level variables may also affect household consumption (C. Zhao et al., 2022). Last, for county macro variables, *Industrial structure*, *Investment*, *Government*, and *GDP* are added to the control variable set (Bano et al., 2022; P. Chen, Clarke, et al., 2022).

4.4 | Descriptive statistics

The detailed definitions and descriptive statistics of the above variables are shown in Table A1. We could see that in the full sample, the mean value of log rural household consumption per capita is 8.93, and 7.99% of counties have implemented the LARTDC policy. From the perspective of householders, the average age was 53.03, more than 89% were male, the average years of education was 7, and 15.15% of the observations were party members. These characteristics are similar to those described in some previous studies (Emran & Hou, 2013; C. Zhao et al., 2022). Also, we found that at the village level, there are an average of 6.07 village officials. In addition, we compared the LARTDCs and non-LARTDCs through the *T*-tests. Rural households in the LARTDC have a higher consumption level, tentatively suggesting that rural tourism development is likely to stimulate household consumption in China. Furthermore, for most of the control variables, the *T*-tests were significant, suggesting that the selection of these variables was reasonable and helped to reduce potential endogeneity.

4.5 | Empirical strategy

The LARTDC policy amounts to a quasi-natural experiment, which provides a unique empirical opportunity to analyze the causal effect of rural tourism on rural household consumption by using the DID method. However, the traditional DID method can only assess the policy effect at a single point in time. Therefore, by drawing on the approach in prior studies (Gertler et al., 2016; Kline & Moretti, 2014), we constructed a staggered DID model with the following form:

$$Consumption_{i,c,t} = \alpha_0 + \alpha_1 LARTDC_{c,t} + \alpha_n X_{i,c,t} + \mu_i + \delta_t + \varepsilon_{i,c,t}. \quad (1)$$

In Equation (1), $Consumption_{i,c,t}$ denotes the consumption status of rural household i in county c in year t. The core explanatory variable $LARTDC_{c,t}$ indicates whether county c has been selected as a LARTDC in year t. The coefficient α_1 of $LARTDC_{c,t}$ is the focus of attention, which measures the net effect of LARTDC policy on rural household consumption. $X_{i,c,t}$ represents a set of control variables, as mentioned above, including a series of householder, household, village, and county characteristics. μ_i and δ_t are households and year fixed effects, respectively, which allows us to adopt a reduced form of the DID model. $\varepsilon_{i,c,t}$ is the random disturbance term.

We have two main concerns under the staggered DID framework. An essential requirement for using a DID model is the parallel trend assumption. The parallel trend assumption means that before LARTDC policy implementations, rural household consumption in the treatment group and the control group had the same trend without significant differences (Khandker, Koolwal, & Samad, 2009; Kline & Moretti, 2014). To test this assumption, we used event analysis to perform parallel trend tests. The following equation is adopted:

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$Consumption_{i,c,t} = \beta_0 + \sum_{k=-6}^{k=+6} \beta_k \times D_{c,t_o+k} + \beta_n X_{i,c,t} + \mu_i + \delta_t + \varepsilon_{i,c,t}. \quad (2)$$

In Equation (2), D_{c,t_o+k} is a series of dummy variables indicating the kth year before or after LARTDC implementations in county c. k greater than 0 means after the policy is implemented; conversely, k less than 0 means before the policy is implemented.[6]

Another core identifying assumption of the DID estimates is the stable unit treatment values assumption (SUTVA), which is the assumption that whether different units are exposed to policy shocks is independent and that the exposure of the treatment group to policy shocks does not affect the outcomes of any other units (Joseph, 2022). That is, spillover effects are absent or negligible to guarantee that the treatment and control groups are relatively clean in the DID cases (Lu et al., 2019). In reality, however, a considerable number of policy shocks have certain externalities.[7] According to previous studies (Joseph, 2022; Lu et al., 2019), we used two methods to assess the possible policy spillover effects. The first was to test whether rural household consumption in other counties adjacent to the LARTDC also increased and the model was as follows:

$$Consumption_{i,c,t} = \theta_0 + \theta_1 LARTDC_Near_{c,t} + \theta_n X_{i,c,t} + \mu_i + \delta_t + \varepsilon_{i,c,t}. \quad (3)$$

In Equation (3), $LARTDC_Near_{c,t}$ is the interaction term between $LARTDC_{c,t}$ and $Near_c$, indicating the spillover effect. $Near_c$ is used to distinguish the treatment group from the control group, assigned a value of 1 if county c is a neighboring county around a LARTDC, and 0 otherwise.[8] The second approach was to exclude non-LARTDCs around the LARTDC from the control group (Kline & Moretti, 2014). We calculated the spatial distance between the non-LARTDCs and the nearest LARTDC, and then set different exclusion rules based on the spatial distance window.

5 | RESULTS

5.1 | The effect of the LARTDC policy on rural household consumption

Table 1 reports the baseline results, estimating the effect of LARTDC policy implementations on rural household consumption. Column (1) showed that the coefficient of *LARTDC* was significantly positive at the 1% level when only household and year-fixed effects were controlled for. That was, the LARTDC policy helped to promote rural household consumption. To further validate this relationship, a series of control variables were added in turn to Columns (2)–(4), and we found the results remain unchanged. Taking our preferred specification Column (4) as an example, results showed that on average, LARTDC policy implementations resulted in a 4.91% increase in rural household consumption per capita.

Our research, aligning with prior studies (Dang et al., 2024; Hwang & Lee, 2015; L. Zhao & Xia, 2020), underscores the positive influence of rural tourism on household welfare. Other literature focusing on rural household consumption in China reveals that broadband infrastructure and urban–rural medical insurance policies can elevate total household consumption by 17% and 14%, respectively (H. Chen, Ding, et al., 2022; Wan et al., 2021). While the impact

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Baseline results: The effect of the LARTDC policy on rural household consumption.

	Household consumption			
	(1)	(2)	(3)	(4)
LARTDC	0.0499*** (0.0163)	0.0513*** (0.0161)	0.0498*** (0.0161)	0.0491*** (0.0161)
Age		−0.0018*** (0.0005)	−0.0011** (0.0005)	−0.0011** (0.0005)
Age squared		0.0001*** (0.00004)	0.0001* (0.00004)	0.0001* (0.00005)
Gender		0.0135 (0.0162)	−0.0092 (0.0165)	−0.0090 (0.0165)
Hukou		−0.0462* (0.0257)	−0.0224 (0.0264)	−0.0226 (0.0264)
Education		0.0051*** (0.0015)	0.0054* (0.0028)	0.0053* (0.0028)
Self-rated health		0.0004 (0.0061)	0.0131** (0.0061)	0.0132** (0.0061)
Party		−0.0315** (0.0151)	−0.0384** (0.0156)	−0.0384** (0.0156)
Total income			0.3264*** (0.0091)	0.3262*** (0.0091)
Current deposit			−0.0198*** (0.0010)	−0.0199*** (0.0010)
Medical care ratio			−0.0130 (0.0121)	−0.0132 (0.0121)
Village income			0.0018 (0.0072)	0.0010 (0.0072)
Village terrain			0.0275 (0.0184)	0.0274 (0.0185)
Village officials			−0.0081*** (0.0024)	−0.0080*** (0.0025)
Village economic regions			−0.0496 (0.0434)	−0.0507 (0.0438)
Industrial structure				0.0022* (0.0013)
Government				−0.1141 (0.1343)
GDP				−0.0014 (0.0260)
Household fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Constant	8.9320*** (0.0027)	9.0347*** (0.0405)	5.7518*** (0.1305)	5.7728*** (0.3739)
R^2	0.6393	0.6392	0.6736	0.6737
N	48,237	47,165	41,986	41,986

Note: Robust standard errors are reported in parentheses.

Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.

***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

coefficients in our study are relatively modest, they are nonetheless significant, especially considering the long-standing consumption challenges faced by rural Chinese households. Overall, our findings demonstrate that rural tourism development, as observed in the quasi-natural experiment of LARTDC policy implementations, significantly contributed to boosting household consumption.

In Column (4), the variable *Age squared* shows a significantly positive coefficient. This indicates a U-shaped relationship between the age of householders and rural household consumption. The positive coefficient of education underscores the beneficial impact of education on consumption, aligning with findings by C. Zhao et al. (2022) and Cheng (2021). Interestingly, our estimates reveal higher consumption in households with health-compromised householders. This trend, noted in previous studies (C. Li et al., 2016; Zhu et al., 2021), might reflect the financial strain of high medical expenses, a significant burden for rural households (C. Zhao et al., 2022; Zhu et al., 2021).

Regarding other economic variables, higher income and lower savings in rural households correlate with increased consumption. This observation is consistent with the broader literature (Cheng, 2021; Zhu et al., 2021). However, there is a notable negative association between *Village officials* and rural household consumption, possibly because an increase in village cadres might lead to a shift in expenditure towards public goods rather than direct financial support or initiatives that increase disposable income, which does not immediately translate into higher consumption levels for rural households. Finally, the positive coefficients on *Industrial structure* indicate that its transformation and upgrading bolster household income and facilitate consumption (X. J. Liu et al., 2022).

5.2 | Parallel trend tests and spillover effects

As mentioned earlier, satisfying parallel trends and the absence of spillover effects are key prerequisites for a DID design. Based on Equation (2), we first tested the parallel trend assumption, and the results are shown in Figure 3.[9] When we assumed that the LARTDC policy occurred in advance (k is negative), the coefficients were all insignificant at the 5% level (zero is included in the confidence interval). These results indicated that in the treatment group, rural household consumption levels did not show significant differences from those in the control group before the onset of the LARTDC policy. Thus, the parallel trend assumption was satisfied and the LARTDC policy could be treated as a quasi-natural experiment, which was suitable for estimation using a staggered DID model in this study. Moreover, the coefficients were significantly positive in the year of LARTDC policy implementation and became more stable as the years progressed. Our estimates suggested that rural tourism might make a long-term contribution to household consumption in China.

According to Equation (3), we then tested the spillover effects of the LARTDC policy. It was clear that the coefficient of the spillover interaction term *LARTDC_Near* was insignificant in Column (1) of Table 2, indicating that the LARTDC policy did not boost rural household consumption in neighboring counties. In addition, results excluding neighboring cities from the control group were reported in the last three columns (Lu et al., 2019). We found that the coefficients on *LARTDC* were still significantly positive. All in all, the above results provided evidence that the spillover effect of the LARTDC policy was negligible in this study, confirming the appropriateness of employing the DID model.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

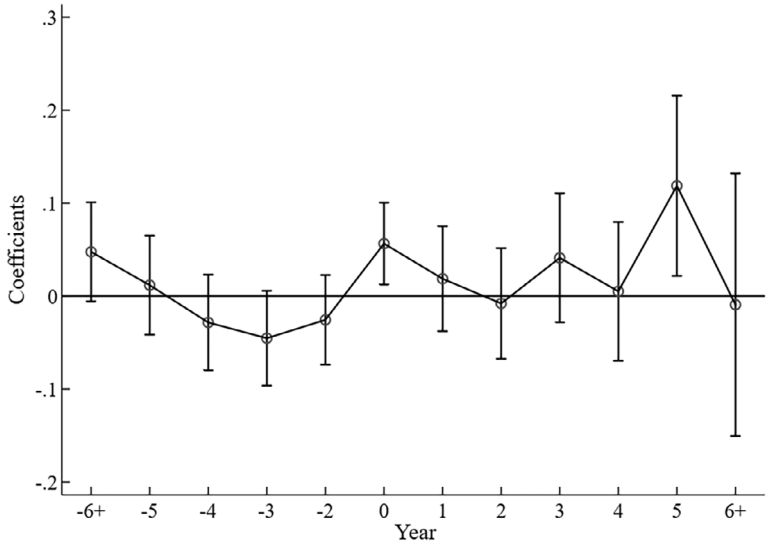


FIGURE 3 Parallel trend tests.

TABLE 2 Spillover effect tests.

	Household consumption			
	(1)	(2)	(3)	(4)
	Full sample	>60 km	>80 km	>100 km
LARTDC_Near	0.0045 (0.0194)			
LARTDC		0.0385** (0.0172)	0.0407** (0.0176)	0.0388** (0.0176)
Control variables	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
R^2	0.6736	0.6744	0.6108	0.6141
N	41,986	40,692	38,591	36,978

Note: Robust standard errors are reported in parentheses.
Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

5.3 | Robustness tests

5.3.1 | Propensity score matching

In the framework of the staggered DID approach, we tested the positive impact of the LARTDC policy on rural household consumption. However, there was still a possibility that the LARTDC was not randomly selected, which could lead to selection bias. Thus, we applied the propensity score matching method to reduce this bias before using the DID method (Gertler

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

et al., 2016). More specifically, we selected a series of county characteristics as covariates and then ran DID regressions based on the matched observations according to propensity scores. We used kernel matching, radius matching, and nearest neighbor matching, and these results were presented sequentially in Columns (1) to (3) of Table A2. We could see that under different matching methods, coefficients on *LARTDC* were all significantly positive, and the magnitude remained largely consistent with the baseline results in Table 1. Therefore, the conclusion that rural tourism was conducive to promoting household consumption in China remained unchanged after overcoming potential selective bias.

5.3.2 | Eliminating disturbing policies

Some similar or related policies emerge cross-over or in parallel in response to a particular objective. If other policies that are more correlated with rural household consumption are implemented during the sample period, there may be some policy overlap effects, which would have an impact on our results. In line with previous research (Chen et al., 2020; C. Zhao et al., 2023), we constructed some variables that measured these disturbing policies and included baseline regressions to observe whether coefficients on *LARTDC* changed significantly. The results are represented in Table A3.

First, to further promote consumption upgrading, China implemented several policies directly related to consumption, such as the Information Consumption Demonstration City (ICDC) and the Cultural Consumption Demonstration City (CCDC). If these related consumption policies at the prefecture city level were not ruled out,[10] it should lead to an overestimation. In Column (1), we found that the coefficients on *LARTDC* remained significantly positive when we added variables for these two policies.

Second, in China, digital technology is spreading rapidly to the socioeconomic sector, making it easy to purchase a wide range of products and services even in rural China (C. Zhao et al., 2022). The positive effect of digital development on household consumption patterns has received widespread attention (Ma et al., 2020; B. Yang et al., 2023). Two policies related to digital development that may affect the consumption of rural households were identified (Hong et al., 2023): Smart City Constructions (SCC) and Broadband China (BC). The results in Column (2) suggested that after controlling the effect of these digital development policies, LARTDC policy implementations still improved rural household consumption.

Third, in recent years, several studies considered the stimulating effect of e-commerce development on the consumption of rural residents. E-commerce development helps shorten the distance between farmers and markets and promotes household online consumption (M. Liu et al., 2021). To accelerate the development of rural e-commerce, China implemented the National Rural E-commerce Demonstration County (NREDC) program in 2014. We excluded the effect of NREDC program in Column (3), finding that the estimates of *LARTDC* were not challenged.

Fourth, the positive effect of entrepreneurial behavior on household consumption structure and consumption characteristics was highlighted in the literature (Priem et al., 2012). Thus, we planned to further eliminate policies related to rural entrepreneurship. Since 2016, the Chinese government has been carrying out the Back-to-Home Entrepreneurship Demonstration County (BEDC) policy in 341 counties in three batches, attracting more migrant labor to start businesses in their hometowns. As before, in Column (4), we controlled the BEDC policy, finding a

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

significant and positive *LARTDC* coefficient, which is consistent with the results in prior columns.

5.3.3 | Bacon decomposition

In a causal interpretation of the two-way fixed effects (TWFE) DID estimates, treatment effects that are constant over time are required (Goodman-Bacon, 2021). Specifically, it is necessary to satisfy the assumption that none of the causal effects of the pretreatment group are time-varying after the treatment (De Chaisemartin & d'Haultfoeuille, 2020). To test the accuracy of our DID results, we drew on the Bacon decomposition proposed by Goodman-Bacon (2021) as a robustness test. Results for Bacon decomposition are represented in Figure A1 and Table A4. Treatment groups were divided into four types of sub-treatment groups at different points in time, with a never-treated group. Table A4 shows that 93.7% of the net effect of LARTDC policy implementations on rural household consumption came from the average treatment effect of the "Treatment vs. Never treated". Also, the weight of the main group ("Later Treatment vs. Early Control") that caused the staggered DID bias was only 1%. These results indicated that the baseline analysis with the never-treated group as the control group explained the vast majority of the LARTDC policy effects. Our staggered DID estimates were not heavily influenced by heterogeneity treatment effects causing bias, suggesting that our baseline results are quite convincing.

5.3.4 | Placebo tests

To avoid the influence of other unobservable contingencies on our results, placebo tests were carried out in robustness checks (C. Zhao et al., 2023). Specifically, we randomly chose the LARTDC policy implementation list from the full sample and set it as a pseudo-treatment group. However, since this pseudo-treatment group was not selected and was randomly generated, the constructed placebo test policy would not have a significant effect on rural household consumption in this case. The results in Figure A2 show us that coefficients based on the randomly generated treatment group, after 1000 regressions, were overwhelmingly insignificant and their mean was close to 0, indicating that this pseudo-LARTDC policy did not promote rural household consumption. Placebo test results implied that our estimates were not obtained by chance, indirectly confirming the reliability of our estimates.

5.3.5 | Alternative specifications

We conducted some robustness checks by using alternative specifications. First, as discussed earlier, the selection of the LARTDCs was not random, although this did not affect the causal inference effect of the DID estimates. To avoid the potential impact of the nonrandomness, an interaction term between county benchmark factors and a linear trend over time was added to Equation (1), drawing on Lu et al. (2017). The county benchmark factors denote a set of dummy variables for county benchmark factors, including *Poverty-stricken county*, *Top 100 county*, *Taobao village*, and *Eastern region*,[11] which make it possible to capture the effect of differences in pre-existing intrinsic characteristics between counties on rural household consumption. In

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Column (1) of Table A5, the coefficient on *LARTDC* remained significantly positive at the 5% level, indicating that our results were robust after taking into account the possible effects of inherent inter-regional differences.

