

数字赋能:大数据发展的治税效应*

范源源¹ 李建军²

(1.东北财经大学财政税务学院,辽宁 大连 116025;

2.西南财经大学财政税务学院,四川 成都 611130)



内容提要:“以数治税”是税收现代化服务中国式现代化的客观要求,是驱动税收执法、服务、监管方式和手段变革的重要基础。本文基于2010—2022年中国上市公司数据,以“国家大数据综合试验区”建设为准自然实验,在数字经济发展背景下考察以大数据为基础的区位导向性政策对公司税收遵从的影响。研究发现:“国家大数据综合试验区”建设显著提高了公司税收遵从水平。机制分析表明,大数据发展通过推动税收征管数字化升级,抑制公司隐匿数字化产品销售收入和异常关联交易,进而提升公司税收遵从水平。异质性分析表明,该效应主要集中于内外部信息环境较差的公司和规模较大的公司。拓展性分析结果表明,大数据发展对公司税收遵从影响的长期效应大于短期效应。从治税收益的角度看,大数据发展显著增加了政府税收收入,为持续推动数字经济发展提供了动力。经济后果检验发现,大数据发展还会增加公司通过避税来化解经营风险的成本。本文结论不仅有助于深化理解数据治理创新在推动税收治理和国家治理现代化中的作用效果,也为进一步推动数字产业政策与新一轮财税体制改革协同发力、形成合力提供了新的证据。

关键词:新质生产力 数字政府 大数据 以数治税 公司避税

中图分类号:F812.42 **文献标志码:**A **文章编号:**1002—5766(2024)12—0038—22

一、引言

自20世纪90年代以来,将数字技术运用于公共治理一直是世界各国政府的主要政策目标(Wilson和Mergel,2022)^[1]。近年来,信息技术与经济社会的交汇融合引发了数据迅猛增长,数据已成为国家基础性战略资源,大数据正日益对全球生产、流通、分配、消费活动以及经济运行机制、社会生活方式和国家治理能力产生重要影响^①。为了加快实施国家大数据战略,我国启动了八大国家大数据综合试验区建设,在推进地区大数据整合、共享以及创新应用等方面取得了显著成效,促进了我国数字经济的发展。据统计,目前我国数字经济规模超过50万亿元,总量稳居世界第二,占GDP比重提升至41.5%^②。数字经济的迅猛发展不仅为实体经济高质量发展提供了新动能,也为打造数字政府、实现政府宏观调控目标提供了新机遇。

收稿日期:2024-03-03

*基金项目:国家社会科学基金重大项目“数字经济时代我国税收制度适配性研究”(23&ZD064)。

作者简介:范源源,男,副教授,研究领域为税收理论与政策,电子邮箱:dufeyy@sina.com;李建军,男,教授,博士生导师,研究领域为财税理论与政策,电子邮箱:lijj@swufe.edu.cn。通讯作者:李建军。

① 国务院关于印发促进大数据发展行动纲要的通知(国发[2015]50号)[EB/OL].https://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm,2015-09-05。

② 数据来源:央视网,《中国城市数字经济发展报告(2023)》发布[EB/OL].<https://sh.cctv.cn/2024/01/05/ARTIAE1bmtEkUUAXaLrd78Mu240105.shtml>,2024-01-05。

已有研究表明,数字经济在推动政府数字治理上发挥了积极作用。陈艳利和蒋琪(2024)^[2]研究发现,数字经济背景下政府数据开放有助于推动企业创新。何雨可等(2024)^[3]研究发现,大数据赋能政府治理能够降低制度性交易成本,进而激发城市创业活力。此外,也有研究探讨了数字经济发展对地方政府腐败(Welch等,2005)^[4]、地方政府治理能力(聂爱云和潘孝虎,2023)^[5]、财政透明度(刘奥和张双龙,2022)^[6]以及污染治理(余典范等,2023)^[7]等的影响。然而,数字经济发展对于财税治理的影响尚不明确。2023年12月中央经济工作会议强调,要增强宏观政策取向一致性,加强财政、货币、就业、产业、区域、科技、环保等政策协调配合。可见,探究数字经济背景下大数据产业政策如何与新一轮财税体制改革形成合力具有重要的现实意义。

