

第三方数字平台能否帮助中小微企业提升经营收益?

——来自百万商户大数据的证据

熊巧琴 汤珂 张丰羽*

摘要: 本文基于某大型流程数字化商贸平台的百万商户大数据, 利用倾向性得分匹配方法, 为服务行业中小微企业通过第三方数字平台开展流程数字化的经营效益提供因果推断与异质性分析。研究发现, 相比同等条件非数字化商户, 数字化商户年均销量和金额均增加。三、四线城市, 企业主为中老年, 以及拥有一定数字经济资源禀赋的中等数字城市企业, 营收提升更明显。本文为服务业中小微企业的流程数字化与缩减数字鸿沟提供政策参考。

关键词: 流程数字化平台; 中小微企业; 倾向性得分匹配

DOI: 10.13821/j.cnki.ceq.2023.05.04

一、引言

习近平总书记指出应以信息化培育新动能, 用新动能推动新发展。^① 作为“国民经济和社会发展的生力军”^②, 中小微企业实现线上和实体经营的融合和重构, 成为适应后疫情时代、推动就业和经济形势的重要议题之一。虽然我国自古有“行商坐贾”和“运筹握算”的流程优化文化, 强调信息传播、口碑经营和精细管理的重要性; 但受技术和规模限制, 商户获取的收益有限。如今, 企业普遍使用信息、计算、沟通和连接等数字技术来优化生产、管理和销售等活动 (BarNir et al., 2003), 极大提高传统中小微商企流程优化的效率, 亦称“流程数字化” (business process digitization)。伴随数字技术的迭代, 基于第三方数字平台和数据的流程创新与价值交互, 成为流程数字化的新特征 (Karimi and Walter, 2015), 也是加快形成“供需互促、产销并进”良性循环的重要渠道。

服务业中小微企业科技创新投入和研发较弱, 普遍依赖低价、低技术和附加值发展路径, 难以独立数字化。第三方数字平台广泛触达中小微企业网络和大众口碑, 能为中小微企业提供更优的技术能力、客户流量、数据资源和规模优势, 从而在发掘顾客需求和细分市场、辅助决策和产品服务创新等方面产生重要价值。中小微企业借助第三方数字平台整合多方数字技术和资源 (De Reuver et al., 2018), 有利于帮助单个企业填补市

* 熊巧琴, 深圳大学经济学院; 汤珂, 清华大学社会科学学院经济学研究所; 张丰羽, 清华大学经济管理学院、清华大学中国现代国有企业研究院。通信作者及地址: 汤珂, 北京海淀区清华大学明斋, 100084; 电话: 13466777332; E-mail: ketang@mail.tsinghua.edu.cn。本文得到中宣部“四个一批人才”暨哲学社会科学领军人才项目和国家自然科学基金重大项目 (72192802) 资助。感谢两位匿名审稿人的建设性修改意见, 文责自负。

① 2018年4月20日至21日, 习近平在全国网络安全和信息化工作会议上发表的重要讲话。

② 2018年8月20日国务院促进中小微企业发展工作领导小组第一次会议。

场的结构性空缺和个人数字能力的相对不足，实现资源聚合、组织合作、生态改良和效益改善 (Li et al., 2018)。

服务业中小微企业利用平台赋能流程数字化，一方面，利用大数据实现产品差异化、生产智能化和产销目标化，有利于缓解信息不对称、提升产品质量、扩大市场范围和提高生产效率，即“数字红利”(陈剑等，2020)。但另一方面，不同地区、行业和信息化基础设施等的禀赋差别又影响企业数字化红利水平，甚至进一步扩大差距，即“数字鸿沟”(Hoffman et al., 2000)。中小微企业基于数字平台参与流程数字化，究竟是扩大鸿沟还是增进红利？现有研究并未得到一致的结论。使用大数据和反事实(counterfactual)因果框架，准确评估第三方流程数字化商贸平台对服务业中小微企业的经营影响及其异质性研究也付之阙如。

本文的百万商户大数据源于某知名第三方数字化商贸平台，它主要提供营销数字化(通过APP、小程序或网页提供折扣广告、位置推送、优质产品推荐等)、生产和管理数字化(智能点餐、支付结算、售后点评、财务记账、数据分析等)等服务。本文在控制企业主、企业和企业所在地基础上，运用倾向性得分匹配法(propensity score matching, PSM)，测算出2018—2019年使用该平台给服务业中小微企业增加15.53%的交易量和11.85%的销售金额。同时，通过对平台效应在不同行业、区域、地形地势、企业主年龄和数字经济水平的异质性分析和稳健性检验，发现处于数字化早中期阶段的数字经济中等发达城市，经济发展三、四线城市，以及企业主为中老年的企业，参与流程数字化平台获得的营收提升更明显；相对平原地区，山区零售业企业的数字化效益有待提高。

本文的贡献包括：第一，首次基于百万级别大数据，通过PSM，辅助双重差分(倾向评分)匹配(PSM-DID)等多种稳健性检验，构造近似随机对照实验，精确测度基于平台开展流程数字化对服务业中小微企业的营收影响及其异质性。第二，本文的大数据具有准确记录企业大部分经营活动的优势，有效避免小样本问卷调查、统计数据的主观偏差(Kleinjans and Soest, 2014)或代表性不足等问题，为国民经济测算提供新思路。第三，基于平台流程数字化的实证研究，不仅丰富了数字化和区域经济、数字鸿沟关系的理论研究，在引导第三方数字平台赋能中小微企业数字转型、促进就业、深化数字化布局等实践层面也有较高的指导价值。

本文安排如下：第二部分为文献综述和研究假设；第三部分介绍数据和研究方法；第四部分为服务业各行业参与流程数字化平台与营收关系的实证研究；第五部分为稳健性检验、异质性和潜在作用机制分析；最后为总结与政策建议。

二、文献综述

关于数字化、信息化与企业营收关系的文献尚未得出统一结论。有学者认为数字化及其投资可通过降本提效、促进创新和出口、降低融资约束、优化组织效率等，直接提高公司主营业务收入和创新成果(Bharadwaj, 2000; 陈剑等, 2020)，或通过优化内外部交流、开放式创新等间接正向调节企业营收(Sorescu, 2017)。也有研究支持“数字化悖论”，认为信息化投资与企业利润、资本回报之间并无稳定联系(Loveman, 1994)。

流程数字化是数字化的子类别,指企业利用数字技术支持或补充现有流程,包括生产、管理和销售数字化(Bharadwaj et al., 2013; Adomako et al., 2021)。随着数字技术迭代,基于第三方数字平台和数据进行流程数字化成为主流(Karimi and Walter, 2015)。服务业中小微企业受规模和技术限制,独立开展流程数字化风险较高;依托规模化和网络化的第三方数字平台,能更好连接、整合多方数据与资源共创价值(Gummesson and Mele, 2010)。例如,整合线上线下营销渠道,避免纯粹在线业务导致信任缺失(Oh et al., 2012),帮助企业进行精准营销,完善动态定价(Cohen et al., 2018)。通过生产和管理数字化,利用数据分析创造价值,改善企业的产品和供应链技术创新(Lee et al., 2018),实现差异化和人性化管理(傅颖等, 2021)等。