Second, we considered the expected effects. The implementation of the LARTDC policy underwent a preparatory process, and local governments and rural residents might have different expectations for this policy. If some counties took early action in anticipation of the LARTDC policy, this could similarly lead to contamination of the control group before policy implementation. Thus, a dummy term for the year before implementing the LARTDC policy was added to Equation (1). As can be seen in Column (2) of Table A5, the estimate of *LARTDC* remained unchanged, suggesting that the expected effects of the LARTDC policy did not have shock on results in Table 1.

Third, to avoid the effect of extreme values on regression results, The pre- and post-5% observations were removed based on the variable *Household consumption*. Not surprisingly, the results in Column (3) of Table A5 show that the coefficient on *LARTDC* was positive and passed the significance test at the 1% level. This finding was consistent with the baseline estimates.

Fourth, we observed that some individuals live in rural areas but have urban *hukou* in Table 1. Although very common for a country with a large population like China, we excluded these observations to avoid the potential impact of such a heterogeneous sample. Results in Column (4) of Table A5 indicate that the LARTDC policy implementations still significantly raised household consumption levels after excluding these observations with an urban *hukou*, suggesting that our results were not sensitive to these special observations.

Last, using TWFE in estimating the dynamic effects of policy may result in coefficients for each period becoming difficult to interpret due to cross-contamination across periods (de Chaisemartin & d'Haultfoeuille, 2020). Based on this consideration, we used the DID_M method proposed by de Chaisemartin and d'Haultfoeuille (2020) to deal with this problem. In Column (5) of Table A5, we found that the estimate of $LARTDC_M$ was largely unchanged from the baseline results, suggesting that the TWFE model bias issue did not materially affect our DID estimates.

5.4 | Mechanism tests

As discussed in Section 3, at the theoretical level, we argued that rural tourism may boost household consumption by promoting nonfarm employment and infrastructure construction. In this subsection, we tested these two mechanisms through empirical analysis.

5.4.1 | Non-farm employment

To verify the mechanism of nonfarm employment, we selected some mechanism variables in terms of work type and non-farm income (Ba et al., 2021). *Work type* is a binary variable, set to 1 if the rural householder worked in a nonagricultural job and 0 otherwise. The dummy variable *Income sources* was assigned to 1 if the main source of household income is from non-farm sectors. Besides, we used *Arable land* as a proxy to capture nonfarm employment from the perspective of agricultural production conditions.[12]

Columns (1) to (3) of Table 3 present the effects of the LARTDC policy on non-farm employment. We observed that the LARTDC's impact was significantly positive in Column (2) but

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/roie.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Mechanism tests: nonfarm employment.

	(1)	(2)	(3)	(4)	(5)	(6)
	Work type	***Income sources***	***Arable land***	***Household consumption***		
LARTDC	−0.0127 (0.0087)	0.0195** (0.0084)	−0.1322*** (0.0381)	0.0382* (0.0196)	0.0242 (0.0177)	0.0491*** (0.0171)
LARTDC × *Work type*				0.0211 (0.0278)		
LARTDC × *Income sources*					0.0983*** (0.0332)	
LARTDC × *Arable land*						−0.0474*** (0.0079)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.7359	0.6515	0.7299	0.6738	0.6738	0.6903
N	44,440	44,440	35,607	41,986	41,986	34,098

Note: Robust standard errors are reported in parentheses.
Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

negative in Column (3). This suggests that rural tourism development fosters nonfarm income growth and decreases the per capita arable land in rural Chinese households. Specifically, compared to non-LARTDCs, LARTDC policy implementations increased the likelihood that rural households' main income came from nonfarming sources by 1.95 percentage points, and reduced the amount of cultivated land per household by 13.22%. These outcomes align with prior research (Buckley et al., 2021; Khadaroo & Seetanah, 2007), highlighting the beneficial role of tourism development in enhancing non-farm employment opportunities.

In the last three columns of Table 3, we constructed interaction terms between the mechanism variables and *LARTDC* to further validate the non-farm employment channel, as done in prior research (Chen et al., 2020; C. Zhao et al., 2022). We could see that coefficients of these interaction terms were also significant in Columns (5) and (6), suggesting that LARTDC policy implementations increased rural household consumption by boosting off-farm income and reducing the arable land area. Taken together, these DID estimates supported our theoretical Hypothesis 1 that rural tourism development stimulated household consumption by promoting non-farm employment in China.

5.4.2 | Infrastructure construction

Productive infrastructure construction mainly influences rural consumption by improving the rural production and investment environment and contributing to higher incomes for farmers, while livelihood infrastructure can directly influence the rural consumption environment (Crawford & Bell, 2012). As a result, we focused on the effect of LARTDC policy

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 Mechanism tests: Infrastructure construction.

	(1)	(2)	(3)	(4)	(5)	(6)
	Roadworks	*Internet usage*	*Lighting conditions*	*Household consumption*		
LARTDC	1.5177*** (0.5146)	0.0035 (0.0077)	0.0083** (0.0039)	0.0920 (0.0733)	0.0581*** (0.0174)	−0.0083 (0.0301)
LARTDC × *Roadworks*				0.1438** (0.0723)		
LARTDC × *Internet usage*					−0.0449 (0.0312)	
LARTDC × *Lighting conditions*						0.0654** (0.0286)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.5210	0.5689	0.9102	0.8186	0.6737	0.6737
N	4699	44,440	44,440	4450	41,986	41,986

Note: Robust standard errors are reported in parentheses.
Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

implementations on rural livelihood infrastructure construction. The basic elements of infrastructure were the systems of transport, communication, and residence facilities (Khadaroo & Seetanah, 2007). To test this mechanism, we have selected three mechanism variables, namely *Roadworks*, *Internet usage*, and *Lighting conditions*.[13]

Table 4 displays results indicating that the LARTDC policy significantly positively affected factors represented in Columns (1) and (3), pointing to improvements in local transport infrastructure and housing conditions. More specifically, LARTDC policy implementations contributed to a 151.77% increase in the proportion of paved roads within villages and boosted the likelihood of rural households using electric lighting by 0.83 percentage points. This is notably significant for enhancing rural infrastructure, thereby stimulating growth in rural household consumption (Zhong et al., 2020). Moreover, similar to previous analyses, we created interaction terms between mechanism variables and *LARTDC*, finding significantly positive coefficients in Columns (4) and (6). This provides evidence that rural tourism, through infrastructure development, can enhance household consumption in China. In conclusion, Hypothesis 2 of this study has been empirically tested and supported.

5.5 | Heterogeneity analysis

The positive role of tourism development in promoting economic growth, reducing poverty, and promoting social equity was widely recognized (e.g., C. Liu et al., 2020; Zhang et al., 2022). Whether rural tourism development had these benefits from the consumption level needs to be further explored. In this subsection, by heterogeneity analysis, we further tested these effects at

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

four aspects: consumption structure, consumption inequality, regions, and household characteristics.

5.5.1 | Consumption structure

The expansion of consumer demand encompasses not just quantitative growth but, more critically, the ongoing refinement of the consumption structure. This structure, as previously discussed, includes subsistence, developmental, and enjoyment consumption (X. Li & Luo, 2021; C. Zhao et al., 2022). The analysis presented in Table 5 reveals that the implementation of the LARTDC policy has significantly elevated household enjoyment consumption, whereas its influence on subsistence and developmental consumption remains marginal. In particular, rural households in areas with the LARTDC policy show a notable 15.86% increase in enjoyment-based consumption compared to their counterparts in counties without this policy.

A possible reason was that as the basic consumption needs of rural households, subsistence, and developmental consumption were characterized by a low elasticity of demand. With the improvement of living conditions in rural China, rural households' needs for food, clothing, shelter, transport, and some recreation have been largely met. Thus, the marginal boosting effect of rural tourism development on subsistence and developmental consumption was weak. In conclusion, our results implied that rural tourism development might expand the content of products and services for rural consumption and help to transform and upgrade the consumption structure.

5.5.2 | Consumption inequality

To obtain a deeper understanding of rural tourism development and its connection to household consumption welfare, it is necessary to examine the consumption inequality effect. Consumption inequality is often considered a more fundamental economic issue (Attanasio &

TABLE 5 The effect of rural tourism on household consumption structure.

	(1)	(2)	(3)
	Subsistence consumption	***Developmental consumption***	***Enjoyment consumption***
LARTDC	0.0083 (0.0187)	0.0349 (0.0255)	0.1586** (0.0762)
Control variables	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
R^2	0.5958	0.6083	0.4428
N	42,007	42,007	42,007

Note: Robust standard errors are reported in parentheses.
Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 Consumption inequality heterogeneity effects.

	(1)	(2)	(3)	(4)
	Total consumption inequality	***Subsistence consumption inequality***	***Developmental consumption inequality***	***Enjoyment consumption inequality***
LARTDC	−0.0078* (0.0042)	0.0071 (0.0044)	−0.0035 (0.0055)	−0.0140** (0.0065)
Control variables	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
R^2	0.6732	0.6954	0.6040	0.4426
N	41,790	41,805	41,803	41,604

Note: Robust standard errors are reported in parentheses.
Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Pistaferri, 2016), which reflects the true welfare gap between residents, as consumption is considered an important aspect of utility and welfare. A measure of consumption is more often preferred for the analysis of inequality (Meyer & Sullivan, 2023). Therefore, it is important to figure out whether rural tourism development played a role in reducing consumption inequality in China.

We used the Kakwani index to measure different types of rural consumption inequality (Guillaud et al., 2020; Kakwani, 1984), and the results are reported in Table 6. We found that the coefficient on *LARTDC* was significantly negative in Column (1), indicating that rural tourism development might alleviate consumption inequality in rural China. Looking at the different types of consumption inequality, we could see that LARTDC policy implementations mainly reduced the enjoyment consumption inequality. In short, our estimation results showed that rural tourism development narrowed the gap in consumption welfare levels between Chinese rural households, further improving social equity.

5.5.3 | Regional heterogeneity

Globally, the development of rural tourism is widely viewed as a key strategy for enhancing living standards in disadvantaged areas (Daugstad, 2008). This approach is particularly relevant in China, where there is a broad spectrum of regional economic development and varying consumption levels. Consequently, assessing the regional disparities in the impact of rural tourism on household consumption becomes essential. To this end, we categorized our sample into three groups based on the economic status of the villages within the counties. Analysis presented in Table 7, Columns (1) to (3),[14] shows that the LARTDC policy's impact was notably significant in the relatively undeveloped group. This indicates a more pronounced positive effect

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 7 Region heterogeneity effects.

	Household consumption				
	(1)	(2)	(3)	(4)	(5)
	Relatively developed	Indifferent	Relatively undeveloped	Moderately prosperous villages	Nonmoderately prosperous villages
LARTDC	−0.0253 (0.0296)	0.0266 (0.0230)	0.1689*** (0.0489)	−0.0033 (0.0256)	0.0710*** (0.0214)
Control variables	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.6784	0.6925	0.6868	0.6754	0.6790
N	13,491	20,716	6549	12,495	28,928

Note: Robust standard errors are reported in parentheses.
Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

of rural tourism development in elevating household consumption in economically lagging areas.

Furthermore, we bifurcated the sample into two additional groups based on the presence of moderately prosperous villages, as shown in Columns (4) and (5) of Table 7.[15] Here, the LARTDC policy's coefficients were positively significant only in the group comprising nonmoderately prosperous villages. In terms of quantifiable impact, LARTDC policy implementation led to a 16.89% increase in per capita consumption in relatively undeveloped districts and a 7.1% rise in nonmoderately prosperous villages. These findings underscore that rural tourism significantly enhances the consumption welfare of disadvantaged groups, showcasing its inclusive benefits in the Chinese context.

5.5.4 | Household heterogeneity

We finally carried out the heterogeneity effect analysis from the perspective of household characteristics. On the one hand, for householders, we chose *Old age*, *Relatively highly educated*, and *Relatively healthier*, as the elderly and those with low education or poor health were considered vulnerable groups (C. Zhao et al., 2022).[16] On the other hand, two variables, namely the *Officers* and *Dependency ratio*, were used to represent household characteristics.[17] Clearly, those who do not serve as officers and have high family dependency ratios are considered vulnerable households. The empirical results are shown in Table 8. We found that the estimated coefficient on interaction term was significantly negative in Column (3), suggesting that LARTDC policy implementations were more likely to stimulate rural household consumption with a relatively poorly healthy householder. In addition, in Columns (4) and (5), we found that LARTDC policy

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 8 Household characteristics heterogeneity effects.

	Household consumption				
	(1)	(2)	(3)	(4)	(5)
LARTDC	0.0533*** (0.0184)	0.0730** (0.0321)	0.1431** (0.0576)	0.0533*** (0.0163)	−0.0092 (0.0208)
LARTDC × *Old age*	−0.0138 (0.0306)				
LARTDC × *Relatively highly educated*		−0.0303 (0.0355)			
LARTDC × *Relatively better health*			−0.1009* (0.0593)		
LARTDC × *Officers*				−0.1234* (0.0704)	
LARTDC × *Dependency ratio*					0.0793*** (0.0240)
Control variables	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.6737	0.6737	0.6737	0.6737	0.6899
N	41,986	41,986	41,986	41,986	40,474

Note: Five variables *Old age*, *Relatively highly educated*, *Relatively better health*, *Officers*, and *Dependency ratio* are controlled from Column (1) to Column (5), respectively. Robust standard errors are reported in parentheses.
Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

implementations were more conducive to boosting consumption among rural households without officers or with a higher dependency ratio. These findings again suggested that rural tourism development played a more significant role in improving the welfare of disadvantaged households in rural China.

6 | FURTHER ANALYSIS

6.1 | Further evidence for rural tourism

In this study, we primarily focused on how rural households' consumption is impacted by the LARTDC policy. However, we did not reveal the effect of policy exposure. That is, in the context of LARTDC policy implementations, the response, and participation of rural households to this policy are inconsistent, and we were not certain that all households in the pilots benefited from rural tourism development (Aaronson & Mazumder, 2011; Lu et al., 2019). To address this gap, given the absence of detailed data at the individual or household level, our research shifts to a village-centric perspective. We identified four key proxy variables to represent the policy exposure, including *Village investments*, *Village enterprises*, *Village environment*, and *Village institutions*.[18]

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The four variables collectively offer a comprehensive view of rural tourism development (Gao & Wu, 2017; Singh et al., 2010; X. S. Yang & Xu, 2022). Village investments reflect the extent of infrastructure development crucial for tourism, while the number of Village enterprises indicates economic vibrancy and service diversity, essential for a thriving tourism sector (Andergassen & Candela, 2013; X. S. Yang & Xu, 2022). The Village environment, measuring the village environment through the use of firewood for fuel highlights concern for environmental sustainability and appeal to tourists (Singh et al., 2010). Lastly, Village institutions represent the organizational capacity for effective tourism management and community benefit, underscoring the importance of governance in fostering a sustainable and culturally rich tourism experience.

We constructed some interaction terms to capture the policy exposure, with the results displayed in Table A6. Notably, the interaction terms in Columns (3) and (4) are significant, offering evidence that the LARTDC policy positively influences rural household consumption. This underscores the policy's effectiveness in boosting the economic welfare of rural households. More importantly, our findings highlight the potential of the LARTDC policy as a catalyst for sustainable and culturally inclusive rural tourism development. However, it is essential to recognize that while our village-centric approach offers valuable insights, further research is needed to deepen our understanding of individual and household-level impacts. Future studies should aim to collect and analyze detailed data at these levels to uncover the nuanced effects of the LARTDC policy on different demographic groups within villages.

6.2 | Potentially competitive interpretations

The above analysis suggests that rural tourism development may contribute to household consumption welfare by creating more non-farm employment opportunities. Crucially, the rural tourism industry is characterized by low entry barriers and strong inclusivity, meaning that even those traditionally less competitive in the labor market, such as elderly residents and women remaining in rural areas, can find employment opportunities (T. Li et al., 2021; Su, 2011; Wilson et al., 2001).

Next, an interesting question is: if engaging more elderly people and women in work reduces their leisure time, what will be the overall impact on people's welfare? This points out that when developing rural tourism, residents trade off economic advantages against potential changes in their lifestyles to realize welfare gains. It emphasizes the importance of considering how increased employment opportunities in rural tourism might affect the leisure time and overall quality of life of the elderly and women in these communities. Due to the lack of data on individual leisure in the NFRS database, we used the China Labor Force Dynamic Survey (CLDS) data for this extended analysis.[19] Given that only the 2018 CLDS provided specific district and county codes, we were able to match these with the LARTDC list to obtain a year's worth of valid data (C. Zhao et al., 2023). To analyze the impact of the LARTDC policy on individual leisure, and considering data availability, we selected four proxy variables: *Weekly Working Hours*, *Annual Working Hours*, *Work Freedom*, and *Work Satisfaction*.[20]

The estimates from Table A7 reveal that the implementation of the LARTDC policy does not significantly increase residents' working hours; in fact, it appears to enhance work freedom, especially for the elderly. Overall, our analysis, leveraging the CLDS, has illuminated a critical aspect of the LARTDC policy's impact: its influence on the leisure time and quality of life of rural residents, particularly the elderly and women. The results reveal a reassuring trend; the

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

implementation of the LARTDC policy, contrary to the above concerns, does not significantly increase working hours. It appears to contribute positively to work freedom. This finding is pivotal as it challenges the traditional narrative of economic development often leading to reduced leisure and quality of life (Aguiar & Hurst, 2007; Stevenson & Wolfers, 2013). These findings suggest that rural tourism, characterized by its low entry barriers and inclusivity, can be a sustainable model that balances economic growth with social welfare.