实践中,许多国家的税务部门都在逐步运用大数据来推动税收征管效率提升。如美国国内收入局(IRS)运用大数据技术将纳税人财务数据与其他数据集进行整合分析,构建纳税人行为预测模型,提高了税收稽查的选案效率(Neuman和Sheu,2022)^[8]。澳大利亚税务局运用大数据技术制定不同层级的风险监控指标,通过可验证的第三方数据开展全面的数据比对稽核,识别出异常的涉税交易事项^①。比利时基于数据挖掘技术设立新系统“交易网络分析”,依托该平台,成员国可以快速交换、协同处理增值税数据,打击增值税逃税^②。近年来,我国也在积极运用大数据提升税收治理现代化水平。其中,“国家大数据综合试验区”(以下简称“试验区”)政策的实施为推动大数据和税收征管融合发展提供了新的契机。一方面,试验区通过持续推进数据中心整合利用,加快经济活动数据化,帮助税务部门及时、准确、完整地收集纳税人及其经济社会活动的各类信息,为实施精准税收监管提供数据基础;另一方面,试验区积极推动大数据技术的加速创新和广泛使用,这也有利于税收征管的数字化升级和智能化改造(李鑫钊,2023)^[9]。然而,目前少有文献探究试验区建设对税收征管效率和公司税收遵从的影响。

基于此,本文以中国上市公司为研究样本,以试验区建设为准自然实验,构建双重差分模型,考察大数据发展的治税效应。研究发现,大数据发展显著提高了公司的税收遵从水平,这一结论在经过一系列稳健性检验后依旧成立。机制分析结果表明,大数据发展通过推动税收征管的数字化升级,抑制了公司隐匿数字化产品的销售收入和异常关联交易,进而提升公司税收遵从水平。基于信息不对称理论和税收分类监管的异质性分析结果表明,该效应对于内外部信息环境较差和规模较大的公司而言更为显著。拓展性分析结果表明,随着大数据和税收征管的深度融合,大数据发展对公司税收遵从水平的提升效应也越强。此外,大数据发展还显著增加了政府的企业所得税收入;经济后果检验发现,大数据发展也会增加公司通过避税化解经营风险的成本。

本文边际贡献在于:第一,本文补充了试验区建设经济后果的文献。以往研究试验区建设经济后果的文献,主要集中在碳排放(Wei等,2023)^[10]、企业数字化转型(孙伟增等,2023)^[11]以及劳动投资效率(游家兴和苏三妹,2024)^[12]等方面,较少有文献关注试验区建设对公司税收遵从的影响。本文从大数据发展推动税收征管数字化水平提高这一视角切入,验证了试验区建设对公司税收遵从的影响,拓展了现有文献的研究框架。第二,本文拓展了公司税收遵从影响因素的相关研究。以往解释公司税收遵从影响因素的文献,多是基于信息透明度(Pomeranz,2015)^[13]、融资约束(魏志华和夏太彪,2020)^[14]和代理问题(严若森等,2018)^[15]。本文结合数字经济背景下涉税事项具有交易隐秘性、高度流动性以及数据依赖性等特点,从异常关联交易和隐匿收入的角度,探究了大数据发展对公司税收决策的具体影响途径,为未来研究公司税收遵从问题提供了新的思路。第三,本文还为推动作为新质生产力重要组成部分的大数据技术和税收征管深度融合,

① 澳大利亚税务局副局长杰里米·赫希霍恩. 澳大利亚:精心优化客户体验 推进征管数字化转型[N]. 中国税务报, 2022-1-5.

② 樊勇,陈飞. 科技税务推动制度变革与领域应用[M]. 北京: 中国财政经济出版社, 2024.

助力税收治理现代化和国家治理现代化提供了政策参考。现有税收征管数字化的相关文献大多关注征管系统数字化的影响(Dečman等,2010^[16];Bhuasiri等,2016^[17];Night和Bananuka,2019^[18];张克中等,2020^[19])。也有部分文献关注电子发票使用(Bellon等,2022)^[20]、采用电子纳税申报(Okunogbe和Pouliquen,2022)^[21]以及区块链技术运用(Anomah等,2024)^[22]等征管方式数字化的影响。然而,对于大数据发展和税收征管数字化的关系还缺乏系统的实证分析。本文通过实证研究发现大数据发展可以通过促进经济社会数字化来推动税收征管数字化,这为进一步建设适应数字经济发展的税收征管体系提供了新的证据。