以上研究增进了对数字平台、流程数字化和企业发展关系的理解,但多为定性研究,或使用访谈、上市公司、百强企业等小样本数据。不仅难以排除内生性影响,其普适性和差异性也需实证补充。本文使用服务行业百万中小微企业的经营大数据,提出假设:

假设 1 企业通过第三方数字平台进行流程数字化,有利于提升服务业中小微企业的营收。

数字技术给全人类带来福祉,也造成社会分化和数字鸿沟^①。腾讯发布的《2020 数字中国指数报告》显示我国数字化发展水平不均衡。罗裕梅和凌鸿(2014)从网民、网页和宽带接入等方面,指出网络信息消费在我国头部与后线城市等不同区域存在较大数字鸿沟。第三产业是山区和后线城市除农业之外的重要经济抓手,产业数字化是它们实现弯道超车的有效依托渠道(杨旭和李竣, 2017);但山区和后线城市服务企业多呈现小、散、乱的分布特征(周滔等, 2003),在营销、生产和管理流程中能充分利用的数字技术和手段有限。依托第三方数字平台的信息整合、规模宣传和线上口碑运维等优势,商户可降低消费者搜索成本,减小与消费者之间的信息不对称性(Bakos, 1997)。经由先试点再推广,一、二线和平原城市比三、四线和山区城市更早使用该第三方数字商贸平台,普及度更高。由于同类竞争对手尚未数字化转型,欠发达地区企业使用平台的渠道整合、数字技术等资源,可能获得更多数字化红利和边际收益。据此,本文提出假设:

假设 2(a) 相对头部城市,后线城市服务业中小微企业参与第三方流程数字化商贸平台的效益更大。

假设 2(b) 相对平原地区,山区城市服务业中小微企业参与第三方流程数字化商贸平台的效益更大。

由于对信息、数字技术的接入和使用程度不同,老年人在互联网、新媒体和大数据应用分析能力上存在劣势,即“代际数字鸿沟”(Büchi et al., 2016; Bourgeois, 2007)。国务院 2020 年印发《关于切实解决老年人运用智能技术困难的实施方案》,要求贯彻落实解决老年人“数字鸿沟”问题。本文所研究的第三方流程数字化商贸平台,通过聚合多方资源和数据来优化“接入”企业的营销和管理流程,无需企业主熟练掌握复杂数字技术。相对年轻企业主,中老年企业主对数字化经营方式的接受度更低,因此同年龄段

① 数字鸿沟分为数字接入鸿沟、数字能力鸿沟和数字产出鸿沟(Warschauer, 2003; Norris, 2001)。

的多数企业主商户未开展数字化运营；同时平台能有效帮助中老年企业主减小其广告投放、媒体推送等多方面劣势。因此，使用流程数字化商贸平台的中老年企业主商户比未数字化的同类型商户，可能获得更多数字化效益。据此，本文提出假设：

假设 2(c) 相对年轻企业主，老年企业主的服务业中小微企业参与第三方流程数字化商贸平台的效益更大。

数字鸿沟可能自我强化，数字资源禀赋越强的地区和社群数字化获益越高。不同社群已有信息化基础、经济金融市场生态和数字技术教育的配套发展，是吸引数字经济投资的前提条件；数字资源禀赋不同的社群间的“数字鸿沟”可能会进一步加剧。例如东西部不同省份的电商水平差异，进而扩大了数字金融、数字征信等方面的鸿沟（星焱，2021）。由于移动支付已全国普及，相对电商、数字征信等行业，服务业如餐饮、婚庆和旅行企业，利用第三方商贸平台提供的技术、资源和数据等开拓市场和改善经营决策，无需该区域或企业本身拥有丰富的数字基础设施和技术。在具备基本的数字经济资源禀赋后，例如移动支付、消费者养成线上消费习惯等，数字经济后线城市的服务企业使用第三方流程数字化平台，可能超越数字资源禀赋相对不足的限制，获得更高边际效益。本文提出假设：

假设 2(d) 相对数字资源禀赋更丰富的数字经济头部城市，数字经济后线城市服务业中小微企业数字化的效益更大。

三、数据和研究方法

（一）样本选择及数据来源

本文研究对象为使用某大型第三方数字平台开展流程数字化的服务业中小微企业。该平台主要为服务业中小微企业提供营销数字化、生产和管理数字化服务和功能，覆盖餐饮、家政维修、美容美护、休闲娱乐、商超宅配、送洗旅行等众多生活消费领域。本文随机选取该平台百万使用移动支付的商户^①（以下简称“支付码商”），通过倾向性得分匹配方法，考察2018年1月前已入驻该平台，且截至2019年年末仍使用该平台的支付码商（实验组），相对截至2019年年末都不曾签约该平台的支付码商（控制组）在2018—2019年的营收表现。

若直接比较两组企业的营收差异，可能存在遗漏变量和自选择偏差，即一些影响企业是否使用该平台的因素未被观测，以及过往营收和所处区域经济更发达的企业更倾向加入该平台。随机分配的自然实验是最理想的因果推断方法，但现实中获取此数据几无可能。因此本文采用平衡性良好的匹配（balanced matching），通过构造反事实因果框架，实现准自然实验的因果推论效果（Rosenbaum and Rubin, 1983）。但匹配若特征变量数目太多，将造成“维数灾难”。对此倾向性得分匹配法提出独立性假设（unconfoundedness assumption）： $Y_i(0), Y_i(1) \perp D_i | X \Rightarrow Y_i(0), Y_i(1) \perp D_i | P_i(X)$ 。本文

^① 为使交易笔数和金额尽可能代表企业全部营收，本文抽样标准为：2017年在某主流移动支付平台日均交易大于5笔、2018—2019年日均有交易且存在某月交易量大于30笔的商户（若企业的交易频率低于1笔/日，很难判断该企业是否仍正常运营）。

中, $Y_i(1)$ 和 $Y_i(0)$ 分别为实验组和控制组企业营收。 D_i 为是否使用该平台的哑变量。 X 为可能影响样本使用该平台的因素, $P_i(X)$ 为由所有协变量测算出的倾向性得分。通过可观测的诸多协变量的共变项/ $P_i(X)$ 可控制遗漏变量问题, 基于倾向性得分的平衡匹配将构造出参与“流程数字化”的实验组企业的反事实结果(许竹青等, 2013)。

在PSM的截面数据性质上, 本文进一步引入时间维度: 确保所有控制变量和异质性分析的分组变量由2018年以前数据构成, 由此研究是否加入该平台所带来的企业2018—2019年营收影响, 通过解决前瞻性问题来更好地识别因果关系。选取此时间节点基于以下几点原因: 第一, 该流程数字化商贸平台在2018年年初基本完成线下商业布局(易观国际报告显示, 2017年12月该数字平台已拥有近1500万活跃用户和百万活跃商家)。第二, 2018年我国两家主流电子支付平台用户已分别超9亿。电子支付成为消费首选方式, 意味着移动支付为该平台上服务业中小微企业微小、难以统计的经营活提供大部分数据记录。第三, 2020年年初新冠疫情使服务业业绩普遍下滑, 不具备历史可比性。