7 | DISCUSSION AND CONCLUSION

7.1 | Main findings

In recent years, rural tourism has received attention from countries around the world as an important way to alleviate unemployment and population loss and to promote rural economic development and industrial transformation. Moreover, the important role of tourism in rural development has also been emphasized by scholars, such as promoting urban–rural exchanges, improving the quality of life, and increasing household income (Dang et al., 2024; Hwang & Lee, 2015; L. Zhao & Xia, 2020). However, there is a gap in the existing literature in assessing the causal relationship between rural tourism and household consumption, especially in developing countries such as China.

By treating the LARTDC policy implementations in batches since 2010 as a quasi-natural experiment, we employed the staggered DID method to explore the impact of rural tourism development on household consumption. We found that LARTDC policy implementations increased the per capita consumption of rural households by 4.91%, confirming the positive effect of rural tourism development in stimulating consumption in China. To ensure the reliability of the staggered DID method, we further considered the potential spillover effects based on the parallel trend tests and used the latest verification methods such as the Bacon decomposition and DID_M estimations. Besides, our baseline results were also supported by other robustness checks, such as using propensity score matching approaches, eliminating disturbing policies, and adopting alternative specifications.

Furthermore, our study identified two main factors through which rural tourism promotes growth in household consumption: expansion of non-farm employment and infrastructure improvements. Rural tourism advances not only increase non-farm job opportunities, boosting household income from these sources but also significantly enhance local infrastructure. This leads to a rise in rural household consumption. Specifically, rural tourism development shifted household income sources toward nonfarm activities. In counties implementing the LARTDC policy, there is a 1.95 percentage points higher likelihood of rural households engaging in nonfarm work and a 13.22% decrease in per capita cultivated land. In addition, rural tourism development substantially improved infrastructure, evidenced by a 151.77% increase in the percentage of paved village roads, and raised the likelihood of rural households using electric lighting by 0.83 percentage points.

Meanwhile, our heterogeneity effect analysis revealed that rural tourism in China plays an inclusive role. The implementation of the LARTDC policy not only enhances consumption patterns but also reduces consumption disparity among rural households. Results show a significant 15.86% rise in enjoyment-based consumption among households in areas with the LARTDC policy. Furthermore, rural tourism development has a notably positive impact on disadvantaged rural households, including those with less healthy members, higher dependency ratios, or those in economically weaker or nonmoderately prosperous villages. Specifically, the

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

LARTDC policy is linked to a 16.89% increase in per capita consumption in less-developed districts and a 7.1% increase in nonmoderately prosperous villages. The further analysis underscores the LARTDC policy's success in boosting household consumption. In addition, it highlights that rural tourism, with its low barriers to entry, substantially improves nonagricultural job prospects for the elderly and women. Crucially, this economic benefit does not compromise their leisure time or overall quality of life.

7.2 | Theoretical implications

Our study had some theoretical implications. First, we provided new evidence for the research field on the relationship between tourism development and consumption. In previous studies, the relationship between tourism and consumption is mainly focused on the tourist perspective rather than on the local population (Deng et al., 2022; Han, 2021). Thus, this research added to the class of literature on the local benefits of rural tourism (e.g., Huang et al., 2023; Hwang & Lee, 2015; T. Li et al., 2021; L. Zhao & Xia, 2020). More importantly, there is controversy over whether tourism development can effectively promote local consumption (Gunter et al., 2022; Rivera et al., 2016), and we reconcile this conflict by using a quasi-natural experiment with LARTDC in China. This study identified the positive effect of rural tourism development in enhancing the consumption welfare of local rural households.

Second, there is widespread academic concern about the high savings and low consumption dilemma in developing countries such as China, especially in rural areas. More explanation from the economic or psychological level in previous literature, including precautionary motives, consumption habits, and credit constraints (e.g., Curtis et al., 2015; Emran & Hou, 2013; Gradín & Wu, 2020; Khanal et al., 2019). We provided evidence that the development of new industrial models, such as rural tourism, may help stimulate the consumption dynamics of rural households and promote social equity in China. Thus, our findings further enriched this research area by giving a new theoretical solution from the perspective of rural tourism development.

Third, as an important method for policy evaluation and causal effect identification, the DID estimates have also received great attention in recent years. However, some recent studies showed that the presence of heterogeneous treatment effects may lead to biased staggered DID estimators (De Chaisemartin & d'Haultfoeuille, 2020; Goodman-Bacon, 2021). Therefore, in addition to the parallel trend tests, we further considered the SUTVA and TWFE model bias issues. By excluding potential spillover effects and using the Bacon decomposition and DID_M estimations, this study further guaranteed the reliability of the results, which provides methodological insights for subsequent research.

Last, by exploring the role of two main channels, non-farm employment, and infrastructure construction, we clarified the potential mechanisms of how rural tourism could enhance rural household consumption. Moreover, our study also related to research about the inclusive characteristics of public policies related to rural development represented by the LARTDC policy. All in all, this research provided some insights for subsequent studies aimed at exploring the relationship between tourism and consumption.

7.3 | Practical implications

Our findings provide some policy implications. For China, first, local governments should consider factors such as environmental carrying capacity, industrial structure, and resource

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

endowment of LARTDCs. In particular, rural tourism has a strong dependence on ecological resources and the environment (MacNeill & Wozniak, 2018). Especially after COVID-19, retaliatory trips may appear in Chinese LARTDCs, and the government should properly protect rural tourism resources to reveal sustainable tourism development. Second, the role of rural tourism in promoting household consumption is further exploited by stimulating nonfarm employment. It is necessary to leverage rural tourism development opportunities to promote employment and consumption (Andergassen & Candela, 2013; C. Liu et al., 2020). Combining local advantages, localized green development of rural tourism creates more nonagricultural jobs and increases training for local tourism workers. Third, taking action to upgrade rural public services should be necessary, such as improving relevant transport infrastructure to create convenient travel conditions, as our results highlighted infrastructure construction was an important way for rural tourism to promote household consumption growth.

In the post-COVID-19 era, similar to China, most developing countries also face the issues of low levels of rural household consumption and a more serious imbalance between urban and rural consumption. Concerns over the potential negative impacts of these issues have created a demand for comprehensive planning and research on the effects of rural tourism development on household consumption. Our findings provided evidence that rural tourism contributed to rural household consumption growth, consumption upgrading, and consumption welfare improvement of rural residents. Therefore, we believed that designing programs and strategies to develop rural tourism in some developing countries could also help them to alleviate the problem of insufficient consumption, promoting integrated urban and rural development and social equity.

7.4 | Limitations

It should be emphasized that there were still some limitations in this paper. First, the existing literature does not have a uniform measure of the level of rural tourism development (Su, 2011), and we used the LARTDC policy as a proxy variable for rural tourism in China (Huang et al., 2023; Zhong et al., 2020). Although using this policy shock helped us to identify the causal effect between rural tourism development and local household consumption growth, the LARTDCs were not the same across locations, such as the type, size, and financial investment of rural tourism. Since this information is not publicly available, it is difficult to further analyze more details and heterogeneous effects. Future research might try to use big data and online intermediary platforms to more accurately measure the level of rural tourism development in each region and provide more empirical evidence for rural tourism policy formulation.

Second, some important mechanisms we could not effectively identify due to NFRS database constraints. For example, changes in consumption habits may be a potential mechanism for rural tourism to promote household consumption (Khanal et al., 2019). Tourism development brings new industrial models and more consumption styles (Hwang & Lee, 2015; Wilson et al., 2001), which may change the consumption philosophy of local rural households. In addition, We lacked a direct discussion of household earnings, opting instead for non-farm employment mechanisms. The relationship between rising incomes and consumer behavior, especially in the context of rural tourism, presents a complex dynamic that warrants a more detailed exploration (Rivera et al., 2016). Therefore, some potential mechanisms are expected to be further explored and tested in the subsequent literature through more detailed survey data.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Third, in the further analysis, we selected proxy variables in an attempt to explore the exposure of different rural households to the LARTDC policy. After all, not all rural households in LARTDCs reflect on rural tourism development. Limited by the database, we cannot directly observe households' tourism economic behaviors, such as engaging in the tourism industry. Therefore, subsequent studies could design a questionnaire related to household and individual tourism-related behaviors to complement the policy effects of LARTDCs in detail.

ACKNOWLEDGMENTS

The authors acknowledge the support from the National Natural Science Foundation of China, Grant/Award Number: 72373043, 72304102; National Social Science Foundation of China, Grand/Award Number: 23&ZD112; Philosophy and Social Science Program Discipline Co-construction Project of Guangdong Province, Grant/Award Number: GD23XYJ68.

CONFLICT OF INTEREST STATEMENT

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Binhui Wei https://orcid.org/0009-0003-5064-834X
Chunkai Zhao https://orcid.org/0000-0003-1918-9598
Mingzhong Luo https://orcid.org/0009-0000-8429-9502

ENDNOTES

[1] Data source: https://www.gmw.cn/01gmrb/2005-10/10/content_313975.htm

[2] To date, the eighth batch of 2017 is the last batch.

[3] Data source: https://www.huoqiu.gov.cn/public/6618481/35256805.html.

[4] Data source: http://www.moa.gov.cn/nybgb/2015/yi/201711/t20171129_5922689.htm

[5] Although we focused on rural households, the NFRS conducts surveys based on place of residence rather than hukou type. Therefore, we could not completely exclude urban householders in the full sample. In the later analysis, we excluded these observations for robustness checks.

[6] The data are from 2009 to 2017, which covered the 8 years before and 7 years after the policy implementation. That is, for the first batch of LARTDC in 2010, the policy has been implemented for 7 years in 2017 and k is marked as +7. For the last batch of LARTDC in 2017, there is an 8-year gap with a k record of −8 relative to 2009. Given the small observations of the treatment group 2 years before and after policy implementations, to avoid full covariance (since we controlled for household fixed effects), we pooled the observations for 2 years before and after LARTDC policy implementations to period −6 (−6) and period 6 (+6).

[7] For example, due to the self-selectivity, if the treatment area is better supported by the policy, individuals from the control area may spontaneously migrate from the control area to the treatment area (Lu et al., 2019), implying that the non-policy target area is also affected by the intervention. Therefore, the existence of spillover effects from intervention policies is a potential threat that should be considered in the empirical study using the DID estimates.

[8] The term "neighboring counties" refers to non-LARTDC that share a common border with the LARTDC.

[9] This means that to avoid full covariance, we removed the −1 period.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[10] Since these two policies were implemented at the prefecture city level, a pilot city means that all lower counties implemented the policy.

[11] These four variables are all dummy variables that are set to 1 if a village meets the corresponding criteria, and 0 otherwise. For instance, if the village is a Taobao village, the variable *Taobao village* is assigned a value of 1.

[12] This variable is expressed in terms of the household arable land area per capita.

[13] *Roadworks* is represented by the proportion of hardened roads in the village. *Internet use* is a dummy variable, which is set to 1 if the household has Internet and 0 otherwise; *Lighting conditions* is assigned a value of 1 if the household has electric lighting and 0 otherwise.

[14] The economic level of villages is divided into five categories in NFRS database: upper, upper-middle, middle, lower-middle and lower. To make the distinction easier, we combined upper and upper-middle as relatively developed, middle as indifferent, and lower-middle and lower as relatively undeveloped.

[15] A village can only be defined as moderately prosperous in China if it meets certain criteria, such as annual income per capita, housing area per capita, and so on. The NFRS database provides information on whether a village is a moderately prosperous village or not in the sample.

[16] *Old age*, *Relatively highly educated*, and *Relatively healthier* are all dummies, set to 1 if the householder is over 60 years old, received formal education or self-rated health as excellent, good, or indifferent, respectively, and 0 otherwise.

[17] *Officers* is a dummy that equals 1 if a family member is an officer (such as the village cadres), and 0 otherwise. The dependency ratio is expressed as the ratio of the number of non-working members of the household to the number of working people.

[18] *Village investments* denotes the productive investments made by the village collective. *Village enterprises* refers to the number of businesses within the village. *Village environment* is represented by the number of households using firewood as fuel. *Village institutions* are indicated by the number of personnel in village collective cooperative economic organization management.

[19] The CLDS database, a pioneering project initiated by Sun Yat-sen University since 2012, represents the first nationally representative labor force tracking survey in China. Conducted biennially, the CLDS focuses on communities (villages), households, and individual laborers to systematically observe the social structures within communities and the evolving dynamics and interactions among households and individual labor participants.

[20] *Work Freedom* integers from 1 to 3, with larger indicating the less freedom to work. *Work Satisfaction* integers from 1 to 5, with larger indicating greater job dissatisfaction.

REFERENCES

Aaronson, D., & Mazumder, B. (2011). The impact of Rosenwald schools on black achievement. *Journal of Political Economy*, *119*(5), 821–888.

Abadi, A., & Khakzand, M. (2022). Extracting the qualitative dimensions of agritourism for the sustainable development of Charqoli village in Iran: The promotion of vernacular entrepreneurship and environment-oriented preservation perspectives. *Environment, Development and Sustainability*, *24*(11), 12609–12671.

Aguiar, M., & Hurst, E. (2007). Measuring trends in leisure: The allocation of time over five decades. *The Quarterly Journal of Economics*, *122*(3), 969–1006.

Aker, J. C., & Mbiti, I. M. (2010). Mobile phones and economic development in Africa. *Journal of Economic Perspectives*, *24*(3), 207–232.

Aldao, C., Blasco, D., Espallargas, M. P., & Rubio, S. P. (2021). Modeling the crisis management and impacts of 21st century disruptive events in tourism: The case of the COVID-19 pandemic. *Tourism Review*, *76*(4), 929–941.

Alrwajfah, M. M., Almeida-García, F., & Cortés-Macías, R. (2020). Females' perspectives on tourism's impact and their employment in the sector: The case of Petra, Jordan. *Tourism Management*, *78*, 104069.

Andergassen, R., & Candela, G. (2013). Less developed countries, tourism investments and local economic development. *Review of Development Economics*, *17*(1), 16–33.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Arroyo, C. G., Barbieri, C., & Rich, S. R. (2013). Defining agritourism: A comparative study of stakeholders' perceptions in Missouri and North Carolina. *Tourism Management*, *37*, 39–47.

Attanasio, O. P., & Pistaferri, L. (2016). Consumption inequality. *Journal of Economic Perspectives*, *30*(2), 3–28.

Ba, M., Anwar, A., & Mughal, M. (2021). Non-farm employment and poverty reduction in Mauritania. *Journal of International Development*, *33*(3), 490–514.

Banerjee, O., Cicowiez, M., Morris, E. J., & Moreda, A. (2018). Boosting tourism's contribution to growth and development: Analysis of the evidence. *Review of Development Economics*, *22*(3), 1296–1320.

Bano, S., Liu, L., & Khan, A. (2022). Dynamic influence of aging, industrial innovations, and ICT on tourism development and renewable energy consumption in BRICS economies. *Renewable Energy*, *192*, 431–442.

Boukas, N., & Ziakas, V. (2016). Tourism policy and residents' well-being in Cyprus: Opportunities and challenges for developing an inside-out destination management approach. *Journal of Destination Marketing & Management*, *5*(1), 44–54.

Bramwell, B., & Lane, B. (2014). The "critical turn" and its implications for sustainable tourism research. *Journal of Sustainable Tourism*, *22*(1), 1–8.

Brida, J. G., Osti, L., & Faccioli, M. (2011). Residents' perception and attitudes towards tourism impacts: A case study of the small rural community of Folgaria (Trentino–Italy). *Benchmarking: An International Journal*, *18*(3), 359–385.

Brooks, W., & Donovan, K. (2020). Eliminating uncertainty in market access: The impact of new bridges in rural Nicaragua. *Econometrica*, *88*(5), 1965–1997.

Buckley, R., Zhong, L., & Martin, S. (2021). Mental health key to tourism infrastructure in China's new megapark. *Tourism Management*, *82*, 104169.

Chen, H., Ding, Y., Tang, L., & Wang, L. (2022). Impact of urban–rural medical insurance integration on consumption: Evidence from rural China. *Economic Analysis and Policy*, *76*, 837–851.

Chen, P., Clarke, N., & Hracs, B. J. (2022). Urban-rural mobilities: The case of China's rural tourism makers. *Journal of Rural Studies*, *95*, 402–411.

Chen, Y., Fan, Z., Gu, X., & Zhou, L. A. (2020). Arrival of young talent: The send-down movement and rural education in China. *American Economic Review*, *110*(11), 3393–3430.

Cheng, Z. (2021). Education and consumption: Evidence from migrants in Chinese cities. *Journal of Business Research*, *127*, 206–215.

Chi, X., & Han, H. (2021). Emerging rural tourism in China's current tourism industry and tourist behaviors: The case of Anji County. *Journal of Travel & Tourism Marketing*, *38*(1), 58–74.

Cole, M. T., & Davies, R. B. (2014). Royale with cheese: Globalization, tourism, and the variety of goods. *Review of Development Economics*, *18*(2), 386–400.

Crawford, C., & Bell, S. (2012). Analysing the relationship between urban livelihoods and water infrastructure in three settlements in Cusco, Peru. *Urban Studies*, *49*(5), 1045–1064.

Curtis, C. C., Lugauer, S., & Mark, N. C. (2015). Demographic patterns and household saving in China. *American Economic Journal: Macroeconomics*, *7*(2), 58–94.

Dang, P., Ren, L., & Li, J. (2024). Does rural tourism reduce relative poverty? Evidence from household surveys in western China. *Tourism Economics*, *30*(2), 498–521.

Daugstad, K. (2008). Negotiating landscape in rural tourism. *Annals of Tourism Research*, *35*(2), 402–426.

De Brauw, A., & Rozelle, S. (2008). Migration and household investment in rural China. *China Economic Review*, *19*(2), 320–335.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996.

Deng, T., Liu, S., & Hu, Y. (2022). Can tourism help to revive shrinking cities? An examination of Chinese case. *Tourism Economics*, *28*(6), 1683–1691.