二、政策背景与理论分析

1. 制度背景

2015年8月,国务院印发的《促进大数据发展行动纲要》(以下简称《纲要》)中明确提到,大数据已成为推动经济转型发展的新动力、重塑国家竞争优势的新机遇,也是提升政府治理能力的途径。为了贯彻落实《纲要》,2015年9月,贵州省启动全国首个大数据综合试验区建设工作。2016年2月,国家发展和改革委员会等三部委发函批复,同意贵州省建设国家大数据(贵州)综合试验区。试验区围绕大数据制度创新、数据中心整合利用以及数据共享开放等开展系统性试验。其中,事关政府治理能力提升的系列措施包括^①:第一,持续推进数据中心整合利用试验。贵州省实施数据资源汇聚工程,打造全国大数据内容中心,这为税收征管数字化水平的提升打下了坚实的数据基础。第二,持续推进数据共享开放试验。贵州省通过推动省内政务信息系统互联互通以及数据共享交换,实现了政务数据的统筹管理。政府信息交换平台的建设正有利于深化税收大数据共享应用,推动税收征管的数字化升级和智能化改造。第三,持续推进大数据创新应用试验。贵州省促进数据开放共享和对跨部门、跨领域数据进行分析,推动税务与海关、金融、公安、市场监督管理等进行跨部门和跨地区的信息交换、数据共享、执法联动和协同监管。

2016年10月,国家发展和改革委员会等三部委批复同意在京津冀等七个区域建设第二批试验区,包括京津冀、珠江三角洲、上海市、河南省、重庆市、沈阳市以及内蒙古。其中,京津冀和珠江三角洲为跨区域类综合试验区,更加注重数据要素流动,以数据流引领资金流、人才流,支撑跨区域公共服务、社会治理和产业转移。上海市、河南省、重庆市、沈阳市为区域示范类综合试验区,更加注重数据资源统筹,加强大数据产业集聚,促进区域协同发展。内蒙古为基础设施统筹发展类综合试验区,依托地理优势加大资源整合、建设大数据中心^②。虽然八大试验区总体定位有所区别,但是,从目标来看,各试验区都强调了形成大数据产业集聚,将大数据运用于政府科学决策和公共治理中。据此,本文以试验区建设为准自然实验,探究大数据发展的治税效应。

2. 理论分析和研究假说

数字技术的不断发展、经济数据的海量生成以及数据存储处理能力的攀升,都为推动税收征管的数字化升级、提高税收遵从水平创造了新的机遇。事实上,我国税务机关一直在探索运用现代化征管技术来推动税收征管的数字化升级和智能化改造。早在20世纪80年代初,我国就开始使用微型计算机来辅助税收征管工作,而后历经金税一期、金税二期、金税三期建设,到目前进入到金税四期建设新阶段。金税工程(税收管理信息系统工程)以计算机网络为依托,通过对已有信息的整合分析来压缩企业逃税空间(张克中等,2020)^[19]。然而,构建以数据为核心驱动要素的新

^① 曾帅,陈玲,江婷婷. 不负使命担重托 乘风破浪正当时——贵州深入推进国家大数据(贵州)综合试验区建设[N]. 贵州日报,2021-6-12。

^② 连玉明等. 中国大数据发展报告(No.1)[M]. 北京:社会科学文献出版社,2017。

型税收征管模式,不仅需要统一规范的应用系统平台,还需要推进大数据等新一代信息技术与税收征管融合发展。目前全球数字经济发展迅速,数字经济发展下的市场主体具有的高流动性、交易隐蔽性、业务模糊性和数据依赖性特征,会从多个方面显著增加税收征管难度,降低现行税收征管制度的效率(李鑫钊,2023)^[9]。虽然数字经济发展给税收征管带来了多种挑战,但以大数据为代表的数字化技术与税收征管的有机结合,不仅可以提升税收征管效率,也能有效提高纳税主体的税收遵从水平。大数据发展影响公司税收遵从的实质是通过经济社会数字化为税收征管数字化升级创造条件,进而抑制公司通过隐匿收入和关联交易等方式进行避税。