独立性假设需要控制所有可能会影响企业是否参与流程数字化的因素。企业家精神、环境变化、企业资源、动态能力四方面因素对传统企业向数字化转型存在显著影响(吉峰和牟宇鹏, 2016)。在中小微企业中, 企业主是关键人物, 他们的创新性、风险偏好、年龄、对IT技术的知识和态度等可能直接影响企业是否数字化转型(Neidleman, 1979; Thong and Yap, 1995)。此外企业规模、竞争环境、信息强度、组织准备程度和可见利益等企业和环境层面因素也会影响其数字化决策(Gatignon and Robertson, 1989; Lacovou et al., 1995)。基于此, 本文从企业主、企业和企业所在地三个维度选取影响企业选择流程数字化商贸平台的协变量。

表1为上述协变量和决策变量的描述性统计。其中省份变量(*Province*)不含港澳台, 表中共列出30个哑变量的统计比较(例如最大值0.15表示商户占比最大省拥有总样本15%的商户)。男性企业主(占比55%)较女性多, 企业主平均年龄约40岁, 农村企业主占比约27%, 约65%的企业主是本省人, 企业平均使用移动支付约一年, 10%的企业使用过数字贷款。企业2017年营业水平方差较大, 因而匹配时控制该变量很重要。超过一半的企业位于一、二线城市, 三线城市占比30%, 四线及以下城市约19%, 省份间参与该平台的商户量不均衡, 12%的中小微商户在2018年前参与该流程数字化商贸平台, 即流程数字化的总比例为12%。

表1 描述性统计

	变量名称	变量定义	均值	最小值	最大值	中位数	标准差
企业主	$Gender_i$	性别(女性=1)	0.45	0.00	1.00	0.00	0.50
	Age_i	年龄	39.88	19.00	92.00	39.00	8.96
	$Huji_i$	户籍类型(农业=1)	0.27	0.00	0.00	1.00	0.45
	$If_migrant_i$	是否本省人(是=1)	0.65	0.00	1.00	1.00	0.48

(续表)

	变量名称	变量定义	均值	最小值	最大值	中位数	标准差
企业	$Dif_validmonth_i$	首次签约码商距 2018年1月的月份	11.08	1.00	123.00	7.00	14.03
	If_drawn_i	是否使用该平台 的数字贷款	0.10	0.00	1.00	0.00	0.30
企业	$Incount_2017_i$	2017年月均交易 笔数(对数/万笔)	6.05	5.02	22.88	5.79	0.97
	$Inamount_2017_i$	2017年月均交易 金额(对数/万元)	14.09	5.98	31.05	13.90	1.54
企业所 在地	$If_citydegree_3$	是否三线城市	0.30	0.00	1.00	0.00	0.46
	$If_citydegree_4$	是否四线及以下城市	0.19	0.00	1.00	0.00	0.39
	$Province_i$	企业所在省	0.03	0.00	0.15	0.02	0.04
决策变量	$Digitalization_i$	是否参与该流程 数字化商贸平台	0.12	0.00	0.00	0.00	0.32

表2为总样本及其细分行业各主要变量的均值比较。大多数 t 检验显著,说明实验组和控制组差异大,原始样本不平衡。表2进一步印证匹配前的实验组和控制组存在自选择偏差,直接使用OLS回归会产生内生性问题。

表2 不同行业数字化和非数字化商户的均值比较

变量	整体			餐饮业			零售业		
	非数 字化 商户	数字化 商户	t 检验	非数 字化 商户	数字化 商户	t 检验	非数 字化 商户	数字化 商户	t 检验
$Gender_i$	0.46	0.38	-56.41	0.46	0.38	-39.20	0.46	0.37	-39.26
Age_i	40.00	39.02	-37.13	39.86	38.33	-47.02	40.10	40.39	6.30
$Huji_i$	0.28	0.24	-33.23	0.26	0.20	-40.50	0.29	0.31	9.55
$If_migrant_i$	0.65	0.64	-11.82	0.63	0.62	-6.74	0.67	0.68	2.68
$Dif_validmonth_i$	10.23	17.51	178.30	7.82	18.59	270.09	12.01	15.32	42.16
If_drawn_i	0.09	0.14	55.15	0.06	0.16	95.43	0.12	0.12	-0.14
$If_citydegree_3$	0.30	0.30	-4.09	0.30	0.30	-6.59	0.30	0.30	1.41
$If_citydegree_4$	0.20	0.14	-51.21	0.20	0.14	-39.55	0.19	0.13	-33.20
$Incount_2017_i$	0.07	246.61	23.45	0.06	344.81	18.56	0.07	49.17	6.25
$Inamount_2017_i$	4.81	8352.60	19.00	2.85	12006.08	14.76	6.26	1007.46	7.05
$Incount_1819_i$	1.62	6794.65	22.33	1.35	10486.38	18.27	1.62	1771.20	5.93
$Inamount_1819_i$	121.00	265061.57	17.61	68.10	375798.75	13.61	160.26	42415.36	6.43

注： $Incount_1819_i$ 和 $Inamount_1819_i$ 表示2018年和2019年两年的平均值。

(二) 倾向性得分匹配

1. Logit 回归

首先利用 Logit 回归, 结合上述协变量估计企业参与流程数字化商贸平台的倾向性:

$$\begin{aligned} \text{Logit}\{P(\text{Digitalization}_i=1|X)\} = & \beta_0 + \beta_1 \text{Gender}_i + \beta_2 \text{Age}_i + \beta_3 \text{Huji}_i + \beta_4 \text{If_migrant}_i \\ & + \beta_5 \text{Dif_validmonth}_i + \beta_6 \text{If_drawn}_i \\ & + \beta_7 \text{Incoun}_t_{2017}_i + \beta_8 \text{Inamoun}_t_{2017}_i \\ & + \beta_9 \text{Province}_i + \beta_{10} \text{citydegree}_i. \end{aligned} \quad (1)$$

因变量为 $\text{Logit}\{P(\text{Digitalization}_i=1|X)\}$, P 是企业使用平台的倾向性。计算得出实验组倾向性得分的均值和标准差分别为 0.20 和 0.16, 控制组为 0.12 和 0.10。图 1 显示实验组倾向性得分的高分段分布显著多于控制组, 低分段则相反, 这表明协变量确实很好地区分了实验组和控制组。

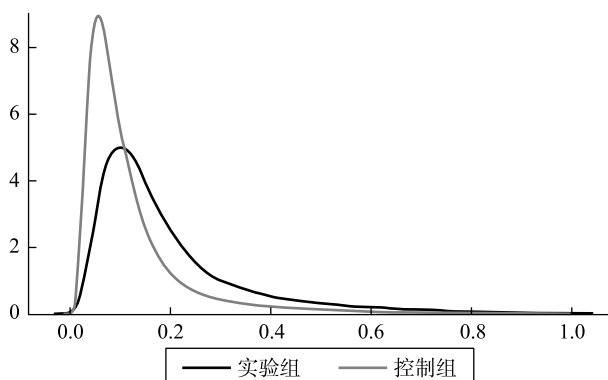


图 1 匹配前实验组和控制组的倾向性得分分布

2. 边际效应分析

根据各省的数字化倾向性得分, 江浙沪、海南省倾向性更高, 表明企业流程数字化倾向性和当地经济水平、第三产业发展程度大致正相关。表 3 为其他特征的数字化倾向性边际效应。性别、户籍类型和过往经营活跃度对数字化倾向的影响最大。商户越早成为码商, 企业主越年轻, 企业所处城市级别越高, 也更倾向于数字化。外省户籍企业主可能社会联系和支持相对较弱, 贷款则可能意味企业经营状况不佳, 因而数字化倾向性更低。