Donaldson, D. (2018). Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*, *108*(4–5), 899–934.

Durbarry, R. (2008). Tourism taxes: Implications for tourism demand in the UK. *Review of Development Economics*, *12*(1), 21–36.

Emran, M. S., & Hou, Z. (2013). Access to markets and rural poverty: Evidence from household consumption in China. *Review of Economics and Statistics*, *95*(2), 682–697.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Fleischer, A., & Tchetchik, A. (2005). Does rural tourism benefit from agriculture? *Tourism Management*, *26*(4), 493–501.

Gao, J., & Wu, B. (2017). Revitalizing traditional villages through rural tourism: A case study of Yuanjia Village, Shaanxi Province, China. *Tourism Management*, *63*, 223–233.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2016). *Impact evaluation in practice*. World Bank Publications.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

Gradín, C., & Wu, B. (2020). Income and consumption inequality in China: A comparative approach with India. *China Economic Review*, *62*, 101463.

Guillaud, E., Olckers, M., & Zemmour, M. (2020). Four levers of redistribution: The impact of tax and transfer systems on inequality reduction. *Review of Income and Wealth*, *66*(2), 444–466.

Gunter, U., Smeral, E., & Zekan, B. (2022). Forecasting tourism in the EU after the COVID-19 crisis. *Journal of Hospitality & Tourism Research*. https://doi.org/10.1177/1096348022112513

Han, H. (2021). Consumer behavior and environmental sustainability in tourism and hospitality: A review of theories, concepts, and latest research. *Journal of Sustainable Tourism*, *29*(7), 1021–1042.

Hatan, S., Fleischer, A., & Tchetchik, A. (2021). Economic valuation of cultural ecosystem services: The case of landscape aesthetics in the agritourism market. *Ecological Economics*, *184*, 107005.

Hong, J., Shi, F., & Zheng, Y. (2023). Does network infrastructure construction reduce energy intensity? Based on the "Broadband China" strategy. *Technological Forecasting and Social Change*, *190*, 122437.

Huang, X., Zhang, K., Wang, H., Xiong, Z., & Hu, B. (2023). Can rural tourism reduce the urban-rural income gap?—Empirical evidence from "National Leisure Agriculture and Rural Tourism Demonstration Counties". *Tourism Tribune*, *38*(2), 16–29. (in Chinese).

Huang, Y. (2010). Debating China's economic growth: The Beijing consensus or the Washington consensus. *Academy of Management Perspectives*, *24*(2), 31–47.

Hwang, J., & Lee, S. (2015). The effect of the rural tourism policy on non-farm income in South Korea. *Tourism Management*, *46*, 501–513.

Joseph, I. L. (2022). The effect of natural disaster on economic growth: Evidence from a major earthquake in Haiti. *World Development*, *159*, 106053.

Kakwani, N. (1984). The relative deprivation curve and its applications. *Journal of Business & Economic Statistics*, *2*(4), 384–394.

Katircioglu, S. T. (2014). International tourism, energy consumption, and environmental pollution: The case of Turkey. *Renewable and Sustainable Energy Reviews*, *36*, 180–187.

Khadaroo, J., & Seetanah, B. (2007). Transport infrastructure and tourism development. *Annals of Tourism Research*, *34*(4), 1021–1032.

Khanal, A. R., Mishra, A. K., & Nedumaran, S. (2019). Consumption, habit formation, and savings: Evidence from a rural household panel survey. *Review of Development Economics*, *23*(1), 256–274.

Khandker, S. R., Bakht, Z., & Koolwal, G. B. (2009). The poverty impact of rural roads: Evidence from Bangladesh. *Economic Development and Cultural Change*, *57*(4), 685–722.

Khandker, S. R., Koolwal, G. B., & Samad, H. A. (2009). *Handbook on impact evaluation: Quantitative methods and practices*. World Bank Publications.

Kline, P., & Moretti, E. (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics*, *129*(1), 275–331.

Lee, J. J., & Sawada, Y. (2010). Precautionary saving under liquidity constraints: Evidence from rural Pakistan. *Journal of Development Economics*, *91*(1), 77–86.

Li, C., Lin, L., & Gan, C. E. (2016). China credit constraints and rural households' consumption expenditure. *Finance Research Letters*, *19*, 158–164.

Li, T., Li, Q., & Liu, J. (2021). The spatial mobility of rural tourism workforce: A case study from the micro analytical perspective. *Habitat International*, *110*, 102322.

Li, X., & Luo, L. (2021). Migration patterns and migrant workers' consumption. *China Agricultural Economic Review*, *13*(4), 781–798.

Liu, C., Dou, X., Li, J., & Cai, L. A. (2020). Analyzing government role in rural tourism development: An empirical investigation from China. *Journal of Rural Studies*, *79*, 177–188.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Liu, M., Min, S., Ma, W., & Liu, T. (2021). The adoption and impact of E-commerce in rural China: Application of an endogenous switching regression model. *Journal of Rural Studies*, *83*, 106–116.

Liu, X. J., Liu, H., Liu, S. Y., Shen, J. H., & Lee, C. C. (2022). Profit sharing, industrial upgrading, and global supply chains: Theory and evidence. *Review of Development Economics*, *26*(4), 2454–2493.

Lu, Y., Tao, Z., & Zhu, L. (2017). Identifying FDI spillovers. *Journal of International Economics*, *107*, 75–90.

Lu, Y., Wang, J., & Zhu, L. (2019). Place-based policies, creation, and agglomeration economies: Evidence from China's economic zone program. *American Economic Journal: Economic Policy*, *11*(3), 325–360.

Lugauer, S., Ni, J., & Yin, Z. (2019). Chinese household saving and dependent children: Theory and evidence. *China Economic Review*, *57*, 101091.

Ma, W., Abdulai, A., & Ma, C. (2018). The effects of off-farm work on fertilizer and pesticide expenditures in China. *Review of Development Economics*, *22*(2), 573–591.

Ma, W., McKay, A., Rahut, D. B., & Sonobe, T. (2023). An introduction to rural and agricultural development in the digital age. *Review of Development Economics*, *27*(3), 1273–1286.

Ma, W., Nie, P., Zhang, P., & Renwick, A. (2020). Impact of Internet use on economic well-being of rural households: Evidence from China. *Review of Development Economics*, *24*(2), 503–523.

MacNeill, T., & Wozniak, D. (2018). The economic, social, and environmental impacts of cruise tourism. *Tourism Management*, *66*, 387–404.

Mahadevan, R., Amir, H., & Nugroho, A. (2017). Regional impacts of tourism-led growth on poverty and income inequality: A dynamic general equilibrium analysis for Indonesia. *Tourism Economics*, *23*(3), 614–631.

Meyer, B. D., & Sullivan, J. X. (2023). Consumption and income inequality in the United States since the 1960s. *Journal of Political Economy*, *131*(2), 247–284.

Mwesiumo, D., Halfdanarson, J., & Shlopak, M. (2022). Navigating the early stages of a large sustainability-oriented rural tourism development project: Lessons from Træna, Norway. *Tourism Management*, *89*, 104456.

Nunkoo, R., & Ramkissoon, H. (2011). Developing a community support model for tourism. *Annals of Tourism Research*, *38*(3), 964–988.

Pandya, R. (2020). Growth and preferences of tourism industry in India: A distinctive study in the framework of Gujarat. *International Journal of Research in all Subjects in Multi Languages*, *8*(6), 13–20.

Pérez-Olmos, K. N., & Aguilar-Rivera, N. (2021). Agritourism and sustainable local development in Mexico: A systematic review. *Environment, Development and Sustainability*, *23*(12), 17180–17200.

Priem, R. L., Li, S., & Carr, J. C. (2012). Insights and new directions from demand-side approaches to technology innovation, entrepreneurship, and strategic management research. *Journal of Management*, *38*(1), 346–374.

Qu, M., McCormick, A. D., & Funck, C. (2022). Community resourcefulness and partnerships in rural tourism. *Journal of Sustainable Tourism*, *30*(10), 2371–2390.

Rivera, M., Croes, R., & Lee, S. H. (2016). Tourism development and happiness: A residents' perspective. *Journal of Destination Marketing & Management*, *5*(1), 5–15.

Rosalina, P. D., Dupre, K., & Wang, Y. (2021). Rural tourism: A systematic literature review on definitions and challenges. *Journal of Hospitality and Tourism Management*, *47*, 134–149.

Saxena, G., Clark, G., Oliver, T., & Ilbery, B. (2007). Conceptualizing integrated rural tourism. *Tourism Geographies*, *9*(4), 347–370.

Schubert, S. F., & Schamel, G. (2021). Sustainable tourism development: A dynamic model incorporating resident spillovers. *Tourism Economics*, *27*(7), 1561–1587.

Šegota, T., Mihalič, T., & Kuščer, K. (2017). The impact of residents' informedness and involvement on their perceptions of tourism impacts: The case of Bled. *Journal of Destination Marketing & Management*, *6*(3), 196–206.

Seyitoğlu, F., & Ivanov, S. (2021). Service robots as a tool for physical distancing in tourism. *Current Issues in Tourism*, *24*(12), 1631–1634.

Singh, G., Rawat, G. S., & Verma, D. (2010). Comparative study of fuelwood consumption by villagers and seasonal "Dhaba owners" in the tourist affected regions of Garhwal Himalaya, India. *Energy Policy*, *38*(4), 1895–1899.

Soltani, M., Soltani Nejad, N., Taheri Azad, F., Taheri, B., & Gannon, M. J. (2021). Food consumption experiences: A framework for understanding food tourists' behavioral intentions. *International Journal of Contemporary Hospitality Management*, *33*(1), 75–100.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Stevenson, B., & Wolfers, J. (2013). Subjective well-being and income: Is there any evidence of satiation? *American Economic Review*, *103*(3), 598–604.

Streimikiene, D., Svagzdiene, B., Jasinskas, E., & Simanavicius, A. (2021). Sustainable tourism development and competitiveness: The systematic literature review. *Sustainable Development*, *29*(1), 259–271.

Su, B. (2011). Rural tourism in China. *Tourism Management*, *32*(6), 1438–1441.

Telfer, D. J., & Sharpley, R. (2015). *Tourism and development in the developing world*. Routledge.

Vatsa, P., Li, J., Luu, P. Q., & Botero-R, J. C. (2023). Internet use and consumption diversity: Evidence from rural China. *Review of Development Economics*, *27*(3), 1287–1308.

Wan, J., Nie, C., & Zhang, F. (2021). Does broadband infrastructure really affect consumption of rural households?—A quasi-natural experiment evidence from China. *China Agricultural Economic Review*, *13*(4), 832–850.

Wang, R., Dai, M., Ou, Y., & Ma, X. (2021). Residents' happiness of life in rural tourism development. *Journal of Destination Marketing & Management*, *20*, 100612.

Wilson, S., Fesenmaier, D. R., Fesenmaier, J., & Van Es, J. C. (2001). Factors for success in rural tourism development. *Journal of Travel Research*, *40*(2), 132–138.

Yang, B., Wang, X., Wu, T., & Deng, W. (2023). Reducing farmers' poverty vulnerability in China: The role of digital financial inclusion. *Review of Development Economics*, *27*, 1445–1480. https://doi.org/10.1111/rode.12991

Yang, J., Yang, R., Chen, M. H., Su, C. H. J., Zhi, Y., & Xi, J. (2021). Effects of rural revitalization on rural tourism. *Journal of Hospitality and Tourism Management*, *47*, 35–45.

Yang, X. S., & Xu, H. (2022). Producing an ideal village: Imagined rurality, tourism and rural gentrification in China. *Journal of Rural Studies*, *96*, 1–10.

Zhang, Y., Guo, Y., & Ji, L. (2022). Going somewhere or for someone? The Sense of Human Place Scale (SHPS) in Chinese rural tourism. *Tourism Management*, *91*, 104530.

Zhao, C., Chen, B., & Li, X. (2023). Rising housing prices and marriage delays in China: Evidence from the urban land transaction policy. *Cities*, *135*, 104214.

Zhao, C., Wu, Y., & Guo, J. (2022). Mobile payment and Chinese rural household consumption. *China Economic Review*, *71*, 101719.

Zhao, L., & Xia, X. (2020). Tourism and poverty reduction: Empirical evidence from China. *Tourism Economics*, *26*(2), 233–256.

Zhong, Y., Tang, L., & Hu, P. (2020). The mechanism and empirical analysis of the integration of agriculture and tourism to promote the optimization and upgrading of rural industrial structure: A case study of National Demonstration Counties of Leisure Agriculture and Rural Tourism. *Chinese Rural Economy*, *36*(7), 80–98. (in Chinese).

Zhu, Z., Ma, W., Leng, C., & Nie, P. (2021). The relationship between happiness and consumption expenditure: Evidence from rural China. *Applied Research in Quality of Life*, *16*, 1587–1611.

How to cite this article: Wei, B., Zhao, C., Luo, M., Jia, Z., & Xia, H. (2024). Beyond the scenery: Understanding the impact of rural tourism development on household consumption in China. *Review of Development Economics*, *28*(3), 1152–1191. https://doi.org/10.1111/rode.13096

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

APPENDIX A

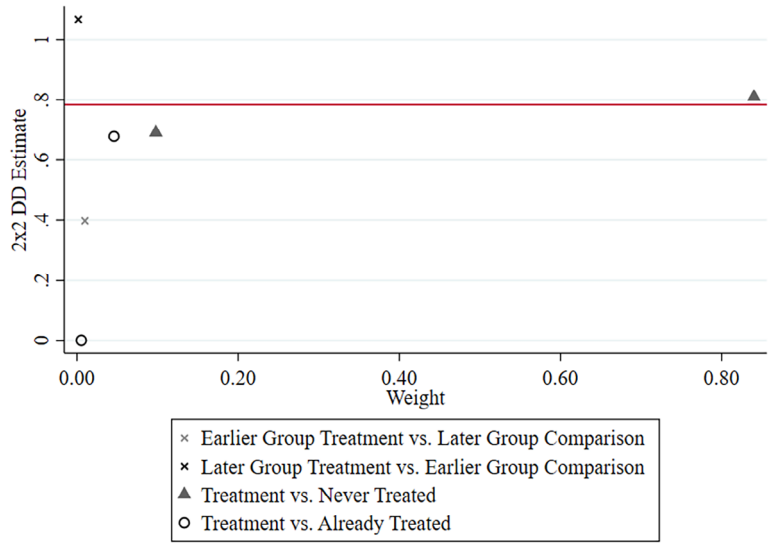


FIGURE A1 Distribution of TWFE estimation bias weights. TWFE, two-way fixed effects. [Colour figure can be viewed at wileyonlinelibrary.com]

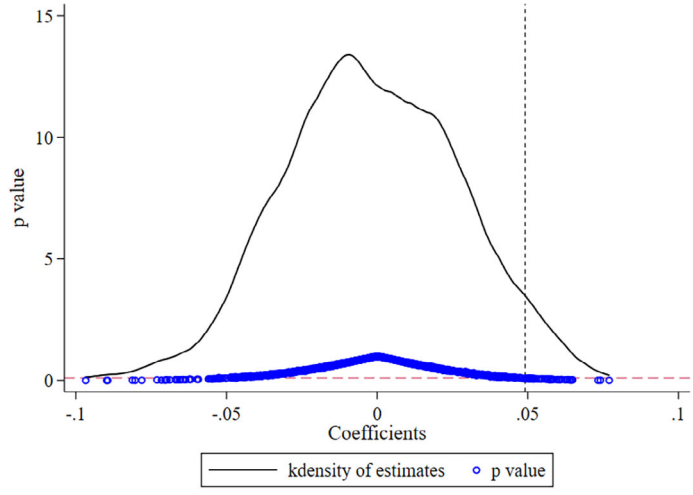


FIGURE A2 Placebo tests. [Colour figure can be viewed at wileyonlinelibrary.com]

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A1 Variable definitions and descriptive statistics.

Variables	Definitions	Full sample			LARTDC = 0	LARTDC = 1	Diff.
		N	Mean	SD	Mean	Mean	Mean
Household consumption	Household consumption expenditure per capita (yuan, logarithm)	48,389	8.9345	0.7993	8.9103	9.2406	−0.3303***
LARTDC	= 1 if a county is the LARTDC, 0 otherwise	48,389	0.0731	0.2603			
Age	Householder age	48,370	53.0299	13.7662	52.9556	53.9723	−1.0167***
Age squared	Square of the age of householder/100	48,370	30.0168	99.1171	29.9970	30.2671	−0.2700
Gender	Householders' gender. Male = 1; female = 0	48,389	0.8902	0.3126	0.8934	0.8502	0.0432***
Hukou	Householders' *hukou* status. Rural = 1; urban = 0	48,389	0.9648	0.1842	0.9663	0.9466	0.0197***
Education	Householders' years of education.	47,385	6.9971	3.0442	6.9522	7.5637	−0.6115***
Self-rated health	Householders' health status. Excellent = 1; good = 2; indifferent = 3; poor = 4; incapacity for work = 5	48,354	1.6895	0.8986	1.6907	1.6743	0.0164
Party	Householders' political identity. Party membership = 1, 0 otherwise	48,389	0.1515	0.3585	0.1460	0.2202	−0.0741***
Total income	Household total income (yuan, logarithm)	48,268	10.2950	0.8267	10.2694	10.6210	−0.3516***
Current deposit	Household current deposit (yuan, logarithm)	48,379	4.9924	5.1515	4.9332	5.7428	−0.8095***
Medical care ratio	Percentage of households participating in new rural cooperative medical care in the village	48,389	0.9805	0.2297	0.9808	0.9767	0.0041
Village income	The average income per person in the village (yuan, logarithm)	44,810	8.8160	0.6242	8.7813	9.3081	−0.5268***
Village terrain	Plain = 1; hilly = 2; mountainous = 3	45,536	1.8823	0.8412	1.8762	1.9696	−0.0934***
Village officials	The actual number of village officials	45,541	6.0710	3.6611	5.9044	8.4315	−2.5271***
Village economic regions	Agricultural area or forestry area = 1; Other types (such as pastoral area) = 0	45,879	0.9529	0.2120	0.9514	0.9729	−0.0215***
Industrial structure	The ratio of value added in the secondary and tertiary sectors to value added in the primary sector	48,389	7.1348	10.5650	7.0557	8.1380	−1.0823***
Government	County local fiscal general budget revenue/GDP	48,389	0.0564	0.0338	0.0554	0.0683	−0.0129***
GDP	10,000 yuan, logarithm	48,389	14.0350	0.9560	13.9935	14.5613	−0.5678***

Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A2 Propensity score matching results.