第一,大数据发展可以推动税源数字化,抑制公司通过隐匿收入等方式进行避税。理论上,信息不对称问题是现代税收征管的核心问题(Pomeranz,2015)^[13]。以往的税收征管模式中,税务机关使用的税费信息主要依赖于纳税人自行申报的数据以及部分通过数据共享平台获取的经济行为类佐证数据,数据种类和体量都显不足(冯绍伍等,2022)^[23]。尤其是在数字经济背景下,计算机软件、音像制品以及信息服务等数字化产品通过互联网以电子形式销售,交易标的具有无形性、隐秘性和开放性特征,导致物流和资金流难以追踪(姚东旭和严亚雯,2023)^[24]。大数据发展推动了税源信息的多样性和丰富性,为加强对数字经济活动的税收监管提供了契机。一方面,大数据发展可以推动经济社会的数据化,促使企业愿意将交易数据全面记录,企业财务信息更加公开透明(Goldfarb和Tucker,2019)^[25]。随着企业交易数据可得性的增加,税源信息变得更加丰富,这为税收征管的数字化升级提供了数据基础。另一方面,大数据发展还推动政府各部门系统互联互通,进而有助于税务机关从多领域、多方面深度挖掘涉税信息。在数字经济时代,海量交易信息涌现,但这些涉税数据标准和口径不统一,限制了税务机关对税收大数据的挖潜利用(高博楠和王志刚,2023)^[26]。此外,信息交换激励机制不完善、信息统计口径不统一以及缺乏信息交换意识等都会制约跨部门涉税信息交换效率,影响信息交换的税收遵从效应的发挥(陈思霞等,2021)^[27]。试验区积极建设统一的数据共享交换平台和开放平台,并编制大数据关键共性标准,因此,当地税务机关可以依托该平台获取的信息进行大数据对比稽核,查实公司是否存在不实申报等避税行为。

第二,大数据发展还可以推动税收征管业务流程的数字化,抑制公司通过关联交易等方式进行避税。已有文献表明,关联交易是企业控制收入和成本,进行盈余操纵和避税的重要手段(洪剑峭和薛皓,2009)^[28]。数字经济的迅猛发展降低了物流运输成本,为企业通过与低税率子公司进行关联交易的方式来逃避税款提供了便利(张乾等,2022)^[29]。税收征管的数字化升级可以提高税务机关识别企业异常关联交易的效率,提升关联交易双方的税务风险(刘慧龙等,2022)^[30]。借助大数据,税务机关能够及时、全面、准确地对经济活动进行智能化监管,推动税收征管模式由传统“以票管税”向“以数治税”分类精准监管转变(李平,2020)^[31]。试验区积极深化大数据共享应用,将大数据运用于完善税收监管的模型和指标体系,推动了税收征管业务流程的数字化,进而可以提高税务机关查实公司是否存在异常关联交易等避税行为的效率。

据此,本文提出研究假说:

H₁:大数据发展可以提升公司税收遵从水平。

三、研究设计

1. 样本与数据来源

本文选取2010—2022年中国A股所有上市公司进行分析,并按照如下原则对初始样本进行筛选:剔除金融保险类上市公司;剔除ST和ST*的公司;剔除大连市、青岛市、深圳市、宁波市和厦门市等计划单列市的样本;剔除主要财务数据缺失的样本。经过上述处理后,本文共得到29327个观测值。由于各变量缺失程度不同,因此各回归中样本个数略有不同。此外,本文还对所有连续变量

进行了上下各1%的缩尾处理。本文数据来自于CSMAR数据库、中国研究数据服务平台(CNRDS)和Wind数据库,省级层面的数据收集整理于《中国统计年鉴》和各省份统计年鉴。

2. 模型设计与变量定义

本文以试验区的设立为准自然实验,运用双重差分法来考察大数据发展对公司税收遵从的影响,原因在于:试验区是国家大数据战略的重要内容,是数字经济迅猛发展背景下建设数字中国的必然要求,因此,其实施更多源于政府的行为和当地的社会经济因素,单个公司的特征很难决定该地区是否成为试验区,因此该试验区的设立对于公司来说,具有一定的外生性,利用该试点政策探究大数据发展与公司税收遵从的因果关系,能够避免研究中反向因果等内生性问题。为此,本文构建的基准模型如式(1)所示:

$$BTD_{sit} = \delta_0 + \delta_1 Dig_s \times Post_t + \beta X_{sit} + \mu_i + \gamma_t + \varepsilon_{sit} \quad (1)$$

其中, BTD_{sit} 用于度量公司税收遵从水平,下标 s 表示省份, i 表示公司, t 表示年份。本文借鉴Desai和Dharmapala(2006)^[32]的研究,用会计税收差异(BTD)作为公司税收遵从的衡量指标,其中 $BTD = (\text{会计利润} - \text{当期所得税费用} \div \text{名义税率}) \div \text{期初总资产}$,该指标越大,表示公司避税的可能性越大,公司税收遵从水平越低。