表 3 影响商户使用该流程数字化商贸平台各因素的边际效应

特征	边际效应
Gender_i	-0.05***
Age_i	-0.01***
Huji_i	-0.04***
If_migrant_i	0.01***
Dif_validmonth_i	0.03***
If_drawn_i	-0.03***
$\text{Incoun}_t_{2017}_i$	0.04***
$\text{Inamoun}_t_{2017}_i$	-0.002***

(续表)

特征	边际效应
<i>If citydegree₃</i>	-0.01***
<i>If citydegree₄</i>	-0.02***

注：*、**和***分别表示在 10%、5%和 1%水平上显著（下同）。

3. 一对一匹配

得到倾向性得分后，可使用多种匹配方法；Caliendo and Kopeinig (2008) 指出不同的匹配方法得到的结果相近。由于大数据基本可为所有实验组匹配良好的控制组，本文着重汇报放回一对一匹配结果。

首先分别对总样本、餐饮和零售业，实验组、控制组企业，按倾向性得分约 0.05—0.2 的半径分成多组。接着检验共同支撑 (common support) 条件^①。对实验组和控制组企业数目差距较大的组，直接予以剔除修剪 (trimming) 以确保最优程度的重合 (overlap)，且该方法的估计更精确 (Imbens and Wooldridge, 2009)。餐饮业倾向性得分高于 0.95、零售业高于 0.70、总样本高于 0.85 的样本均被剔除，这意味着线下零售业“十分倾向于数字化”的实验组和控制组商户差异很大，不适合做匹配。继而对修剪后各组别按无放回近邻匹配原则一对一匹配，同时对个别难以平衡的变量，进行广义精确匹配 (coarsened exact matching, CEM)。一般使用 *t* 检验比较均值差异，但它易受样本量影响，结果不稳定。本文运用 Imbens and Rubin (2015) 提出的归一化差异 (normalized difference) 方法，若

$$\frac{\bar{x}_{control} - \bar{x}_{treatment}}{\sqrt{\frac{s_{control}^2 + s_{treatment}^2}{2}}} < 0.25,$$

则每组里控制组和实验组各 *X* 变量的差异不显著。表 4 为零售业的倾向性得分和各协变量（不含省份）在组内的均值差和归一化差异，括号内为归一化差异检验值。匹配后每组实验组和控制组差异均不显著，可通过平衡性检验。^② 因此，匹配后实验与控制组只在“是否参与流程数字化商贸平台”不同，其余特征均相似。

表 4 零售业平台与非数字化商户特征的均衡检验表

	均值差与归一化差异								
	(0.00, 0.05]	(0.05, 0.10]	(0.10, 0.15]	(0.15, 0.20]	(0.20, 0.30]	(0.30, 0.40]	(0.40, 0.50]	(0.50, 0.60]	(0.60, 0.70]
P-score	-0.002 (0.15)	-0.001 (0.09)	-0.001 (0.08)	-0.001 (0.07)	-0.002 (0.07)	-0.002 (0.08)	-0.003 (0.10)	-0.000 (0.01)	0.002 (0.06)
企业主									
<i>Gender_i</i>	0.000 (0.00)	-0.000 (0.00)	-0.004 (0.01)	-0.001 (0.00)	0.002 (0.01)	-0.001 (0.00)	-0.004 (0.01)	0.008 (0.02)	-0.005 (0.03)

① 共同支撑假设是倾向得分匹配法另一重要假设： $0 < \text{Probability}(D=1 | X) < 1$ ，即拥有相同协变量特征 *X* 的样本成为实验组或控制组的概率为正。

② 除样本量缺失或不足的台湾等个别省和自治区，其他省均通过平衡性检验；总样本、餐饮业的平衡性检验也均通过。详细结果见附录 I。篇幅所限，附录未在正文列示，感兴趣的读者可在《经济学》(季刊)官网 (<https://ceq.ccer.pku.edu.cn>) 下载。

(续表)

	均值差与归一化差异								
	(0.00, 0.05]	(0.05, 0.10]	(0.10, 0.15]	(0.15, 0.20]	(0.20, 0.30]	(0.30, 0.40]	(0.40, 0.50]	(0.50, 0.60]	(0.60, 0.70]
<i>Age_i</i>	-0.006 (0.01)	-0.010 (0.01)	-0.017 (0.02)	-0.021 (0.02)	-0.018 (0.02)	-0.034 (0.03)	-0.017 (0.02)	-0.048 (0.05)	-0.064 (0.07)
<i>Huji_i</i>	0.000 (0.00)	-0.001 (0.00)	-0.000 (0.00)	-0.002 (0.00)	0.000 (0.00)	0.001 (0.00)	0.009 (0.04)	0.005 (0.02)	-0.021 (0.12)
<i>Ifmigrant_i</i>	0.000 (0.00)	0.000 (0.00)	0.002 (0.00)	0.001 (0.00)	-0.001 (0.00)	0.002 (0.00)	0.003 (0.01)	0.000 (0.00)	-0.021 (0.06)
企业									
<i>Dif_validmonth_i</i>	-0.037 (0.05)	-0.059 (0.08)	-0.091 (0.14)	-0.096 (0.15)	-0.105 (0.17)	-0.106 (0.13)	-0.153 (0.17)	-0.038 (0.04)	0.040 (0.05)
<i>If_drawn_i</i>	0.000 (0.00)	-0.001 (0.00)	-0.002 (0.01)	-0.002 (0.01)	-0.000 (0.00)	0.002 (0.01)	0.005 (0.02)	0.003 (0.01)	0.032 (0.20)
<i>Incoun_t_2017_i</i>	-0.011 (0.01)	-0.013 (0.01)	-0.013 (0.01)	-0.020 (0.02)	-0.025 (0.02)	-0.047 (0.04)	-0.054 (0.03)	-0.126 (0.07)	-0.145 (0.10)
<i>Inamoun_t_2017_i</i>	0.005 (0.01)	0.001 (0.00)	-0.003 (0.00)	-0.018 (0.02)	-0.018 (0.02)	-0.020 (0.03)	-0.034 (0.04)	-0.016 (0.01)	-0.031 (0.03)
企业所在地									
<i>Citydegree_i</i>	-0.001 (0.00)	-0.006 (0.00)	-0.007 (0.01)	-0.009 (0.01)	-0.002 (0.00)	-0.010 (0.01)	-0.007 (0.01)	-0.003 (0.01)	-0.069 (0.14)

最终得到 200 004 个有效样本, 包含 117 908 个餐饮企业和 82 076 个零售企业。由于采取一对一匹配, 匹配后实验组和控制组每组样本量相似。图 2、图 A1 和图 A2^① 为总样本、餐饮业和零售业匹配后 P-score 频数图。餐饮业倾向性得分的各分段均有数千样本分布; 零售业样本则随倾向性得分增加而减少。