	Household consumption		
	(1)	**(2)**	**(3)**
	Kernel matching	**Radius matching**	**Nearest neighbor matching**
LARTDC	0.0472*** (0.0167)	0.0342** (0.0169)	0.0471*** (0.0167)
Control variables	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
R^2	0.6746	0.6730	0.6739
N	41,061	40,706	40,919

Note: Robust standard errors are reported in parentheses.
Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

TABLE A3 Eliminating disturbing policies.

	Household consumption			
	(1)	**(2)**	**(3)**	**(4)**
	ICDC and CCDC	**SCC and BC**	**NREDC**	**BEDC**
LARTDC	0.0487*** (0.0162)	0.0439*** (0.0163)	0.0594*** (0.0162)	0.0486*** (0.0161)
Control variables	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
R^2	0.6737	0.6738	0.6740	0.6738
N	41,986	41,986	41,986	41,986

Note: Robust standard errors are reported in parentheses.
Abbreviations: BC, Broadband China; BEDC, Back-to-Home Entrepreneurship Demonstration County; CCDC, Cultural Consumption Demonstration City; ICCD, Information Consumption Demonstration City; LARTDC, Leisure Agriculture and Rural Tourism Demonstration County; NREDC, National Rural E-commerce Demonstration County; SCC, Smart City Constructions.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A4 Bacon decomposition results.

	Estimates	Weight
Early Treatment versus Later Control	0.398	0.010
Later Treatment versus Early Control	1.068	0.001
Treatment versus Never treated	0.797	0.938
Treatment versus Already treated	0.608	0.051

TABLE A5 Alternative specifications.

	Household consumption				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
	Additional controls	**Considering the expected effects**	**Excluding extreme values**	**Excluding observations with an urban *hukou***	**DID_M estimations**
LARTDC	0.0476*** (0.0164)	0.0541*** (0.0180)	0.0474*** (0.0141)	0.0469*** (0.0165)	
$LARTDC_M$					0.0425** (0.0210)
Control variables	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.6738	0.6737	0.6377	0.6741	
N	41,986	41,986	37,773	40,486	41,986

Note: Robust standard errors are reported in parentheses.
Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A6 Further evidence for rural tourism.

	Household consumption			
	(1)	**(2)**	**(3)**	**(4)**
LARTDC	0.0582*** (0.0207)	0.0394 (0.0255)	0.0558*** (0.0184)	0.5075*** (0.1398)
LARTDC × *Village investments*	−0.0043 (0.0104)			
LARTDC × *Village enterprises*		0.1629 (1.8625)		
LARTDC × *Village environment*			−0.0001* (0.00004)	
LARTDC × *Village institutions*				0.0915*** (0.0237)
Control variables	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
R^2	0.7145	0.6800	0.6791	0.6884
N	22,232	22,264	37,097	18,710

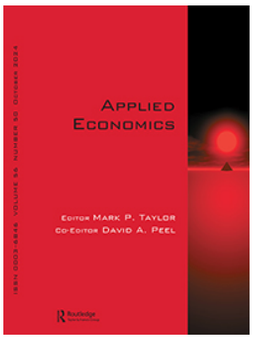
Note: Four variables *Village investments*, *Village enterprises*, *Village environment*, and *Village institutions* are controlled from Column (1) to Column (4), respectively. Robust standard errors are reported in parentheses.
Abbreviation: LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

TABLE A7 Potentially competitive interpretations: using the CLDS data.

	(1)	**(2)**	**(3)**	**(4)**
	Weekly working hours	**Annual working hours**	**Freedom of employment**	**Work satisfaction**
Panel A: All observations				
LARTDC	−4.9381 (5.9386)	−5.837 (4.1808)	−0.0612** (0.0266)	0.0366 (0.0261)
N	11,185	11,185	11,184	13,289
Panel B: The elderly				
LARTDC	7.3280 (16.3577)	−2.6274 (11.8144)	−0.1348*** (0.0492)	0.0762 (0.0598)
N	2120	2120	2120	2872
Panel C: Female labors				
LARTDC	−5.1985 (7.9243)	−4.9381 (5.9386)	−0.0608 (0.0393)	0.0226 (0.0381)
N	5157	5157	5157	6583

Note: The control variables include gender, political profile, literacy, social status, marital status, and number of siblings. We added province fixed effects in all columns. Robust standard errors are reported in parentheses.
Abbreviations: CLDS, China Labor Force Dynamic Survey; LARTDC, Leisure Agriculture and Rural Tourism Demonstration County.
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

14679361, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/rode.13096 by South China Agricultural, Wiley Online Library on [02/07/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



Applied Economics

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/raec20

Teachers' administrative positions and students' non-cognitive abilities in China: evidence from a quasi-natural experiment

Boou Chen & Chunkai Zhao

To cite this article: Boou Chen & Chunkai Zhao (2024) Teachers' administrative positions and students' non-cognitive abilities in China: evidence from a quasi-natural experiment, Applied Economics, 56:50, 6091-6108, DOI: 10.1080/00036846.2023.2267819

To link to this article: https://doi.org/10.1080/00036846.2023.2267819

Published online: 09 Oct 2023.

Submit your article to this journal

Article views: 145

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=raec20

APPLIED ECONOMICS
2024, VOL. 56, NO. 50, 6091–6108
https://doi.org/10.1080/00036846.2023.2267819

Routledge
Taylor & Francis Group

Check for updates

Teachers' administrative positions and students' non-cognitive abilities in China: evidence from a quasi-natural experiment

Boou Chen [a] and Chunkai Zhao [b]

[a]School of Economics, Nanjing University of Finance and Economics, Nanjing, Jiangsu Province, China; [b]College of Economics and Management, South China Agricultural University, Guangzhou, Guangdong Province, China

ABSTRACT

This research investigates the causal relationship between teachers' administrative positions and students' non-cognitive abilities (NCAs) in China using a quasi-natural experiment of random student-classroom assignment. The results suggest that headteachers' administrative positions significantly improve students' NCAs, especially the extraversion and neuroticism. This positive effect is more pronounced in headteachers with higher administrative positions, middle titles, and more teaching experience. Meanwhile, girls and urban students benefited more from headteachers' administrative positions. Furthermore, the mechanism analysis indicates this positive effect can be explained by the improvement of teachers' job satisfaction, higher frequency of using more advanced teaching equipment, more parental respect, increased sense of belonging to the class, and peer effects. Our findings further enrich the impact of teacher characteristics on students' NCAs and highlight the potential benefits of teachers' administrative positions on student human capital. Also, this paper had policy implications for promoting students' human capital development by optimizing teacher assignment and allocation.

KEYWORDS

Teachers' administrative positions; students' non-cognitive abilities (NCAs); human capital; quasi-natural experiment; China

JEL CLASSIFICATION

I21; J13

I. Introduction

In recent years, the role of non-cognitive abilities (NCAs) in human capital development and individual success has been widely recognized (e.g. Blanden, Gregg, and Macmillan 2007; Heckman and Rubinstein 2001). Distinguished from cognitive abilities, NCAs are a collection of personality traits, emphasizing individual soft power, such as self-confidence, perseverance, concentration, self-regulation, and emotional control (Cobb-Clark and Schurer 2012; Costa and McCrae 1992). Scholars emphasized the role of non-cognitive abilities in positive economic, health, education, and psychological outcomes (e.g. Almlund et al. 2011; Bowles, Gintis, and Osborne 2001; Gong, Lu, and Song 2021; Rouse and Hunziker 2020). For students in particular, some previous studies confirmed that NCAs could help improve students' academic achievements (e.g. Heckman and Rubinstein 2001; West et al. 2016). Clearly, teacher characteristics are closely related to children's human capital accumulation, such as teacher gender, quality, income, and job satisfaction (Bhai and Horoi 2019; Chen and Zhao 2022; Gong, Lu, and Song 2018; Horoi and Bhai 2018; Loyalka et al. 2019), which are also likely to be the important factors influencing the development of students' NCAs (Blazar and Kraft 2017; Doherty and Conolly 1985; Jackson 2018).

Prior research on teachers and students' NCAs has been conducted in two main ways. One body of literature linking teachers and students' academic performance through NCAs has not yielded uniform findings. Specifically, some studies concluded that teachers have a weak effect on some indicators of students' NCAs (Cheng and Zamarro 2018; Kraft 2019), while others found that teachers significantly enhance students' NCAs (Araujo et al. 2016; Rouse and Hunziker 2020), even greater than their effect on the test score (Jackson 2018). Another category of literature focused on the effects of teacher characteristics on students' NCAs, such as teacher gender (Gong, Lu, and Song 2018), teaching quality (Gershenson 2016), and teacher-student interactions (DeAngelis 2021). Unlike these previous studies, this paper aimed to explore the effect of

CONTACT Chunkai Zhao ckz_zhao@163.com College of Economics and Management, South China Agricultural University, No.483, Wushan, Tianhe District, Guangzhou, Guangdong Province, China

© 2023 Informa UK Limited, trading as Taylor & Francis Group

teachers' administrative positions on students' NCAs, especially in the Chinese education system where it is very common for teachers to take on administrative positions.

In China and some other Asian countries, teachers in compulsory education[1] are required to take on some additional administrative positions beyond teaching (Kim 2019; Okumura 2017). In China, teachers are encouraged to hold administrative positions (Zheng and Sun 2020; Zhong and Liu 2021), and many teachers are motivated to hold administrative positions due to the potential administrative benefits (Li 2018). According to recent studies, roughly 16% of teachers in middle schools in China are also employed as administrators (Zheng and Sun 2020). However, the effect of teachers' administrative positions on students' NCAs may be twofold. On the one hand, administrative work may drain teachers' energy excessively, reduce their amount of time available for instruction (Chen and Zhao 2022; Kim 2019), exacerbate job burnout and work exhaustion (Li 2018; Zhong and Liu 2021), and even cause occupational stress (Sun and Zhou 2022). These results in a decline in teaching quality (Ballet, Kelchtermans, and Loughran 2006), which is detrimental to the improvement of students' NCAs (Rouse and Hunziker 2020). On the other hand, Chinese schools favour senior or more successful instructors for administration roles due to the impact of classical Confucianism. Thus, taking on administrative positions may allow teachers to be perceived as entrusted with important responsibilities, thereby improving their teaching efforts (Zheng and Sun 2020), which could enhance students' NCAs (Gershenson 2016). Furthermore, teachers with administrative positions have greater authority than other teachers in compulsory education (Gerrard and Farrell 2014), and their words and actions are more likely to have a positive modelling effect on students.

A possible endogenous issue could have an impact on this conflict from earlier investigations. The availability of administrative positions for teachers is not necessarily random but rather selected (Chen and Zhao 2022; Sun and Zhou 2022). Many confounding variables, including school characteristics, teacher characteristics, and some other unobserved features that are associated with admission into administrative positions, may also affect the identification of the causal link between teachers' administrative positions and students' NCAs. However, these confounders may also possibly have an impact on students' NCAs, but they are challenging to completely control (Araujo et al. 2016; DeAngelis 2021).

We use a quasi-natural experiment of random student-classroom assignment to determine the causal effect of teachers' administrative positions on students' NCAs to overcome the endogeneity issues. In recent years, this random assignment approach has been widely adopted in the research field of economics of education (e.g. Chen and Zhao 2022; Gong, Lu, and Song 2018; Hu 2018; Sun and Zhou 2022). Moreover, given China's particular headteacher system, we mainly consider the impact of headteachers holding administrative positions on students' NCAs. Although headteachers are widely available in Asian and European countries (Kashy-Rosenbaum, Kaplan, and Israel-Cohen 2018), Chinese headteachers are not only responsible for all aspects of the classroom, but may also take on teaching tasks and administrative positions. Previous literature also confirmed the important role of headteachers in the development of students' NCAs (Gong, Lu, and Song 2018; Kashy-Rosenbaum, Kaplan, and Israel-Cohen 2018). As the most representative teacher in a Chinese compulsory education class, taking headteachers as a case is fully convincing.

This paper contributes to the prior research in the following ways. First, to the best of our knowledge, this study is among the first to examine the role of headteachers' administrative positions in improving students' NCAs. Thus, our study mainly complements two types of literature. One involves the effect of teacher characteristics on students' human capital development (Araujo et al. 2016; Bhai and Horoi 2019; Gong, Lu, and Song 2018); another category focuses on measuring students' NCAs in terms of the Big Five Personality Model (BFPM) (Zhao and Chen 2022), which helps to fully capture students' noncognitive performance through multiple

[1]Primary and junior high schools are included in China's compulsory education system.

dimensions (Costa and McCrae 1992). Second, we explore the mechanisms by which headteachers' administrative positions enhanced students' NCAs, which provides an explanation for the influence of this special teacher characteristic on students' NCAs during compulsory schooling in the special Chinese setting. We find that these mechanisms include teachers' attitude and teaching method, students' sense of belonging to the class, and peer effects. Moreover, heterogeneity effect results from teacher, and student levels also provide some guidance for allocating and planning administrative tasks as well as some insights for improving the relevant education policies.

II. Teachers' administrative positions in Chinese compulsory education

In Chinese compulsory schools, it is encouraged for teachers to have administrative jobs concurrently for reasons including ease of administration, authorized strength compression, and reduction of educational expenditures. In 2001, the General Office of the State Council, the Ministry of Education, and the Ministry of Finance of China jointly issued 'Opinions on the Formulation of Staffing Standards for Primary and Secondary Schools', stating that as far as feasible, teachers should handle the administration of primary and secondary schools. Schools that do need to have dedicated administrative staffs to support teaching, logistics and school administration, these personnel make up no more than 15% of the teaching staff in junior high schools and no more than 9% in elementary schools.

Moreover, as parents and society have gradually placed greater emphasis on compulsory education, the administrative work faced by schools has increased significantly (Zhang and Pang 2016). In addition to traditional instructional management, schools need more staff to implement relevant educational policies, handle external public relations, and strengthen school infrastructure and soft power. Furthermore, in the context of advancing the reform of quality education, teachers in the compulsory education stage are required to not only pay attention to students' academic performance, but also organize a number of extracurricular activities to promote students' all-round development (Dello-Iacovo 2009). Because of the conflict between the requirement for more complicated administrative activities and the lack of administrative employees, many teachers are now filling administrative roles.

Excluding these objective factors, teachers also have a desire to take initiative in administrative positions. First, some schools offer additional compensation to teachers who hold administrative positions, such as administrative stipends and benefits. Thus, some teachers may be willing to take on these additional administrative tasks to earn more revenue. Second, administrative positions in Chinese schools, which are a unique form of political capital, also provide teachers with additional implicit benefits, such as increased training opportunities, improved research environments, and priority access to special resources (García-Gallego et al. 2015). Even more significant, many schools view administrative work as a kind of performance evaluation that gives teachers an advantage in title promotion (Zheng and Sun 2020). In summary, due to these factors, it is currently common for teachers in compulsory education in China to hold administrative positions.

III. Data and variables

Data

The data used in our study come from the baseline wave of the China Education Panel Survey (CEPS) in the 2013–2014 academic year, a large nationally representative tracking survey, which aims to reveal the impact of household, school, community, and other macro social structures on student educational outcomes. The CEPS focuses on 7th and 9th grade middle school students in China, containing the information about characteristics of schools, classes, teachers, students, and families (Sun and Zhou 2022; Zhao and Chen 2023). More particularly, the CEPS investigates whether the headteacher holds an administrative position, allowing us to identify the main explanatory factor (Chen and Zhao 2022), and also contains the test results of personality traits, which can be used to measure students' NCAs. In addition, the CEPS includes many questions involving the teachers' teaching methods and attitudes, the

relationship between students and teachers, and students' subjective attitude, which ensures the feasibility of mechanism analysis. Therefore, the CEPS facilitates our study on exploring the effect of teachers' administrative positions on students' NCAs in Chinese compulsory schools.

Furthermore, the CEPS collected the information about the method of student – classroom assignment of each school, which mainly includes two categories of random assignment and assignment in terms of academic performance. The 'random student-classroom assignment' is used as a quasi-natural experiment to determine the causal relationship between the administrative positions held by headteachers and students' NCAs. Thus, in our study, we only retained schools where administrators randomly or averagely allocated students to classes.

Sample

The students and headteachers in the sample were from 344 classes of 93 middle schools in the 7th and 9th grades. In our retained sample, we include 8286 students and 344 headteachers. Among them, some headteachers hold administrative positions, accounting for about 19.6%. Analyzing the differences in student NCAs between these two groups was of interest. For students ($N = 8286$, 50.9% of girls), 43.8% were from urban areas and 46.8% were 9th graders. In terms of headteachers, their mean age and years of teaching experience were 37.24 years old and 15.6 years, respectively. In addition, the percentage of male headteachers was only 38.4%.