本文以试验区设立作为大数据发展水平的外生冲击,设定虚拟变量 Dig_s ,当 s 省份属于试验区试点,赋值为1,否则设为0。 $Post_t$ 为试验区政策实施前后的虚拟变量,2016年之前设为0,2016年及之后设为1。

X_{sit} 代表控制变量,借鉴以往研究(李青原等,2022^[33];Nguyen和Nguyen,2020^[34]),包括公司层面特征和宏观层面特征。具体来说,公司层面特征包括:(1)公司规模($SIZE$),公司规模会影响税收遵从,但影响方向并不明确。部分学者认为,公司规模越大,其税收遵从水平越高。这是因为对于大企业来说,其作为“纳税大户”通常会受到税务部门的重点管理(于文超等,2018)^[35]。也有研究认为,大公司有更大动机从事避税活动,因为大型企业财务更复杂,更有能力构建复杂的避税交易(Zimmerman,1983)^[36]。(2)现金持有量($CASH$),一方面,公司避税存在成本和风险,因此,在公司现金流充裕的情况下,公司无需承担避税风险,避税动机自然会减弱。另一方面,也有研究认为避税程度较高的公司会倾向持有更多的现金,以应对未来可能来自税务机关的处罚(Hanlon等,2017)^[37]。(3)固定资产比率(FR),公司固定资产的折旧方法和折旧年限会影响企业所得税税基,公司可以通过弹性选择不同的固定资产折旧方法和年限来进行税收筹划。(4)无形资产比率(IR),由于研发费用可以在计算企业所得税时税前扣除,因此无形资产比率(IR)也会影响公司避税水平(陈作华和方红星,2018)^[38]。(5)账面市值比(MB),以往研究认为,公司会出于风险权衡进行税收筹划,账面市值比越高的企业,避税动机相应减弱(丁方飞等,2019)^[39]。(6)公司年龄(AGE),成立年限越长的公司,税收策略越趋于保守,避税动机相对较小(陈德球等,2016)^[40]。不同成立年限的公司,享受的税收优惠有所差异,公司经营状况、税收负担和现金流状况也不同,公司税收决策可能也存在差异。除了公司自身的微观特征,公司所在地区的环境因素也会影响其税收遵从。为此,本文控制了产业结构(INS)、人均GDP的对数($LNGDP$)和财政收支占比($STRESS$),以控制地区宏观经济因素对公司税收遵从的影响。此外,本文还控制了公司固定效应 μ_i 和年份固定效应 γ_t 。 δ_0 和 ε_{sit} 分别是常数项和残差项。相关变量定义及描述性统计如表1所示。

表1 主要变量定义及描述性统计

变量	变量定义	均值	标准差	最小值	最大值
BTD	如上文所示	0.0414	0.0826	-0.2467	0.348
$SIZE$	公司总资产的自然对数	22.2399	1.3046	18.9433	26.1791

续表 1

变量	变量定义	均值	标准差	最小值	最大值
CASH	现金及现金等价物/总资产	0.1586	0.1222	0.007	0.6445
FR	固定资产/总资产	0.2139	0.1585	0.0018	0.6908
IR	无形资产/总资产	0.0468	0.0508	0	0.3234
MB	总资产/市值	0.628	0.2517	0.1098	1.184
AGE	所在年度—企业注册年份+1	19.696	5.7435	7	34
INS	第一产业增加值/第二产业增加值	0.1554	0.1392	0.0085	0.9153
LNGDP	人均地区生产总值取对数	10.5401	0.7551	7.9336	11.7685
STRESS	省级财政支出/省级财政收入	1.7849	0.7212	1.0795	4.7521

四、实证结果

1. 基准回归结果

为了考察大数据发展对公司税收遵从的影响,本文对基准模型进行回归,回归结果如表 2 所示。其中,列(1)为单变量回归,大数据发展变量($Dig \times Post$)的系数为-0.0118,在 1% 水平显著;列(2)控制公司层面特征变量后,大数据发展变量($Dig \times Post$)的系数为-0.0082,同样在 1% 水平显著;列(3)进一步控制宏观层面的特征后,此时大数据发展变量($Dig \times Post$)的系数为-0.0074,在 1% 水平显著。意味着,就平均值而言,试验区建设促使试点地区公司的税收遵从水平大约提高 17.87% ($0.0074 \div 0.0414 = 17.87\%$)。以上结果表明,大数据发展能够发挥治税效应,提高公司税收遵从水平。假设 H_1 得到验证。