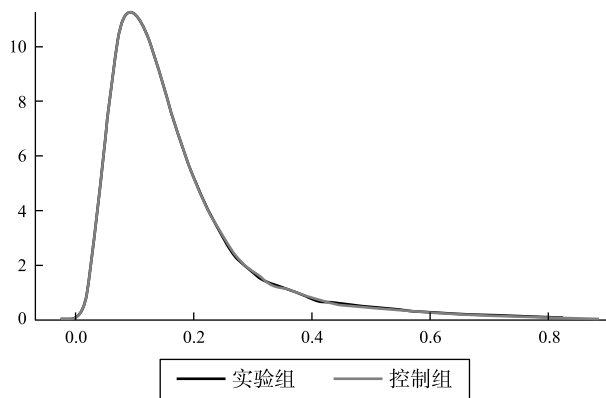


图 2 总样本匹配后的倾向性得分频数图

① 餐饮业和零售业匹配后的倾向性得分匹配图见附录 II 图 A1 和图 A2。

4. 基准回归模型

匹配后可进行倾向性得分匹配 (propensity matching)、倾向性得分调整 (propensity adjustment) 和倾向性得分加权 (propensity-based weighting) 等分析。当实验组和控制组倾向性得分完全相等时, 匹配后两组均值差与真实效应一致。但由于两组倾向性得分不可能完全相等, 直接计算两组均值差仍可能有偏。本文结合倾向性得分匹配和调整 (Rubin and Thomas, 2000; Kurth et al., 2006), 对匹配后样本进行以下回归以确保双重稳健:

$$Y_i = \alpha + \delta Digitalization_i + \beta p(X_i) + u_i, \quad (2)$$

$$Y_i = \alpha + \delta_1 Digitalization_i + \delta_2 Digitalization_i \times p(X_i) + \beta p(X_i) + u_i, \quad (3)$$

其中, Y_i 表示企业 2018—2019 年的交易金额或笔数 (取对数)^①, α 和 u_i 分别是截距项和随机扰动项。对半对数回归的系数解读, Halvorsen and Palmquist (1980) 指出其真实效应为 $e^\beta - 1$, 其中 β 为回归系数。^②

四、基准实证结果与经济解释

表 5 第 (1)、(3) 列和第 (2)、(4) 列分别为全样本基准回归模型 (2) 和模型 (3) 的结果, 第 (1)、(2) 列和第 (3)、(4) 列分别将交易笔数和金额作为因变量。相对非数字化商户, 服务业使用该流程数字化商贸平台的商户 2018—2019 年年均销量和交易额显著增加 15.53% 和 11.85%。这表明服务业企业流程数字化对营收促进作用明显, 接入和运用第三方流程数字化商贸平台是中小微企业数字化转型的有效途径之一。交互项系数显著为正, 表明数字化总体可能性越高的企业数字化收益越大, 也意味着同等条件下, 适合流程数字化商贸平台的企业若未及时线上化带来的损失可能会更大。^③ 随着经商环境和消费观念的发展, 数字运维能力将构成商户的必需能力之一。

表 5 总样本加入流程数字化商贸平台对 2018—2019 年营收的影响

	交易笔数		交易金额	
	(1)	(2)	(3)	(4)
是否加入平台	0.14*** (0.00)	0.06*** (0.01)	0.11*** (0.00)	0.03** (0.00)
倾向性得分	4.06*** (0.00)	3.81*** (0.00)	4.24*** (0.00)	4.01*** (0.00)
是否加入平台 × 倾向性得分		0.50*** (0.00)		0.47*** (0.00)
样本量	200 004	200 004	200 004	200 004

注: 括号内是汇报 p 值。系数为回归系数, 但正文系数解读为 Halvorsen and Palmquist (1980) 方法的结果。总样本倾向性得分均值为 0.1723。后表相同部分不再赘述。

① 由于交易金额是交易笔数和单笔订单金额的乘积, 单笔订单金额受其他随机因素影响, 文章主要解读交易笔数, 交易金额作为辅助参考指标。

② 若自变量是连续变量, 系数 δ 和 $\delta_1 + \delta_2 \overline{p(X_i)}$ 乘 100 表示测算出的企业由于参与流程数字化商贸平台的营收的百分比效应, 其中 $\overline{p(X_i)}$ 为样本倾向性得分均值。

③ 是否适合数字化经营, 可通过成本收益分析判别。如, 早期信息化基础设施的接入标准是否一致, 将影响该行业后期数字化发展的建设成本。

表6为餐饮和零售业分别匹配、检验和回归的结果。餐饮业数字化商户比非数字化商户的年均销量增加12.03%，年均销售额增加10.12%；零售业年均销量增加27.46%，年均销售额增加23.65%。同时，相对零售业，餐饮业商户数字化收益稍低，反映出零售业企业流程数字化的红利可观。另有两点值得注意，一是第(2)列餐饮业商户交易笔数包含交互项的回归，“是否加入平台”的系数为-0.04，考虑到餐饮业倾向性得分均值为0.2605，因此加入平台平均收益较高；但若某企业的倾向性得分很低(低于0.070)，该企业加入平台反而会使其交易笔数减少。这反映出餐饮业商户间的竞争较激烈，商户只能不断提高综合素质和竞争能力才能从流程数字化中获得营收增益。二是第(6)列零售业包含交互项的交易笔数回归，交互项系数并不显著，即零售业加入平台对交易笔数的影响在商户间的分化不大，此阶段只要“接入”和使用流程数字化商贸平台都会产生较显著的营收提升。总样本和分行业样本的基准回归结果皆印证了本文假设1。

表6 餐饮和零售业加入流程数字化商贸平台对商家2018—2019年营收的影响

	餐饮业				零售业			
	交易笔数		交易金额		交易笔数		交易金额	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
是否加入平台	0.11*** (0.00)	-0.04*** (0.00)	0.10*** (0.00)	0.01 (0.40)	0.24*** (0.00)	0.21*** (0.00)	0.24*** (0.00)	0.16*** (0.00)
倾向性得分	2.17*** (0.00)	3.97*** (0.00)	3.80*** (0.00)	1.88*** (0.00)	3.07*** (0.00)	3.06*** (0.00)	2.51*** (0.00)	2.31*** (0.00)
是否加入平台× 倾向性得分		0.60*** (0.00)		0.33*** (0.00)		0.03 (0.63)		0.39*** (0.00)
样本量	117 908	117 908	117 908	117 908	82 076	82 076	82 076	82 076

注：餐饮业和零售业的倾向性得分均值为0.2605和0.1352。

五、异质性分析和稳健性检验

为增强研究稳健性，本文补充多种匹配方法、PSM-DID、安慰剂(placebo)检验和子样本匹配回归等反驳(refute)分析。以上检验的结果与假设1的结论一致，本文核心结论依然成立。^①

(一) 异质性分析

将总体和分行业的匹配后样本，按企业所在城市级别、平原山地、企业主年龄和数字经济级别分组，再将分组子样本以实证模型(2)为基础分别回归，测算企业参与流程数字化商贸平台的效应异质性。