Variables and summary statistics

As mentioned earlier, prior studies on teacher characteristics influencing students' NCAs were mainly limited to certain personality traits, lacking a comprehensive indicator. As shown in Figure 1, we use the BFPM to construct a comprehensive indicator for student NCAs (Cobb-Clark and Schurer 2012; Costa and McCrae 1992),[2] which includes five sub-dimensions: conscientiousness, agreeableness, extraversion, openness, and neuroticism. Based on CEPS questionnaires, a series of indicators were selected to reflect these personality traits (Zhao and Chen 2022). For example, students were asked how often they felt unhappy, which could be considered a sub-indicator for neuroticism. The five sub-dimensions and the items they contain are all scored from 1 to 4, with higher scores indicating a stronger ability. The average scores of these five sub-dimensions were used to calculate the NCAs by equal weighting.

The CEPS records whether headteachers concurrently hold an administrative position.[3] The explanatory variable *Administrative position* is set to 1 if the headteacher had administrative positions, and 0 otherwise. It is important to note that the appointment of teachers to administrative positions may be non-randomly selected, which means it may be associated with other teacher characteristics. And it may lead to administrative positions that capture high-quality teachers, who are likely to be more beneficial to students' NCAs. Considering the above concerns, we have controlled many other characteristics of headteachers in the baseline regressions, including the teacher's age, teaching experience, gender, education, and title, to avoid the interference of other teacher characteristics as much as possible. In addition, relevant studies have confirmed a significant negative effect of teachers' administrative positions on student academic achievement (Chen and Zhao 2022; Sun and Zhou 2022), and this implies that the association between teachers' administrative positions and high quality is not necessarily significant. Both of these studies used the same data (CEPS) of our research, which can also prove the reliability of the data we used. Moreover, based on existing literature related to the determinants of students'

[2]The BFPM consists of five personality traits: conscientiousness, agreeableness, extraversion, openness, and neuroticism. Conscientiousness mainly reflects the individual's sense of achievement, competence, and self-discipline etc. Agreeableness is used to measure trust in others, compliance, and modesty etc. Extraversion reflects individual enthusiasm, activity, and positive emotions etc. Openness reflects the individual's ideas, curiosity and values etc. Neuroticism reflects the individual's emotional state and usually involves anxiety, depression etc (Costa and McCrae 1992). The mean-level changes in Big-Five personality traits are small and do not vary substantially across age groups (Cobb-Clark and Schurer 2012).

[3]Administrative positions in the CEPS primarily fall into three categories: grade team leader (charged with overseeing various grade-related matters), head of teaching and research section (in charge of overseeing the teaching and research of a specific subject), school leader (vice principal and director of teaching affairs).

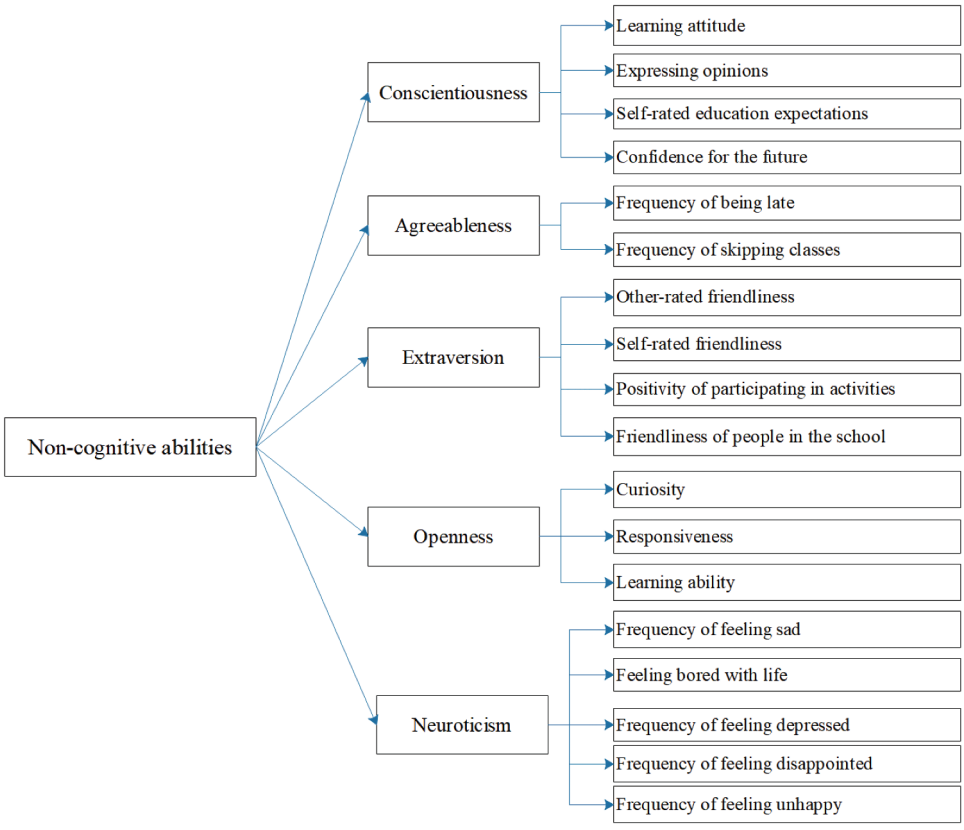


Figure 1. The indicator system for students' NCAs. A higher score for each item in Neuroticism means a lower frequency, indicating a greater ability to control emotions.

NCAs (Zhao and Chen 2022), we also controlled for relevant student and family characteristics, including the grade of student, gender, NCA endowment, *Hukou*,[4] education of their parents, and education expenditure, to avoid confounding by these factors, and to ensure the accuracy and reliability of our results. The definition and summary statistics of the above-mentioned variables are reported in Table 1.

IV. Empirical model and strategy

We develop the following baseline empirical model to investigate the causal impact of headteachers' administrative positions on the NCAs of students:

$$NCA_{ics} = \alpha + \beta HTA_{cs} + \gamma T_{cs} + \delta P_{ics} + \lambda H_{ics} + \theta S_s + \varepsilon_{ics} \quad (1)$$

where the explained variable NCA_{ics} is the NCAs of student i in class c and school s. HTA_{cs} refers to the explanatory variable, suggesting whether or not the headteacher of class c in school s also holds administrative positions. T_{cs}, P_{ics}, and H_{ics} denote three sets of control variables: characteristics of headteachers, students, and households. S_s represents the school fixed effect. ε_{ics} is the random error term.

However, it should be mentioned that in some schools, the assignment of headteachers is not random, that means, headteachers and students are not randomly matched, and will cause two problems. Firstly, the accuracy of the estimation results would be affected because there might be a correlation between the administrative positions of headteachers and other traits of classes, students, and their families. Secondly, if students are not randomly assigned, such as by placing students with high scores or NCAs in a class, this will further skew the results. Thus, we use the 'random student-

[4]The *hukou* system is a basic national administrative system in China. This system divides household registration attributes into agricultural (rural) and non-agricultural (urban) *hukou* based on geography and family members' relationships.

Table 1. Variable definitions and summary statistics.

Variable	Definitions and description	Obs.	Mean	S.D.
NCA	A comprehensive index of NCAs	8286	3.264	0.357
Conscientiousness	The average score across dimensions of conscientiousness	8286	3.015	0.520
Agreeableness	The average score across dimensions of agreeableness	8286	3.846	0.432
Extraversion	The average score across dimensions of extraversion	8286	3.071	0.684
Openness	The average score across dimensions of openness	8286	3.240	0.587
Neuroticism	The average score across dimensions of neuroticism	8286	3.150	0.632
Administrative positions	= 1 if a headteacher had administrative positions; 0 otherwise	8286	0.196	0.397
Age	The teacher's age	8286	37.24	6.932
Teaching experience	The years of teaching experience	8286	15.60	7.532
Gender	The gender of the teacher (0 = female; 1 = male)	8286	0.384	0.486
Education	Teacher's year of education (9 = junior middle school; 12 = senior high school; 15 = junior college; 16 = bachelor; 19 = master)	8286	15.70	0.535
Title	Teacher's Title (0 = no level; 1 = third level; 2 = second level; 3 = first level; 4 = senior; 5 = professor)	8286	2.666	0.991
Gender of student	0 = girl; 1 = boy	8286	0.491	0.500
NCA endowment	Children's NCAs in Grade 6 (1–4 points). This indicator is obtained by taking the average of 7 questions about students' NCAs in their sixth grade.	8286	3.249	0.525
Hukou	0 = rural hukou; 1= urban hukou	8286	0.438	0.496
Grade 9	0 = 7th Grade; 1 = 9th Grade	8286	0.468	0.499
Education expenditure	Expenditures on schooling (yuan), as a logarithm	8286	5.393	3.324
Education of father	Father's year of education (0 = Uneducated; 6 = primary school; 9 = junior middle school; 12 = senior high school; 15 = junior college; 16 = bachelor; 19 = master)	8286	10.29	3.085
Education of mother	Mother's year of education (0 = Uneducated; 6 = primary school; 9 = junior middle school; 12 = senior high school; 15 = junior college; 16 = bachelor; 19 = master)	8286	9.474	3.549

The specific definitions of the five dimensions *Conscientiousness*, *Agreeableness*, *Extraversion*, *Openness*, *Neuroticism* have been explained in Figure 1.

classroom assignment' as a quasi-natural experiment, which is widely used in previous studies (e.g. Chen and Zhao 2022; Gong, Lu, and Song 2018; Hu 2018; Sun and Zhou 2022), to address the endogeneity brought about by the aforementioned issues. In compulsory education in China, the headteacher is predetermined and can be considered part of the classroom identity (Chen and Zhao 2022). Previous literature also confirmed that the impact of teacher administrative positions on students is quite convincing in terms of causal effects in the framework of random class-student assignment (Chen and Zhao 2022; Sun and Zhou 2022). Therefore, according to this empirical strategy, given a random pairing of headteachers and students, we can obtain the unbiased estimate β of HTA_{cs} by estimating the Equation (1) with the OLS method.

Before regressions, it is important to verify that the variance in headteachers holding administrative positions should be random in controlling for school fixed effects. We conduct a t-test between randomly assigned and non-randomly assigned classes, and balancing tests to prove that in the randomly assigned sample, headteachers holding administrative positions is unrelated to any specified background features of the class and its students (Chen and Zhao 2022; Sun and Zhou 2022), which also implies that the randomly assigned classes in the sample did randomly assign students. The summary statistics of predetermined background variables in the balancing tests are reported in Table A1 in the Appendix, and Tables A2 and A3 in the Appendix list the results of t-test and balancing tests, respectively. First, the t-test in Table A2 indicated that when the headteacher has an administrative position, non-randomly assigned classes have a stronger selectivity in the assignment of students than randomly assigned classes, as evidenced by the overall better performance of the students in Math at grade 6, and a higher proportion of parents working in government or public institutions.

Further, the results of the balancing tests in Table A3 show that in the sample of randomly assigned classes (Panel A), headteachers' administrative positions are independent of most class-level variables, which implies that the randomness of headteachers appointments is real. In contrast, in classes with non-random assignment (Panel B), headteachers' administrative positions are strongly correlated with many class characteristics, particularly the proportion of parents working in the public sector. And the above results imply that when schools adopt non-random class assignments, some parents working in the public sector may draw on their political capital or social

connections to get their children into classes with teachers who have administrative positions, possibly because they are subjectively convinced that these teachers perform better in teaching, even though this may not be the reality (Chen and Zhao 2022; Sun and Zhou 2022). Combining the t-test and the balancing tests, the reliability of random student-classroom assignment in the data of our study was confirmed. Therefore, we can eliminate almost all observed associations between the administrative positions of headteachers and the class-level background traits by retaining randomly assigned classes, and ensure the accuracy and reliability of our results.

V. Results

Headteachers' administrative positions and students' NCAs

Table 2 lists the regression coefficients for headteachers' administrative positions on students' NCAs using OLS estimations. The features of headteachers, students, and their households all serve as control factors, and the school fixed effects are also controlled. In column (1), we examine the effect of headteachers' administrative positions on the comprehensive indicator of NCAs, and find that headteachers' administrative positions improve students' NCAs by 0.0278 points. By converting to the standard deviation form (Zhao and Li 2022), it corresponds to a 0.0309 standard deviation improvement in the students' NCAs. Some existing studies have also obtained similar results that teachers' quality and NCAs have a significant impact on students' NCAs (Cheng and Zamarro 2018; Kraft 2019). In general, administrative positions are usually assigned to teachers with greater managerial and social skills, all of which are part of the NCAs. Teachers with administrative positions usually have stronger NCAs and greater authority than other teachers in compulsory education (Gerrard and Farrell 2014), and are likely to gain the respect and trust of students more easily. Therefore, these teachers are more likely to have a positive modelling effect on students' words and deeds, and students' NCAs.

Table 2. The effect of headteachers' administrative positions on students' NCAs.

	(1)	(2)	(3)	(4)	(5)	(6)
		Sub-dimensions of NCAs				
	NCA	Conscientiousness	Agreeableness	Extraversion	Openness	Neuroticism
Administrative positions	0.0278**	0.0229	−0.0070	0.0694***	0.0124	0.0413*
	(0.0108)	(0.0162)	(0.0149)	(0.0226)	(0.0191)	(0.0221)
Age	0.0020	0.0047	0.0046	0.0002	−0.0044	0.0047
	(0.0022)	(0.0033)	(0.0034)	(0.0046)	(0.0041)	(0.0046)
Teaching experience	−0.0020	−0.0038	−0.0030	−0.0001	0.0016	−0.0048
	(0.0021)	(0.0030)	(0.0031)	(0.0042)	(0.0037)	(0.0042)
Gender	−0.0320***	−0.0429***	0.0025	−0.0524***	−0.0245	−0.0427**
	(0.0088)	(0.0133)	(0.0120)	(0.0183)	(0.0158)	(0.0178)
Education	0.0018	−0.0092	−0.0022	0.0094	0.0201	−0.0089
	(0.0089)	(0.0133)	(0.0114)	(0.0175)	(0.0157)	(0.0181)
Title	0.0128*	0.0208*	0.0105	0.0008	0.0303**	0.0017
	(0.0072)	(0.0107)	(0.0103)	(0.0144)	(0.0125)	(0.0141)
Gender of student	−0.0122*	−0.1067***	−0.0324***	−0.0531***	0.0790***	0.0522***
	(0.0068)	(0.0101)	(0.0093)	(0.0140)	(0.0120)	(0.0136)
NCA endowment	0.2154***	0.2396***	0.0763***	0.3320***	0.2379***	0.1913***
	(0.0096)	(0.0127)	(0.0104)	(0.0177)	(0.0143)	(0.0158)
Hukou	−0.0002	−0.0094	−0.0145	0.0221	0.0388**	−0.0378**
	(0.0086)	(0.0129)	(0.0121)	(0.0176)	(0.0151)	(0.0172)
Grade 9	−0.0294***	−0.0609***	−0.0673***	0.0282*	0.0439***	−0.0911***
	(0.0074)	(0.0109)	(0.0100)	(0.0153)	(0.0132)	(0.0150)
Education expenditure	0.0012	0.0058***	0.0008	0.0030	0.0004	−0.0039
	(0.0014)	(0.0022)	(0.0020)	(0.0029)	(0.0026)	(0.0028)
Education of father	0.0065***	0.0171***	−0.0031	0.0070**	0.0093***	0.0022
	(0.0016)	(0.0024)	(0.0023)	(0.0034)	(0.0029)	(0.0034)
Education of mother	0.0101***	0.0151***	0.0019	0.0113***	0.0137***	0.0086***
	(0.0015)	(0.0022)	(0.0022)	(0.0030)	(0.0026)	(0.0029)
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.3967***	2.0642***	3.5326***	1.9220***	2.0835***	2.3811***
	(0.1717)	(0.2526)	(0.2453)	(0.3282)	(0.2902)	(0.3551)
R-squared	0.2644	0.2419	0.0412	0.1669	0.1575	0.0685
N	8286	8286	8286	8286	8286	8286

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are presented in parentheses.

Further, we explore the impacts on five sub-dimensions of NCAs measured by the BFPM, respectively (Costa and McCrae 1992). Obviously, there is a big difference between the capabilities of the above five aspects. Although teachers' administrative positions affect students' overall NCAs, they do not necessarily have a significant impact on the sub-indices of each dimension. As shown in columns (2) to (6), among the five major elements of NCAs, headteachers' administrative positions mainly improve students' extraversion by 0.0694 points and neuroticism by 0.0413 points, while there are no significant effects on conscientiousness, agreeableness, and openness.

Robustness tests

Numerous robustness tests are carried out. Firstly, we exclude some special groups to ensure the representativeness of the sample, such as students from highly educated families, students from extremely poor and wealthy families, unhealthy students, schools in distant places (Hu 2018; Zhao and Chen 2022). The above-mentioned characteristics are likely to have potential effects on adolescents' NCAs that may interfere with the estimation results. Considering that some schools may not assign administrative duties to headteachers due to sufficient teachers, they were also eliminated. Similarly, the administrators of some private schools are usually off-campus personnel, so the sample of private schools is also excluded. Moreover, to avoid the effects on NCAs by the administrative positions of other teachers in the class (non-headteacher), the sample of ordinary teachers who worked in administration was also eliminated. As reported in columns (1) to (7) of Table A4 in the Appendix, the estimates in each column are generally consistent with the baseline regression results regardless of the special groups excluded, with only small changes in the estimated coefficients.

Secondly, in column (8) of Table A4, we compare the regression results with a sample of non-random assignment classes to the baseline regression. It shows that the positive effect of headteachers' administrative positions on students' NCAs is equally significant in the non-random assignment classes, and the coefficient (0.0541) is nearly twice as high as in the random assignment classes. Combined with the balancing tests above (Table A3), it can be inferred that the greater effect in the non-random assignment classes is likely due to the higher proportion of parents in the public sector in the class. Specifically, in the balancing tests, we have confirmed that in the non-random assignment classes, teacher's administrative position is strongly associated with the proportion of parents working in the public sector, and public officials tend to be more educated overall and possess richer social and political capital, which clearly has a positive effect on their children's NCAs development. Thus, this justifies the need to ensure that teachers' administrative positions are independent of class characteristics through the identification strategy of random student-classroom assignment.