表 2 大数据发展与公司税收遵从

变量	(1)	(2)	(3)
$Dig \times Post$	-0.0118*** (0.0028)	-0.0082*** (0.0026)	-0.0074*** (0.0027)
SIZE		0.0203*** (0.0019)	0.0202*** (0.0019)
CASH		0.1506*** (0.0077)	0.1507*** (0.0077)
FR		-0.0814*** (0.0096)	-0.0815*** (0.0096)
IR		-0.0570* (0.0291)	-0.0565* (0.0292)
MB		-0.0744*** (0.0046)	-0.0743*** (0.0046)
AGE		0.0014 (0.0017)	0.0014 (0.0017)
INS			-0.0307 (0.0253)
LNGDP			0.0108 (0.0096)
STRESS			0.0073 (0.0047)
常数项	0.0645*** (0.0019)	-0.3550*** (0.0464)	-0.4693*** (0.1067)

续表 2

变量	(1)	(2)	(3)
年份固定效应	是	是	是
公司固定效应	是	是	是
观测值	29327	28784	28778
R ²	0.0268	0.0945	0.0948

注: 括号内为公司层面的聚类稳健标准误; **、*、* 分别代表系数在 1%、5% 和 10% 水平显著, 下同

2. 稳健性检验

(1) 动态效应分析。表 2 仅列示了大数据发展对公司税收遵从的平均效应, 而动态效应检验能够清晰地展示试验区建设的动态影响。为此, 本文构建模型(2):

$$BTD_{sit} = \alpha_0 + \sum_{s=-5}^6 \alpha_s \times Dig_s \times year_{it} + \beta X_{sit} + \mu_i + \gamma_t + \varepsilon_{sit} \quad (2)$$

其中, $year_{it}$ 是时间虚拟变量, 在试点建设前(后)第 t 年时为 1, 否则为 0。以试点建设前一年 ($t = -1$) 为基准期。图 1 直观地列示了大数据发展对公司税收遵从的动态影响, 从中可以看到, 在试验区建设以前, 系数均不显著, 表明试点地区与非试点地区公司的税收遵从水平不存在显著差异, 说明不能拒绝平行趋势假设成立。在试验区建设以后, 系数均显著为负, 表明试验区建设可以提高公司税收遵从水平。

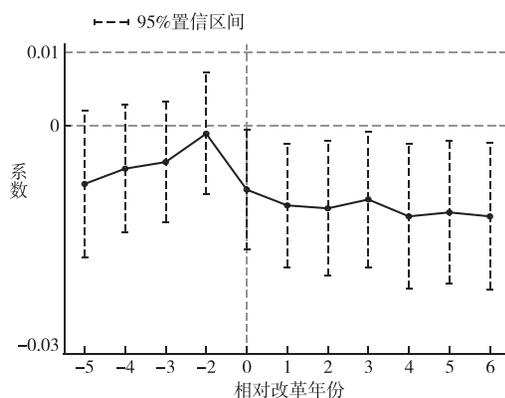


图 1 动态效应分析

现有研究表明, 传统的处理前趋势检验可能造成估计偏误, 进而影响平行趋势假设成立的结论 (Roth 等, 2023)^[41]。为此, 本文借鉴许文立和孙磊 (2023)^[42] 的研究, 采用相对偏离程度限制和平滑限制方法对平行趋势进行敏感性检验。图 2 分别列示了采用上述两种方法进行平行趋势敏感性检验的结果。可以看到, 处理后的点估计量的置信区间均不包含 0 值, 这意味着, 在相对偏离程度限制下, 政策实施的税收遵从水平提升效应稳健; 采用平滑限制方法下处理前趋势偏差 20% 时, 政策实施效应依旧稳健。

(2) 异质性处理效应检验。当存在异质性处理效应时, 即使不能拒绝平行趋势假设成立, 处理效应的估计结果也会存在偏误, 且这一偏误在政策统一时间实施的场景下依旧存在 (许文立和孙磊, 2023)^[42]。为此, 本文采用多期多个体倍分模型进行检验。具体来说, 将位于试验区的公司视作处理组, 位于非试点地区的公司作为控制组, 将处理组公司试点实施后的税收遵从度与其反事实结果进行比较, 从而得到处理效应, 在加权平均后估计出政策转换效应。图 3 列示了采用该模型的动态效应检验结果, 可以看到, 在试点实施以前, 政策效应并不显著, 而在试点启动后, 政策效应显著为负。以上结果表明, 本文基准回归结论稳健。