^① 受篇幅限制，不同匹配方法、PSM-DID和安慰剂检验的回归结果汇报于附录Ⅲ稳健性检验。

1. 不同区域的数字鸿沟演化

结合假设 2(a) 对不同城市级别行政区划的服务业企业流程数字化的经营效益设定，将商户按注册地所在城市分成：一、二线，三线和四线城市。表 7 城市经济级别异质性结果显示，使用流程数字化商贸平台的营收影响随城市级别降低而增高，且零售业数字化效应更大。总样本在一、二线，三线和四线城市数字化商户的年均交易金额分别比非数字化商户高 11.52%、10.69% 和 15.46%；年销量分别高 12.82%、17.17% 和 22.16%。餐饮业一、二线城市企业年均销量比非数字化商户高 10.41%，三线和四线城市这一比例分别为 12.33% 和 17.03%；零售业则分别为 22.61%，29.87%，43.27%。年均销售额结果类似。相对一、二线城市，三、四线城市的服务业中小微企业流程数字化获得更高收益，其下沉潜力更大，验证了假设 2(a)。

表 7 城市经济级别异质性

	交易笔数		交易金额		样本量 (5)
	是否加入平台 (1)	倾向性得分 (2)	是否加入平台 (3)	倾向性得分 (4)	
总样本					
一、二线	0.12*** (0.00)	3.50*** (0.00)	0.11*** (0.00)	3.64*** (0.00)	121 167
三线	0.16*** (0.00)	6.08*** (0.00)	0.10*** (0.00)	5.93*** (0.00)	67 589
四线及以下	0.20*** (0.00)	6.45*** (0.00)	0.14*** (0.00)	5.93*** (0.00)	32 722
餐饮业					
一、二线	0.10*** (0.00)	1.97*** (0.00)	0.10*** (0.00)	3.66*** (0.00)	63 890
三线	0.12*** (0.00)	2.38*** (0.00)	0.08*** (0.00)	4.40*** (0.00)	34 786
四线及以下	0.16*** (0.00)	2.85*** (0.00)	0.11*** (0.00)	4.68*** (0.00)	19 232
零售业					
一、二线	0.20*** (0.00)	2.71*** (0.00)	0.20*** (0.00)	2.25*** (0.00)	45 873
三线	0.26*** (0.00)	4.04*** (0.00)	0.22*** (0.00)	2.86*** (0.00)	25 360
四线及以下	0.36*** (0.00)	4.48*** (0.00)	0.27*** (0.00)	2.75*** (0.00)	10 843

结合各企业所在 338 个城市的平均海拔高度,将样本分为平原(海拔 200m 以下,总计 63 个城市)和山地(包括山地、丘陵和高原,275 个城市)两组,以检验假设 2(b),结果展示于表 8。在餐饮业和总样本中,平原和山地使用平台的企业,相对同等条件下未数字化企业,显著提升约 10% 的交易规模;山地企业交易笔数提升比平原高 2%,交易金额提升比平原低 2%—4%,这可能与山区单笔交易金额较低有关。在零售业,平原地区企业通过流程数字化获得的交易笔数和金额提升分别为 25% 和 20%,山区企业则为 20% 和 17%。因此假设 2(b)对餐饮业企业总体成立,对零售业企业不成立,因而应重视山区零售业企业数字化转型和经营发展。

表 8 平原山地异质性

	交易笔数		交易金额		样本量 (5)
	是否加入平台	倾向性得分	是否加入平台	倾向性得分	
	(1)	(2)	(3)	(4)	
总样本					
山地	0.14*** (0.00)	4.61*** (0.00)	0.10*** (0.00)	4.05*** (0.00)	107 013
平原	0.12*** (0.00)	2.85*** (0.00)	0.12*** (0.00)	2.63*** (0.00)	99 011
餐饮业					
山地	0.14*** (0.00)	2.47*** (0.00)	0.08*** (0.00)	3.50*** (0.00)	58 400
平原	0.12*** (0.00)	1.92*** (0.00)	0.12*** (0.00)	2.95*** (0.00)	62 941
零售业					
山地	0.20*** (0.00)	3.51*** (0.00)	0.17*** (0.00)	2.72*** (0.00)	48 134
平原	0.25*** (0.00)	2.17*** (0.00)	0.20*** (0.00)	1.62*** (0.00)	34 072

表 7 和表 8 的平台赋能中小微企业缩减区域数字鸿沟现象,作用机制可能源于山区和欠发达地区本身地形崎岖,人流量分散,规模化不足,服务业中小微企业更易面临“好酒也怕巷子深”的瓶颈。借助第三方平台流程数字化,能更充分享受资源整合、规模化和在线信任机制的数字红利,帮助商家超越实体店的辐射空间,实现更广触达。

2. 代际数字鸿沟的演化

结合假设 2(c),将商户按注册业主年龄分成三组:小于 30 岁、30—50 岁和 50 岁及以上。表 9 显示中高年龄企业主的企业流程数字化经营收益更高。以总样本使用平台的交易笔数影响为例,30 岁以下为 6.84%,其余两组则分别提升 16.86% 和 20.24%;餐饮和线下零售业的交易金额效益分别为:—1.13%、12.45% 和 16.61%; 15.83%、23.85% 和 31.28%。

表9 企业主年龄异质性

	交易笔数		交易金额		样本量
	是否加入平台	倾向性得分	是否加入平台	倾向性得分	
	(1)	(2)	(3)	(4)	
总样本					
18—29岁	0.07*** (0.00)	4.23*** (0.00)	-0.02 (0.13)	4.09*** (0.00)	36 002
30—49岁	0.16*** (0.00)	3.96*** (0.00)	0.13*** (0.00)	4.24*** (0.00)	164 852
50岁及以上	0.18*** (0.00)	4.88*** (0.00)	0.21*** (0.00)	4.89*** (0.00)	20 624
餐饮业					
18—29岁	0.07*** (0.00)	2.47*** (0.00)	-0.01 (0.42)	4.02*** (0.00)	22 207
30—49岁	0.12*** (0.00)	2.07*** (0.00)	0.12*** (0.00)	3.91*** (0.00)	87 203
50岁及以上	0.16*** (0.00)	3.04*** (0.00)	0.15*** (0.00)	5.08*** (0.00)	8 498
零售业					
18—29岁	0.19*** (0.00)	3.17*** (0.00)	0.15*** (0.00)	2.30*** (0.00)	10 769
30—49岁	0.29*** (0.00)	3.07*** (0.00)	0.21*** (0.00)	2.62*** (0.00)	60 624
50岁及以上	0.28*** (0.00)	3.16*** (0.00)	0.27*** (0.00)	2.19*** (0.00)	10 683

表9反映出老年人企业主利用平台参与服务业数字化，可帮助克服其数字经营能力短板，提高店铺营收。由于第三方平台聚合多方资源、数据和技术来产生价值，无需企业主掌握复杂数字技术，数字运营普及度较低的中老年企业主，现阶段开展流程数字化能获得更高边际收益，证实了假设2(c)。未来可进一步面向数字能力相对弱势的老年企业主，降低其使用数字技术和资源的门槛，通过第三方“打包”供给数字技术和服务等方式，帮助跨越个体数字能力的限制，鼓励老年人再就业，并促进全体公民共享数字红利。