Thirdly, we further test the robustness of our findings by adding control variables (Zhao and Chen 2023). The summary statistics of extra control variables in robustness tests are reported in Table A5 in the Appendix. In column (1) of Table A6 in the Appendix, to further overcome the estimation bias due to the correlation between teacher's administrative position and teacher quality, we add some control variables related to teacher quality, finding that the results remained robust. In addition, we add a series of control variables related to schools, households, and students in columns (2) to (4), respectively, and include all additional control variables simultaneously in column (5), and find that all of these results are generally consistent with the baseline regression.

Finally, in Table A7 in the Appendix, we standardize the comprehensive indicator of NCAs and its five sub-dimension indicators for robustness checks. The results show that the positive regression coefficients and their significance remain largely consistent with the results of the baseline regression after replacing the standardized NCA indicators. Combining the above multiple robustness tests, it is clear that the results of the baseline regression are more robust.

Heterogeneity analysis

First, we explore the heterogeneity by types of administrative positions. Different types of administrative positions have large differences in the workload and job characteristics, which

may also have different effects on students' NCAs. The results of Table A8 in the Appendix show that students' NCAs are significantly benefited by the headteachers who are also the school leaders or heads of teaching and research section. These two different positions' influence may result from several causes. In Chinese secondary schools, the head of teaching and research section is usually held by a teacher with sufficient teaching experience and advanced teaching skills (Zhong and Liu 2021), which may promote students to participate in learning more actively, and help improve their NCAs. School leaders often have more outstanding management skills and leadership, which are more likely to serve as demonstration effects among students (Zheng and Sun 2020), improving their NCAs.

Second, headteachers' administrative works only have significant effects among the headteachers with middle title. Specifically, columns (1) to (3) of Table A9 in the Appendix suggest that headteachers with middle title holding administrative tasks improve their students' NCAs by 0.0454 points. In contrast, headteachers with junior title are often inexperienced and may ignore the cultivation of students' noncognitive abilities. Headteachers with senior titles usually have to undertake heavier tasks in teaching (Zheng and Sun 2020; Zhong and Liu 2021), and are likely to pay more attention to students' academic performance rather than NCAs.

Third, we also test the heterogeneity caused by teachers' age and teaching experience, which may further affect teaching methods and the influence on students' abilities (Tran and Do 2022). The results of columns (4) to (7) in Table A9 show that administrative positions have relatively little impact on young headteachers as compared to middle-aged headteachers. Similarly, the effect is only significant among the headteachers with 15 years of teaching and above, but not among headteachers with less teaching experience. It can be explained that although the headteacher with an administrative position contributes to the development of students' NCAs, the lack of teaching experience of young teachers will partly offset this positive effect.

At last, we examine the differences caused by students' gender and household registration status (*Hukou*). It is a universal empirical fact across nations that there is a clear gender bias in the development of students' abilities, such as girls performing better in languages, and boys performing better in mathematics (Protivínský and Münich 2018). We discover that the administrative positions held by headteachers only significantly affect the NCAs of girls, as shown in columns (1) and (2) of Table A10 in the Appendix. The possible reason is that girls have more advantages in the development of languages (Protivínský and Münich 2018), while it is strongly related to the social skills and organizational skills, which belong to the part of NCAs. Moreover, the results of columns (3) and (4) indicate that urban-rural heterogeneity is significant. The positive effect of administrative positions is very limited among the rural students, but very obvious in urban students. The main reasons may be the lack of their own ability and the limitation of the quality of teachers (Zhao and Chen 2022).

Mechanisms

We further explore multiple potential mechanisms, including teacher and student levels and peer effects at the class level. The summary statistics of variables in mechanisms are shown in Table A11 in the Appendix.

Teachers' attitude and teaching method

Administrative positions may affect teachers in several ways. First, although administrative positions bring additional administrative work (Kim 2019), teachers are unlikely to compress teaching time, but rather trade off their teaching tasks and administrative duties (Chen and Zhao 2022; Zheng and Sun 2020). More importantly, administrative positions may provide teachers with some hidden benefits, such as increased job satisfaction (Klassen and Chiu 2010; Troesch and Bauer 2017). Second, considering that teachers with administrative positions have priority access to some resources in schools (García-Gallego et al. 2015; Sun and Zhou 2022), they are likely to use some of the more advanced information technology teaching equipment in their classes, which also has

Table 3. Administrative positions and teachers' attitude and teaching method.

	(1) Satisfaction	(2) Frequency of using the projector	(3) Frequency of using the internet	(4) Parents' respect
Administrative positions	0.1321***	0.2362***	0.1367***	0.0616***
	(0.0222)	(0.0221)	(0.0276)	(0.0183)
Baseline controls	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
N	8236	8286	7999	8251

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are presented in parentheses.

positive effects on students' NCAs (Shute, Ventura, and Ke 2015). Third, as mentioned earlier, the selection targets of administrative positions in Chinese schools are mainly teachers with better teaching quality. They are more respected and favoured by students and parents, which helps to form a kind of positive feedback and motivate teachers to improve their self-teaching requirements (Lam and Chan 2017). In addition, teachers with administrative positions have more authority in the minds of students and parents (Zheng and Sun 2020), which is more conducive to enhancing students' NCAs and promoting human capital development (Cheng and Zamarro 2018).

The results of Table 3 show that administrative positions have positive effects on teachers' job satisfaction, the frequency of using advanced equipment in teaching, and the respect of students' parents. On the one hand, improved job satisfaction and more respect from parents will motivate teachers to improve their self-teaching requirements and care about their students (Braun, Schonert-Reichl, and Roeser 2020), which may further benefit students' NCA. On the other hand, more advanced and enriched teaching methods have a positive effect on improving students' NCAs by increasing student participation and increasing teacher-student interaction in class (Shute, Ventura, and Ke 2015).

Students' sense of belonging to the class

Several studies confirmed that students who reported a greater sense of belonging to schools and classes tended to be more motivated, confident, and successful in their academic endeavours (Goodenow and Grady 1993). Teachers with administrative positions may increase students' sense of belonging to their schools and classes by creating a better classroom climate, higher quality classroom management, a more prominent disciplinary climate, and better instructional equipment (Li 2018). In Table 4, the results prove that head-teachers with administrative positions help promote students to participate more actively in class activities and like the class and classmates. And one possible explanation is that in China, the education departments of local governments still advocate quality education, that is, not only focusing on students' grades, but also on the development of their NCAs. However, due to the fierce competition in the senior high school entrance examination, most teachers still focus more on students' grades, and ignore the development of their NCAs. Differently, teachers with administrative positions are more likely to abide by the regulations and initiatives of the education departments, balancing students' academic achievements with the development of their NCAs.

Peer effects

Prior studies supported the impact of class and grade variables on students' NCAs (Gong, Lu, and

Table 4. Administrative positions and students' sense of belonging to the class.

	(1) Frequency of participating in class activities	(2) Liking the class	(3) Liking classmates
Administrative positions	0.1032***	0.0835***	0.0584**
	(0.0333)	(0.0286)	(0.0271)
Baseline controls	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes
N	8286	8260	8286

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are presented in parentheses.

Table 5. Administrative positions and peer effects.

	(1) Class average NCAs	(2) Friends skipping class	(3) Friends fighting	(4) Friends violating school rules
Administrative positions	0.0358***	0.0083	0.0036	0.0063
	(0.0031)	(0.0117)	(0.0133)	(0.0121)
Baseline controls	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
N	8286	8212	8223	8223

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are presented in parentheses.

Song 2021; Neidell and Waldfogel 2010), while teachers play an important role in the peer effects in schools (e.g. Duflo, Dupas, and Kremer 2011; Feng and Li 2016). Moreover, as classroom supervisors, headteachers have additional methods to address the peer effect in their classes, such as contacting parents, arranging student seating, appointing class officers, and organizing classroom activities (Feng and Li 2016). Considering that teachers with administrative positions may be more active in taking some measures to influence the class peer effects, we further test whether teachers' administrative positions influence students' NCAs through class peer effects. In Table 5, we examine peer effects from two aspects: class average NCAs (except for the observed student) and the bad behaviour of the student's friends. Results reveal that although headteachers' administrative positions have no discernible impact on the negative behaviours of students' peers, it significantly improves the average NCAs of students in the class. According to previous research, the improvement of the overall NCAs of the class is helpful to the development of individual NCAs (Shure 2021).

VI. Conclusion

As an important component of human capital, NCAs have received attention from various social science fields. Although many studies have explored the influence of teachers on students' NCAs (e.g. Araujo et al. 2016; Kraft 2019; Rouse and Hunziker 2020), few studies focused on the impact of teachers' administrative positions. Therefore, we use nationwide survey data from China to explore the causal effects of teachers' administrative positions on students' NCAs. For the identification of causality, we primarily use a quasi-natural experiment of random student – classroom assignment to ensure a randomized match between teachers and students.

Based on the empirical results, we draw several conclusions. First, by constructing a composite NCA indicator, baseline results suggest that headteachers' administrative positions significantly improve students' NCAs by 0.0278 points, and have a significant positive effect on the subdimensions of extraversion and neuroticism in particular. Main results are proved to be robust after a series of tests. Second, heterogeneity analysis shows that teachers with higher administrative positions, middle titles, middle-aged, and long teaching years held administrative positions that are more conducive to students' NCAs. Girl students and urban students benefited more from headteachers' administrative positions than boys and rural students. Third, the results of the mechanism analysis show that headteachers' administrative positions significantly contribute to students' NCAs by improving their attitudes and teaching methods, increasing students' sense of belonging to the class, and enhancing positive peer effects in classes.

According to our knowledge, this study is among the first to explore the causal effect of headteachers' administrative positions on students' NCAs in China. Our research focuses on the positive impact of teachers' administrative positions on the personality traits of students, particularly as classroom managers of headteachers. This study supplements the research field on the impact of teacher characteristics on student NCAs through a Chinese case. Moreover, our findings differ from previous studies that emphasize the negative effects of teachers' administrative positions on students (Chen and Zhao 2022; Sun and Zhou 2022). Instead, our study confirms the positive effect of headteacher administrative positions on students'

NCAs, suggesting that the role of teachers' administrative positions on student development needs to be viewed more holistically.

The following policy implications can be drawn from our findings. For schools, to balance students' academic performance with the development of NCAs, a more flexible way can be adopted to assign the headteacher. In addition, heterogeneity analysis shows that the NCAs of rural students are not significantly affected by headteachers' administrative positions. The main reasons may be the lack of their own ability and the limitation of the quality of teachers. Therefore, it is suggested that the government can promote the improvement of rural students' NCAs by optimizing the allocation of teachers. Lastly, it can be found that administrative positions can help improve teachers' job satisfaction and promote their better care of students. Based on this finding, local governments can implement some teacher incentive policies and increase teacher bonuses to improve their satisfaction, promoting the development of students' NCAs. Furthermore, even with the above positive effects, it is also important to note that the workload of administrative positions should be moderate, avoiding excessive burdens on teachers.

Acknowledgments

We acknowledge the National Survery Research Center of Renmin University of China for providing the China Education Panel Survey (CEPS) data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the National Natural Science Foundation of China [72303090; 72304102]; Jiangsu Education Science Planning Project [C/2022/01/45]; Philosophy and Social Science Program Discipline Co-construction Project of Guangdong Province [GD23XYJ68].

ORCID

Boou Chen http://orcid.org/0000-0003-1735-7535
Chunkai Zhao http://orcid.org/0000-0003-1918-9598

Data availability statement

The authors do not have permission to share data.

Ethical statement

This paper does not contain any studies with human participants performed by any of the authors. The data we used in this study come from the China Education Panel Survey (CEPS). This survey and data are publicly released.

References

Almlund, M., A. L. Duckworth, J. Heckman, and T. Kautz. 2011. "Personality Psychology and Economics." In *Handbook of the Economics of Education*, 1–181. Elsevier.

Araujo, M. C., P. Carneiro, Y. Cruz-Aguayo, and N. Schady. 2016. "Teacher Quality and Learning Outcomes in Kindergarten." *The Quarterly Journal of Economics* 131 (3): 1415–1453. https://doi.org/10.1093/qje/qjw016.

Ballet, K., G. Kelchtermans, and J. Loughran. 2006. "Beyond Intensification Towards a Schotarship of Practice: Analysing Changes in teachers' Work Lives." *Teachers & Teaching* 12 (2): 209–229. https://doi.org/10.1080/13450600500467415.

Bhai, M., and I. Horoi. 2019. "Teacher Characteristics and Academic Achievement." *Applied Economics* 51 (44): 4781–4799. https://doi.org/10.1080/00036846.2019.1597963.

Blanden, J., P. Gregg, and L. Macmillan. 2007. "Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education." *The Economic Journal* 117 (519): C43–C60. https://doi.org/10.1111/j.1468-0297.2007.02034.x.

Blazar, D., and M. A. Kraft. 2017. "Teacher and Teaching Effects on students' Attitudes and Behaviors." *Educational Evaluation and Policy Analysis* 39 (1): 146–170. https://doi.org/10.3102/0162373716670260.

Bowles, S., H. Gintis, and M. Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature* 39 (4): 1137–1176. https://doi.org/10.1257/jel.39.4.1137.

Braun, S. S., K. A. Schonert-Reichl, and R. W. Roeser. 2020. "Effects of teachers' Emotion Regulation, Burnout, and Life Satisfaction on Student Well-Being." *Journal of Applied Developmental Psychology* 69:101151. https://doi.org/10.1016/j.appdev.2020.101151.

Cheng, A., and G. Zamarro. 2018. "Measuring Teacher Non-Cognitive Skills and Its Impact on Students: Insight from the Measures of Effective Teaching Longitudinal Database." *Economics of Education Review* 64:251–260. https://doi.org/10.1016/j.econedurev.2018.03.001.

Chen, B., and C. Zhao. 2022. "More is Less: Homeroom teachers' Administrative Duties and students' Achievements in China." *Teaching and Teacher Education* 119:103857. https://doi.org/10.1016/j.tate.2022.103857.

Cobb-Clark, D. A., and S. Schurer. 2012. "The Stability of Big-Five Personality Traits." *Economics Letters* 115 (1): 11–15. https://doi.org/10.1016/j.econlet.2011.11.015.

Costa, P. T., and R. R. McCrae. 1992. "Four Ways Five Factors are Basic." *Personality and Individual Differences* 13 (6): 653–665. https://doi.org/10.1016/0191-8869(92)90236-I.

DeAngelis, C. A. 2021. "Divergences Between Effects on Test Scores and Effects on Non-Cognitive Skills." *Educational Review* 73 (4): 503–514. https://doi.org/10.1080/00131911.2019.1646707.

Dello-Iacovo, B. 2009. "Curriculum Reform and 'Quality education' in China: An Overview." *International Journal of Educational Development* 29 (3): 241–249. https://doi.org/10.1016/j.ijedudev.2008.02.008.

Doherty, J., and M. Conolly. 1985. "How Accurately Can Primary School Teachers Predict the Scores of Their Pupils in Standardised Tests of Attainment? A Study of Some Non-Cognitive Factors That Influence Specific Judgements." *Educational Studies* 11 (1): 41–60. https://doi.org/10.1080/0305569850110105.

Duflo, E., P. Dupas, and M. Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review* 101 (5): 1739–1774. https://doi.org/10.1257/aer.101.5.1739 .

Feng, H., and J. Li. 2016. "Head Teachers, Peer Effects, and Student Achievement." *China Economic Review* 41:268–283. https://doi.org/10.1016/j.chieco.2016.10.009.

García-Gallego, A., N. Georgantzís, J. Martín-Montaner, and T. Pérez-Amaral. 2015. "(How) Do Research and Administrative Duties Affect University professors' Teaching?" *Applied Economics* 47 (45): 4868–4883. https://doi.org/10.1080/00036846.2015.1037438.

Gerrard, J., and L. Farrell. 2014. "Remaking The Professional Teacher: Authority And Curriculum Reform." *Journal of Curriculum Studies*, 46 (5): 634–655. https://doi.org/10.1080/00220272.2013.854410.

Gershenson, S. 2016. "Linking Teacher Quality, Student Attendance, and Student Achievement." *Education Finance and Policy* 11 (2): 125–149. https://doi.org/10.1162/EDFP_a_00180.

Gong, J., Y. Lu, and H. Song. 2018. "The Effect of Teacher Gender on students' Academic and Noncognitive Outcomes." *Journal of Labor Economics* 36 (3): 743–778. https://doi.org/10.1086/696203.

Gong, J., Y. Lu, and H. Song. 2021. "Gender Peer Effects on students' Academic and Noncognitive Outcomes: Evidence and Mechanisms." *Journal of Human Resources* 56 (3): 686–710. https://doi.org/10.3368/jhr.56.3.0918-9736R2 .

Goodenow, C., and K. E. Grady. 1993. "The Relationship of School Belonging and friends' Values to Academic Motivation Among Urban Adolescent Students." *Journal of Experimental Education* 62 (1): 60–71. https://doi.org/10.1080/00220973.1993.9943831.

Heckman, J. J., and Y. Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review* 91 (2): 145–149. https://doi.org/10.1257/aer.91.2.145.

Horoi, I., and M. Bhai. 2018. "New Evidence on National Board Certification as a Signal of Teacher Quality." *Economic Inquiry* 56 (2): 1185–1201. https://doi.org/10.1111/ecin.12525.

Hu, F. 2018. "Migrant Peers in the Classroom: Is the Academic Performance of Local Students Negatively Affected?" *Journal of Comparative Economics* 46 (2): 582–597. https://doi.org/10.1016/j.jce.2017.11.001.

Jackson, C. K. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes." *Journal of Political Economy* 126 (5): 2072–2107. https://doi.org/10.1086/699018.