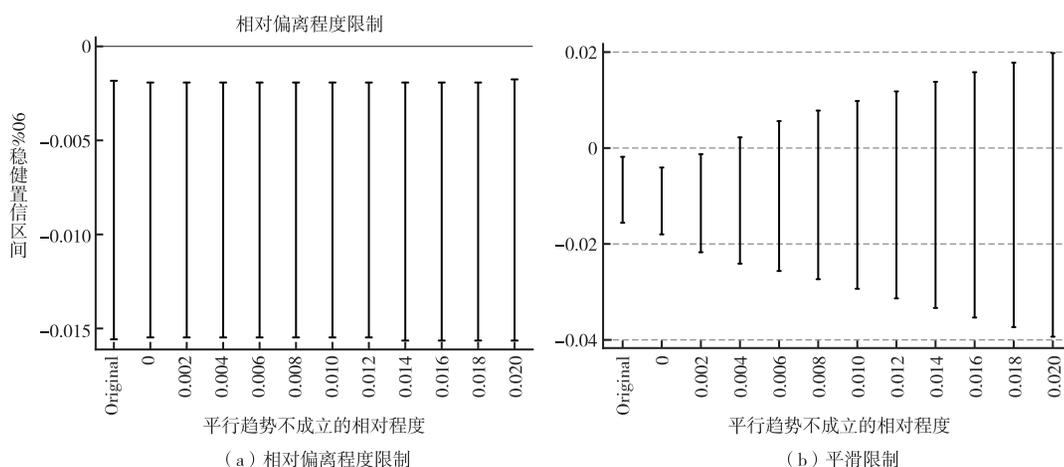


图2 平行趋势假设的敏感性检验

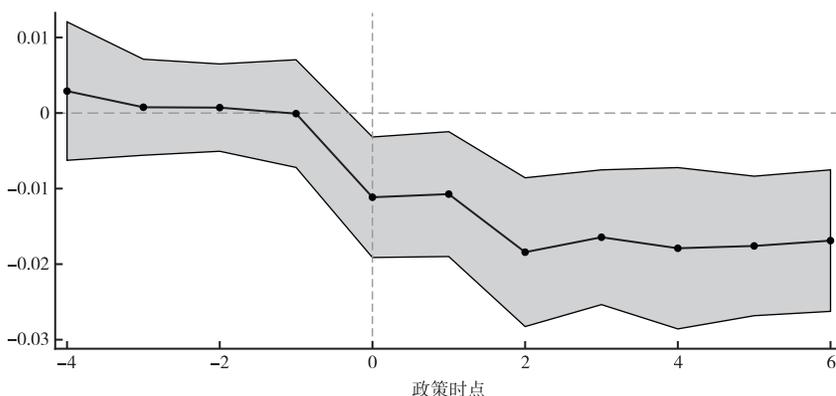


图3 异质性处理效应检验结果

(3)工具变量法。为了避免存在与试验区空间选址类似的其他区位导向性政策对估计结果的影响,本文采用工具变量法来解决试验区选址的内生性问题。借鉴Lyu等(2024)^[43]的研究做法,采用历史上各城市1984年每百人固定电话数量与前一年的全国网民数量(百万)相乘得到的交互项(iv)作为工具变量。主要原因在于,试验区建设需要良好的数字基础设施,通信网络则是数字经济时代的关键基础设施。因此,历史上的电信基础设施可能会从技术和行为习惯上影响试验区的选址,符合相关性要求。此外,理论分析表明,公司税收决策是管理层对避税的收益和风险进行权衡后的结果(Allingham和Sandmo,1972)^[44]。因此,历史上的电信基础设施不会影响当前公司税收遵从水平,满足外生性要求。进一步,本文在“恰度识别”(工具变量的个数和内生性变量的个数相等)的情况下,无法从统计上来严格论证工具变量的外生性假设。为此,本文借鉴Conley等(2012)^[45]的研究,放松工具变量严格外生性假设进行估计,构建如下两阶段方程:

$$BTD_{sit} = \theta_0 + \theta_1 Dig_s \times Post_t + \theta_2 iv_{st} + \beta X_{sit} + \mu_i + \gamma_t + \varepsilon_{sit} \quad (3)$$

$$Dig_s \times Post_t = \tau_0 + \tau_1 iv_{st} + \beta X_{sit} + \mu_i + \gamma_t + \varepsilon_{sit} \quad (4)$$