3. 不同数字经济禀赋社群的数字鸿沟演化

结合假设2(d)，根据《中国城市数字经济指数蓝皮书(2021)》对全国242个城市的数字经济指数排名^①，分为数字经济一线、新一线和二线城市(简称“数字经济一、二线城市”)，数字经济三线 and 四线城市(“数字经济三、四线城市”)和未进入该数字化分级评估城市(“无数字化等级城市”)。表10展示这三组分类的基准回归式(2)结果。

① 中国信息通信研究院和新华三集团依据数据及信息化基础设施、城市治理和产业融合等综合因素排名。

表10 不同数字经济禀赋的社群异质性

	交易笔数		交易金额		样本量
	是否加入平台	倾向性得分	是否加入平台	倾向性得分	
	(1)	(2)	(3)	(4)	
总样本					
数字一、二线	0.16*** (0.00)	2.59*** (0.00)	0.06*** (0.00)	2.52*** (0.00)	72 202
数字三、四线	0.18*** (0.00)	5.68*** (0.00)	0.13*** (0.00)	4.32*** (0.00)	19 269
无等级	0.10*** (0.00)	5.24*** (0.00)	0.13*** (0.00)	4.22*** (0.00)	112 075
餐饮业					
数字一、二线	0.17*** (0.00)	1.86*** (0.00)	0.08*** (0.00)	2.80*** (0.00)	44 013
数字三、四线	0.18*** (0.00)	2.75*** (0.00)	0.12*** (0.00)	3.56*** (0.00)	12 237
无等级	0.08*** (0.00)	2.33*** (0.00)	0.10*** (0.00)	3.44*** (0.00)	63 280
零售业					
数字一、二线	0.24*** (0.00)	2.42*** (0.00)	0.17*** (0.00)	1.88*** (0.00)	28 169
数字三、四线	0.34*** (0.53)	3.80*** (0.00)	0.30*** (0.00)	2.79*** (0.00)	7 538
无等级	0.19*** (0.00)	3.17*** (0.00)	0.19*** (0.00)	2.45*** (0.00)	45 602

表10数字经济禀赋的社群异质性显示,子样本数字化效益系数在6%—35%之间,零售业效应相对更大。未进入数字经济城市排名的城市企业使用该平台也能促进交易笔数和金额提升,侧面印证假设1和假设2(a)。从交易笔数增益来看,有一定数字经济资源禀赋、但发展程度相对较低的三、四线数字经济城市,比数字经济一、二线和无等级城市的企业收益更大。这与假设2(d)的线性关系不完全一致,虽然流程数字化能给企业带来正收益,同时相对数字经济发达地区,数字经济三、四线城市增益更大,这表明当具备一定的数字经济资源禀赋(如基础数据及信息化基础设施和产业融合条件等)时,企业流程数字化可减弱数字鸿沟的自我强化效应;但数字经济三、四线城市的企业数字化效益显著高于数字禀赋资源最低地区,启示政府和大型企业平台需提供基础的信息化设施、数字产业融合条件,以促进区域经济发展。

(二) 异质性组间差异显著性检验

以上异质性在子样本组内显著,为进一步检验组间差异的显著性,本文补充Chow检验(连玉君和廖俊平,2017):

$$Y_i = \alpha + \delta_1 Digitalization_i + \beta p(X_i) + \delta_k if_group_k + \gamma_k Digitalization_i \times if_group_k + u_i, \quad (4)$$

其中 if_group_k 为不同分组的哑变量。若交乘项系数 γ_k 显著，则表明该异质性组间差异显著。回归发现，总样本、餐饮业和零售业的Chow检验结果的交互项的系数大小变化和前文异质性回归趋势一致，且均具有统计显著性，进一步印证了前文异质性分析结果的稳健性。^①

六、结论与政策启示

本文首次利用百万大数据、PSM、PSM-DID等方法，研究发现服务业使用流程数字化商贸平台的中小微商户年均销量和交易额总体年均增幅分别为15.53%和11.85%，且数字化增益与区域异质性、数字化发展程度和个体异质性有关。总体上，当具备基础数字资源条件时，经济发展和数字技术相对落后区域和在传统竞争渠道不占优的企业，通过第三方数字平台进行流程数字化的获益更大；处于一、二线城市，数字经济头部城市，以及企业主为年轻人的商家收益趋向稳定，但依然可观；但山区零售业企业流程数字化的收益能力仍有待提升。本文具有以下政策启示：第一，继续推进企业数字化进程，利用数字化技术赋能实体经济。第二，注重数字经济在缓减区域经济发展不平衡和数字鸿沟等问题的角色和功能，开展动态化、差异化数字化战略，一方面对相对成熟数字化行业不断作细分业态和场景渗透，另一方面挖掘具有数字化深耕潜力但仍处早期或盲区的地域、行业和企业。

本文受采集的大数据变量限制，未能从实证角度为服务业中小微企业利用第三方平台开展流程数字化获益的作用机制提供证据，例如进一步辨析成本节约路径、效率提升路径等。同时未来研究还可从数据上拓展除数字商贸平台以外的数字生活、金融等平台，并综合其他大数据支付平台数据，全面测算消费和产业数字化效益，为产业数字化、智慧城市建设提供参考。

参考文献

- [1] Adomako, S., J. Amankwah-Amoah, S. Y. Tarba, and Z. Khan, "Perceived Corruption, Business Process Digitization, and SMEs' Degree of Internationalization in Sub-Saharan Africa", *Journal of Business Research*, 2021, 123, 196-207.
- [2] Bakos, J. Y., "Reducing Buyer Search Costs: Implications for Electronic Market-Places", *Management Science*, 1997, 43 (12), 1676-1692.
- [3] BarNir, A., J. M. Gallagher, and P. Auger, "Business Process Digitization, Strategy, and the Impact of Firm Age and Size: The Case of the Magazine Publishing Industry", *Journal of Business Venturing*, 2003, 18 (6), 789-814.
- [4] Bharadwaj, A. S., "A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation", *MIS Quarterly*, 2000, 24 (1), 169-196.