Kashy-Rosenbaum, G., O. Kaplan, and Y. Israel-Cohen. 2018. "Predicting Academic Achievement by Class-Level Emotions and Perceived Homeroom teachers' Emotional Support." *Psychology in the Schools* 55 (7): 770–782. https://doi.org/10.1002/pits.22140.

Kim, K. N. 2019. "Teachers' Administrative Workload Crowding Out Instructional Activities." *Asia Pacific Journal of Education* 39 (1): 31–49. https://doi.org/10.1080/02188791.2019.1572592.

Klassen, R. M., and M. M. Chiu. 2010. "Effects on teachers' Self-Efficacy and Job Satisfaction: Teacher Gender, Years of Experience, and Job Stress." *Journal of Educational Psychology* 102 (3): 741. https://doi.org/10.1037/a0019237.

Kraft, M. A. 2019. "Teacher Effects on Complex Cognitive Skills and Social-Emotional Competencies." *Journal of Human Resources* 54 (1): 1–36. https://doi.org/10.3368/jhr.54.1.0916.8265R3.

Lam, Y. Y., and J. C. Y. Chan. 2017. "Effects of Social Persuasion from Parents and Teachers on Chinese students' Self-Efficacy: An Exploratory Study." *Cambridge Journal of Education* 47 (2): 155–165. https://doi.org/10.1080/0305764X.2016.1143448.

Li, Y. 2018. "A Study on Job Satisfaction of "Double-shouldered" Teachers in Primary and Secondary Schools." *Journal of Teaching and Management* 28 (4): 31–33.

Loyalka, P., S. Sylvia, C. Liu, J. Chu, and Y. Shi. 2019. "Pay by Design: Teacher Performance Pay Design and the Distribution of Student Achievement." *Journal of Labor Economics* 37 (3): 621–662. https://doi.org/10.1086/702625.

Neidell, M., and J. Waldfogel. 2010. "Cognitive and Noncognitive Peer Effects in Early Education." *Review of Economics and Statistics* 92 (3): 562–576. https://doi.org/10.1162/REST_a_00012.

Okumura, S. 2017. "Homeroom Teachers or Specialist Teachers?: Considerations for the Workforce for Teaching English as a Subject at Elementary Schools in Japan." *Asian Journal of Education and Training* 3 (1): 1–5. https://doi.org/10.20448/journal.522/2017.3.1/522.1.1.5.

Protivínský, T., and D. Münich. 2018. "Gender Bias in teachers' Grading: What is in the Grade." *Studies in Educational Evaluation* 59:141–149. https://doi.org/10.1016/j.stueduc.2018.07.006.

Rouse, K., and B. Hunziker. 2020. "Child Bodyweight and Human Capital: Test Scores, Teacher Assessments and Noncognitive Skills." *Economics of Education Review* 79:102042. https://doi.org/10.1016/j.econedurev.2020.102042.

Shure, N. 2021. "Non-Cognitive Peer Effects in Secondary Education." *Labour Economics* 73:102074. https://doi.org/10.1016/j.labeco.2021.102074.

Shute, V. J., M. Ventura, and F. Ke. 2015. "The Power of Play: The Effects of Portal 2 and Lumosity on Cognitive and Noncognitive Skills." *Computers & Education* 80:58–67. https://doi.org/10.1016/j.compedu.2014.08.013.

Sun, Y., and X. Zhou. 2022. "The Effect of Teacher's Concurrent Administrative Position on students' Academic Outcomes: Evidence and Mechanisms." *China Economic Review* 71:101735. https://doi.org/10.1016/j.chieco.2021.101735.

Tran, T. T. T., and T. X. Do. 2022. "Student Evaluation of Teaching: Do Teacher Age, Seniority, Gender, and Qualification Matter?" *Educational Studies* 48 (4): 443–470. https://doi.org/10.1080/03055698.2020.1771545.

Troesch, L. M., and C. E. Bauer. 2017. "Second Career Teachers: Job Satisfaction, Job Stress, and the Role of Self-Efficacy." *Teaching and Teacher Education* 67:389–398. https://doi.org/10.1016/j.tate.2017.07.006.

West, M. R., M. A. Kraft, A. S. Finn, R. E. Martin, A. L. Duckworth, C. F. Gabrieli, and J. D. Gabrieli. 2016. "Promise and Paradox: Measuring students' Non-Cognitive Skills and the Impact of Schooling." *Educational Evaluation and Policy Analysis* 38 (1): 148–170. https://doi.org/10.3102/0162373715597298.

Zhang, J., and S. N. Pang. 2016. "Investigating the Development of Professional Learning Communities: Compare Schools in Shanghai and Southwest China." *Asia Pacific Journal of Education* 36 (2): 217–230. https://doi.org/10.1080/02188791.2016.1148851.

Zhao, C., and B. Chen. 2022. "Parental Migration and Non-Cognitive Abilities of Left-Behind Children in Rural China: Causal Effects by an Instrumental Variable Approach." *Child Abuse & Neglect* 123:105389. https://doi.org/10.1016/j.chiabu.2021.105389.

Zhao, C., and B. Chen. 2023. "ICT in Education Can Improve students' Achievements in Rural China: The Role of Parents, Educators and Authorities." *Journal of Policy Modeling* 45 (2): 320–344. https://doi.org/10.1016/j.jpolmod.2023.02.007.

Zhao, C., and X. Li. 2022. "Living Under the Shadow: Adverse Childhood Experiences and Entrepreneurial Behaviors in Chinese Adults." *Journal of Business Research* 138:239–255. https://doi.org/10.1016/j.jbusres.2021.09.016.

Zheng, L., and L. Sun. 2020. "Does the Part-Time Administrative Weaken the Teaching Work? A PSM Estimate Based on the Data of 730 Junior High School Teachers." *Contemporary Education Sciences* 35 (11): 52–58.

Zhong, J., and Y. Liu. 2021. "The Paradox of Reducing teachers' Burden: The Dilemma of de-Professionalization and the Trap of Re-Specialization." *Tsinghua Journal of Education* 42 (6): 80–90.

Appendix

Table A1. Summary statistics of predetermined background variables at class-level in balancing tests.

Variable	Definition	N	Mean	S.D.
Average teaching experience	Average years of teaching experience of teachers in the class	344	15.57	5.919
Teachers with senior title	Number of teachers with senior (or professor) titles in the class (0 = none; 1 = one and more)	344	0.360	0.481
Grade 9	0 = 7th Grade; 1 = 9th Grade	344	0.494	0.501
Class size	Total number of students	344	43.96	13.98
Ratio of boys		344	0.489	0.121
Percentage of local students		344	0.783	0.228
Percentage of rural students		344	0.490	0.305
Average education expenditure	Average education expenditure of the class (yuan), as a logarithm	344	5.641	2.221
Percentage of tutoring	Percentage of students attending tutoring in the class	344	0.324	0.287
Proportion of fathers working in the public sector	Percentage of students whose fathers are leaders or staff in government or public institutions in the total number of students in the class	344	0.050	0.077
Proportion of mothers working in the public sector	Percentage of students whose mothers are leaders or staff in government or public institutions in the total number of students in the class	344	0.032	0.060
Percentage of students who had difficulty learning Math in grade 6	The percentage of students with learning difficulties in Math before entering junior high school (i.e. grade 6) in the class	344	0.097	0.096
Percentage of students who had difficulty learning Chinese in grade 6	The percentage of students with learning difficulties in Chinese before entering junior high school (i.e. grade 6) in the class	344	0.029	0.041
Percentage of students who had difficulty learning English in grade 6	The percentage of students with learning difficulties in English before entering junior high school (i.e. grade 6) in the class	344	0.153	0.131

Table A2. T-test for class-level variables in classes where the headteacher holds an administrative position.

	Full sample (Classes where the headteacher holds an administrative position, $N = 79$)		
	Non-random assignment Mean	Random assignment Mean	T-test Mean Diff.
Average teaching experience	18.383	15.957	2.427
Teachers with senior title	0.400	0.290	0.110
Grade 9	0.400	0.594	−0.194
Class size	49.300	41.188	8.112*
Ratio of boys	0.513	0.479	0.033
Percentage of local students	0.891	0.784	0.107
Percentage of rural students	0.502	0.529	−0.027
Average education expenditure	5.466	5.073	0.393
Percentage of tutoring	0.317	0.252	0.065
proportion of fathers working in the public sector	0.076	0.038	0.038*
proportion of mothers working in the public sector	0.059	0.021	0.038**
Percentage of students who had difficulty learning Math in grade 6	0.050	0.113	−0.063*
Percentage of students who had difficulty learning Chinese in grade 6	0.020	0.027	−0.007
Percentage of students who had difficulty learning English in grade 6	0.155	0.171	−0.016
N	10	69	–

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively.

Table A3. Balancing tests.

Class-level variables	(1) Average teaching experience	(2) Teachers with senior title	(3) Grade 9	(4) Class size	(5) Ratio of boys	(6) Percentage of local students	(7) Percentage of rural students
Panel A. Classes with random assignment							
Administrative positions	1.1591	−0.0316	0.1857**	−0.8882	0.0029	−0.0033	−0.0048
	(0.8419)	(0.0696)	(0.0907)	(1.0311)	(0.0175)	(0.0183)	(0.0211)
N	344	344	344	344	344	344	344
Panel B. Classes with non-random assignment							
Administrative positions	3.0065*	0.0779	−0.1299	4.1299**	0.0456	−0.0077	−0.0965
	(1.7123)	(0.1644)	(0.2304)	(1.6844)	(0.0326)	(0.0261)	(0.0631)
N	62	62	62	62	62	62	62

Class-level variables	(8) Average education expenditure	(9) Percentage of tutoring	(10) Proportion of fathers working in the public sector	(11) Proportion of mothers working in the public sector	(12) Percentage of students who had difficulty learning Math in grade 6	(13) Percentage of students who had difficulty learning Chinese in grade 6	(14) Percentage of students who had difficulty learning English in grade 6
Panel A. Classes with random assignment							
Administrative positions	0.1340	−0.0061	0.0017	−0.0013	0.0166	−0.0077	0.0208
	(0.1782)	(0.0213)	(0.0070)	(0.0050)	(0.0133)	(0.0059)	(0.0164)
N	344	344	344	344	344	344	344
Panel B. Classes with non-random assignment							
Administrative positions	−0.0469	−0.0107	0.0554***	0.0452***	−0.0233	−0.0047	0.0130
	(0.2579)	(0.0544)	(0.0183)	(0.0154)	(0.0256)	(0.0102)	(0.0463)
N	62	62	62	62	62	62	62

Robust standard errors are presented in parentheses. ***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. School fixed effects are controlled for each column of results in each panel.

Table A4. Robustness tests by eliminating some special groups and comparing to non-randomly assigned classes.

	(1) Eliminating students from highly educated families	(2) Eliminating students from extremely poor and wealthy families	(3) Eliminating unhealthy students	(4) Eliminating schools in distant places	(5) Eliminating schools without headteachers holding administrative positions	(6) Eliminating private schools	(7) Eliminating classes with non-headteacher holding administrative positions	(8) Using a sample of classes with non-random assignment
Administrative positions	0.0257**	0.0225**	0.0282***	0.0380***	0.0264**	0.0333***	0.0372***	0.0541**
	(0.0109)	(0.0110)	(0.0109)	(0.0125)	(0.0114)	(0.0112)	(0.0124)	(0.0259)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	8107	7937	7972	6978	4195	7460	7128	1728

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are presented in parentheses.

Table A5. The summary statistics of extra variables in robustness tests.

Variable	Definitions and description	Obs.	Mean	S.D.
Panel A. Teacher quality-related characteristics				
Normal university	Whether the headteacher graduated from a normal university or major (0 = no; 1 = yes)	8286	0.929	0.257
Teaching awards from the education department	Whether the headteacher has received teaching awards from the county education department or above (0 = no; 1 = yes)	8286	0.815	0.389
Teaching awards from the school	Whether the headteacher has received teaching awards from the school (0 = no; 1 = yes)	8286	0.818	0.386
Panel B. School characteristics				
Public education finance	The logarithm of public finance spending on education per student	8063	6.422	1.391
Computer in classroom	Whether each classroom is equipped with computers (0 = no; 1 = yes)	8286	0.688	0.463
Percentage of headteachers holding administrative positions	Percentage of headteachers who concurrently hold administrative positions (%)	8286	19.44	22.02
Location of school	The distance from the school to the centre of the city (1 = extremely close; 2 = somewhat close; 3 = average; 4 = far; 5 = very far)	8286	2.638	1.534
School ranking	School ranking in the district or county (1 = being the worst; 2 = not good; 3 = common; 4 = good; 5 = being the best)	8286	4.039	0.789
Rebuilding of school	whether or not the school was reconstructed in the last five years (0 = no; 1 = yes)	8276	0.572	0.495
Panel C. Household characteristics				
Location of home	The distance from the student's home to the centre of the city (1 = extremely close; 2 = somewhat close; 3 = average; 4 = far; 5 = very far)	8207	3.158	1.766
Home surroundings	Self-rated home surroundings (0 = not good or worse; 1 = good)	8286	0.680	0.467
Home computer	There is a computer at home (0 = no; 1 = yes)	8286	0.679	0.467
Internet access	Availability of Internet at home (0 = no; 1 = yes)	8286	0.593	0.491
Father's job	0 = farmers, ordinary workers, self-employed, and unemployed; 1 = employees of national businesses and institutions, enterprise managers, engineers, doctors, teachers, and attorneys	8286	0.190	0.392
Panel D. Student characteristics				
Only child	0 = no; 1 = yes	8286	0.416	0.493
Rank among children	Rank among children in the family	8286	1.351	0.532
Kindergarten	= 1 if the student has been to kindergarten before; 0 otherwise	8286	0.798	0.401
Cognitive ability scores	Scores on cognitive abilities that are standardized for students	8286	0.046	0.852
Boarding at school	0 = no; 1 = yes	8286	0.324	0.468

Table A6. Robustness tests by adding control variables.

	(1) Adding teacher quality-related variables	(2) Adding school-related variables	(3) Adding household-related variables	(4) Adding student-related variables	(5) Adding all control variables
Administrative positions	0.0283***	0.0272**	0.0250**	0.0312***	0.0282***
	(0.0108)	(0.0109)	(0.0108)	(0.0107)	(0.0109)
Baseline controls	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes
N	8286	8053	8207	8286	7976

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are presented in parentheses.

Table A7. Robustness tests by standardized NCA indicators.

	(1) NCA (std)	(2) Conscientiousness (std)	(3) Agreeableness (std)	(4) Extraversion (std)	(5) Openness (std)	(6) Neuroticism (std)
Administrative positions	0.0780**	0.0441	−0.0161	0.1014***	0.0212	0.0652*
	(0.0303)	(0.0312)	(0.0345)	(0.0330)	(0.0324)	(0.0350)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	8286	8286	8286	8286	8286	8286

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are presented in parentheses. NCA (std), Conscientiousness (std), Agreeableness (std), Extraversion (std), Openness (std), and Neuroticism (std) are the standardized scores of NCA and each sub-dimension indicator in the baseline regression. The standardized score of each indicator has a mean of 0 and a standard deviation of 1.

Table A8. Heterogeneity effects by different positions.

	(1)	(2)	(3)
Grade team leader	0.0117		
	(0.0143)		
Head of teaching and research section		0.0595**	
		(0.0233)	
School leaders			0.0407**
			(0.0183)
Baseline controls	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes
N	7443	6989	7178

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are presented in parentheses.

Table A9. Heterogeneity effects by the teacher characteristics.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Title			Age		Teaching experience	
	Junior	Middle	Senior	Young	Middle-aged	Less than 15 years	15 years and above
Administrative positions	−0.0179	0.0454**	−0.0060	0.0169	0.0394**	0.0057	0.0475***
	(0.0219)	(0.0231)	(0.0599)	(0.0205)	(0.0158)	(0.0179)	(0.0167)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2971	3938	1377	3575	4711	3921	4365

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are presented in parentheses.

Table A10. Heterogeneity effects by students' gender and *Hukou*.

	(1)	(2)	(3)	(4)
	Gender		*Hukou*	
	Girl	Boy	Rural	Urban
Administrative positions	0.0452***	0.0072	0.0168	0.0418**
	(0.0143)	(0.0166)	(0.0141)	(0.0175)
Baseline controls	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
N	4218	4068	4658	3628

***, **, * indicate significance at the levels of 1%, 5%, and 10%, respectively. Robust standard errors are presented in parentheses.

Table A11. The summary statistics of variables in mechanism analysis.

Variable	Definition and description	Obs.	Mean	S.D.
Satisfaction	Satisfaction with the headteacher work (1 = very dissatisfied; 2 = dissatisfied; 3 = common; 4 = satisfied; 5 = very satisfied)	8236	3.576	0.818
Frequency of using the projector	1 = never; 2 = occasionally; 3 = sometimes; 4 = often; 5 = always	8286	3.760	0.978
Frequency of using the internet	1 = never; 2 = occasionally; 3 = sometimes; 4 = often; 5 = always	7999	2.518	1.062
Parents' respect	What percentage of parents in the class adhere to the headteacher (1 = almost none; 2 = few; 3 = about half; 4 = most; 5 = all)	8251	4.473	0.669
Frequency of participating in class activities	1= never; 2 = occasionally; 3 = frequently; 4 = very frequently	8286	2.791	1.008
Liking the class	1 = disagree; 2 = partially disagree; 3 = agree; 4 = strongly agree	8260	3.192	0.852
Liking classmates	1 = disagree; 2 = partially disagree; 3 = agree; 4 = strongly agree	8286	3.307	0.779
Class average NCAs	The class average NCA other than the student observed	8286	3.264	0.167
Friends skipping class	1= never; 2 = occasionally; 3 = frequently	8212	1.083	0.318
Friends fighting	1= never; 2 = occasionally; 3 = frequently	8223	1.116	0.355
Friends violating school rules	1= never; 2 = occasionally; 3 = frequently	8223	1.110	0.351