① 受篇幅限制，此处回归结果汇报于附录IV异质性组间差异显著性检验。

- [5] Bharadwaj, A., O. A. El Sawy, P. A. Pavlou, and N. V. Venkatraman, "Digital Business Strategy: Toward a Next Generation of Insights", *MIS Quarterly*, 2013, 37 (2), 471-482.
- [6] Bourgeois, E. W., *The Race-Related Digital Divide: A Comparison between Youth in Nova Scotia and Nunavut*. Ottawa: Library and Archives Canada, 2007.
- [7] Büchi, M., N. Just, and M. Latzer, "Modeling the Second-Level Digital Divide: A Five-Country Study of Social Differences in Internet Use", *New Media & Society*, 2016, 18 (11), 2703-2722.
- [8] Caliendo, M., and S. Kopeinig, "Some Practical Guidance for the Implementation of Propensity Score Matching", *Journal of Economic Surveys*, 2008, 22 (1), 31-72.
- [9] 陈剑、黄朔、刘运辉, "从赋能到使能——数字化环境下的企业运营管理", 《管理世界》, 2020年第2期, 第117—128页。
- [10] Cohen, M. C., R. Lobel, and G. Perakis, "Dynamic Pricing through Data Sampling", *Production and Operations Management*, 2018, 27 (6), 1074-1088.
- [11] De Reuver, M., C. Sørensen, and R. C. Basole, "The Digital Platform: A Research Agenda", *Journal of Information Technology*, 2018, 33 (2), 124-135.
- [12] 傅颖、徐琪、林嵩, "在位企业流程数字化对创新绩效的影响——组织惰性的调节作用", 《研究与发展管理》, 2021年第1期, 第78—89页。
- [13] Gatignon, H., and T. S. Robertson, "Technology Diffusion: An Empirical Test of Competitive Effects", *Journal of Marketing*, 1989, 53 (1), 35-49.
- [14] Gummesson, E., and C. Mele, "Marketing as Value Co-creation through Network Interaction and Resource Integration", *Journal of Business Marketing Management*, 2010, 4 (4), 181-198.
- [15] Halvorsen, R., and R. Palmquist, "The Interpretation of Dummy Variables in Semilogarithmic Equations", *American Economic Review*, 1980, 70 (3), 474-475.
- [16] Hoffman, D. L., T. P. Novak, and A. E. Schlosser, "The Evolution of the Digital Divide: How Gaps in Internet Access May Impact Electronic Commerce", *Journal of Computer-Mediated Communication*, 2000, 5 (3), 1-55.
- [17] Imbens, G. W., and D. B. Rubin, *Causal Inference for Statistics, Social, and Biomedical Sciences*. New York: Cambridge University Press, 2015.
- [18] Imbens, G. W., and J. M. Wooldridge, "Recent Developments in the Econometrics of Program Evaluation", *Journal of Economic Literature*, 2009, 47 (1), 5-86.
- [19] 吉峰、牟宇鹏, "基于扎根理论的传统企业互联网化转型影响因素研究"《湖南社会科学》, 2016年第6期, 第141—146页。
- [20] Karimi, J., and Z. Walter, "The Role of Dynamic Capabilities in Responding to Digital Disruption: A Factor-Based Study of the Newspaper Industry", *Journal of Management Information Systems*, 2015, 32 (1), 39-81.
- [21] Kleinjans, K. J., and A. V. Soest, "Rounding, Focal Point Answers and Nonresponse to Subjective Probability Questions", *Journal of Applied Econometrics*, 2014, 29 (4), 567-585.
- [22] Kurth, T., A. M. Walker, R. J. Glynn, K. A. Chan, J. M. Gaziano, K. Berger, and J. M. Robins, "Results of Multivariable Logistic Regression, Propensity Matching, Propensity Adjustment, and Propensity-Based Weighting under Conditions of Nonuniform Effect", *American Journal of Epidemiology*, 2006, 163 (3), 262-270.
- [23] Lacovou, C. L., I. Benbasat, and A. S. Dexter, "Electronic Data Interchange and Small Organizations: Adoption and Impact of Technology", *MIS Quarterly*, 1995, 19 (4), 465-486.
- [24] Lee, D., K. Hosanagar, and H. S. Nair, "Advertising Content and Consumer Engagement on Social Media: Evidence from Facebook", *Management Science*, 2018, 64 (11), 5105-5131.
- [25] Li, L., F. Su, W. Zhang, and J. Y. Mao, "Digital Transformation by SME Entrepreneurs: A Capability Perspective", *Information Systems Journal*, 2018, 28 (6), 1129-1157.
- [26] 连玉君、廖俊平, "如何检验分组回归的组间系数差异?", 《郑州航空工业管理学院学报》, 2017年第6期, 第97—109页。
- [27] Loveman, G. W., "An Assessment of the Productivity Impact of Information Technologies", In: Allen, T. J. and

- M. Scott-Morton (eds.), *Information Technology and the Corporation of the 1990s: Research Studies*. New York: Oxford University Press, 1994, 84-113.
- [28] 罗裕梅、凌鸿, “我国网络信息消费中信息鸿沟的数字化解读”, 《社会科学》, 2014年第1期, 第53—63页。
- [29] Neidleman, L. D., “Computer Usage by Small and Medium Sized European Firms: An Empirical Study”, *Information and Management*, 1979, 2 (2), 67-77.
- [30] Norris, P., *Digital Divide: Civic Engagement, Information Poverty and the Internet World-Wide*. Cambridge: Cambridge University Press, 2001.
- [31] Oh, L. B., H. H. Teo, and V. Sambamurthy, “The Effects of Retail Channel Integration through the Use of Information Technologies on Firm Performance”, *Journal of Operations Management*, 2012, 30 (5), 368-381.
- [32] Rosenbaum, P. R., and D. B. Rubin, “The Central Role of the Propensity Score in Observational Studies for Causal Effects”, *Biometrika*, 1983, 70 (1), 41-55.
- [33] Rubin, D. B., and N. Thomas, “Combining Propensity Score Matching with Additional Adjustments for Prognostic Covariates”, *Journal of the American Statistical Association*, 2000, 95 (450), 573-585.
- [34] Sorescu, A. “Data-Driven Business Model Innovation”, *Journal of Product Innovation Management*, 2017, 34 (5), 691-696.
- [35] Thong, J. Y. L., and C. S. Yap, “CEO Characteristics, Organizational Characteristics and Information Technology Adoption in Small Businesses”, *Omega International Journal of Management Science*, 1995, 23 (4), 429-442.
- [36] Warschauers, M., *Technology and Social Inclusion: Rethinking the Digital Divide*. London: The MIT Press, 2003.
- [37] 星焱, “农村数字普惠金融的‘红利’与‘鸿沟’”, 《经济学家》, 2021年第2期, 第102—111页。
- [38] 许竹青、郑风田、陈洁, “‘数字鸿沟’还是‘信息红利’? 信息的有效供给与农民的销售价格——一个微观角度的实证研究”, 《经济学》(季刊), 2013年第4期, 第1513—1536页。
- [39] 杨旭、李竣, “县域电商公共服务资源投入与治理体系”, 《改革》, 2017年第5期, 第95—105页。
- [40] 周滔、杨庆媛、刘筱非, “西南丘陵山区农村居民点整理: 难点与对策”, 《中国土地科学》, 2003年第5期, 第45—49页。

Can the Third-Party Digital Platforms Help Small, Medium, and Micro Enterprises Improve Their Sales Performance? —Evidence from Millions of Merchants’ Big Data

XIONG Qiaoqin

(Shenzhen University)

TANG Ke* ZHANG Fengyu

(Tsinghua University)

Abstract: Using the big data of one million firms on a popular business-process-digitizing commercial platform, this paper presents a causal and heterogeneity analysis of firm digitization benefits in the tertiary industry by propensity score matching. It shows the adoption of this platform has a significant positive

* Corresponding Author: Tang Ke, Institute of Economics, School of Social Science, Tsinghua University, Mingzhai building, Tsinghua University, Haidian District, Beijing 100084, China; Tel: 86-13466777332; E-mail: ketang@mail.tsinghua.edu.cn.

effect on firms' average annual sales and trading volume compared with similar but non-digitized Small, Medium, and Micro Enterprises (SMMEs). Moreover, SMMEs in third- and fourth-tier cities, with middle-aged and elderly owners and less developed digital cities with the basic digital infrastructure, benefit more. It provides policy implications on firms' business process digitization and narrowing the digital divide.

Keywords: business process digital platform; SSMEs; propensity score matching

JEL Classification: L86, M21, O14