其中, iv 代表工具变量(历史上各城市1984年每百人固定电话数量与前一年网民数量交互项),其余变量定义与上文一致。若 $\theta_2 \neq 0$,表明 iv 并不严格外生;若 $\theta_2 = 0$,表明 iv 近似外生。本文使用置信区间集合方法(UCI),在符合经济逻辑的前提下,预先给 θ_2 设置一个预期范围,再计算得到 θ_1 的点估计与置信区间。

表3列(1)和列(2)列示了使用工具变量的回归结果。列(1)列示了第一阶段的估计结果,可见,工具变量(iv)的系数在1%水平显著为正,即历史上的电信基础设施和当地成为试验区的概率

正相关,进一步结合 K-Prk Wald F 统计量检验结果可知,本文选取的工具变量不存在弱工具变量问题。列(2)列示了第二阶段的估计结果,可见,此时大数据发展变量($Dig \times Post$)的系数仍然显著为负,进一步证明了本文研究结论的稳健性。此外,基于 UCI 方法得到的大数据发展变量($Dig \times Post$)的置信区间为(-453232.87, -836.00783),这表明,即使在工具变量并不完全外生的前提下,大数据发展依然显著提高了公司税收遵从水平。

表 3 放松“严格外生性”要求的工具变量法估计结果

变量	(1)	(2)	(3)	
	第一阶段	第二阶段	UCI	
$Dig \times Post$		-0.0319*** (0.0089)	-453232.87	-836.00783
iv	0.0004*** (0.0000)			
K-Prk Wald F	331.105			
年份固定效应	是	是		
公司固定效应	是	是		
控制变量	控制	控制		
观测值	24312	23926		
R ²	0.5324	0.0895		

此外,现实中 θ_2 的分布和取值会发生变化,而非始终沿着本文所设定的某一区间变化。为此,本文还绘制了大数据发展变量($Dig \times Post$)的回归系数随着 θ_2 变动而变动的分布区间图。如图 4 所示,尽管随着 θ_2 的增加,即工具变量外生性逐渐减弱的情况下, θ_1 的系数出现了递减的趋势,但仍然显著为负。可知,即使 θ_2 的值持续变化,大数据发展变量($Dig \times Post$)对公司税收遵从的影响也始终存在。综上,即使本文所选工具并非完全外生,本文的结论依旧稳健。

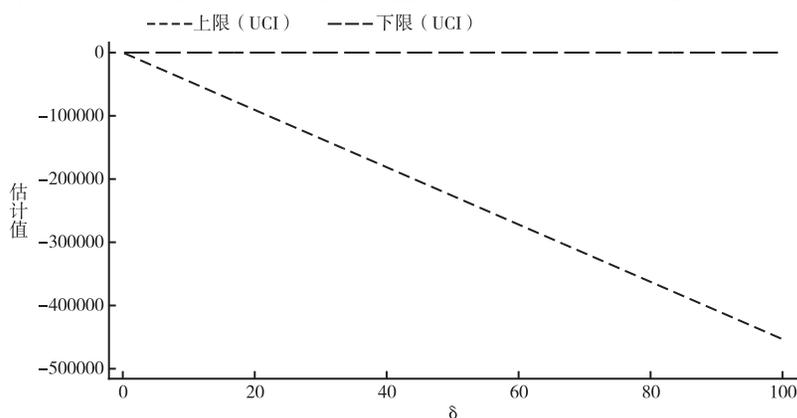


图 4 UCI 假设下的置信区间

(4)考虑样本选择偏误。本文以试验区建设为准自然试验,运用双重差分法考察了大数据发展对公司税收遵从的影响。然而,由于试验区政策并非完全随机,本文的结果可能是由位于试验区的公司自身特征所导致。为了避免样本选择偏误引起的内生性问题,本文以倾向得分匹配-双重差分法和熵平衡-双重差分法展开进一步检验。

首先,采用倾向得分匹配(PSM)方法为位于试验区的公司寻找配对样本。PSM的基本思路是,在未处于试验区的控制组中找到某个公司 j ,使得 j 与位于试验区的处理组公司 i 的可观测变量尽可能相似(匹配),即 $X_j \approx X_i$,以便能够进行相互比较。由于截面匹配可能会混淆数据的时间特征,本文借鉴 Blundell 和 Costa Dias (2005)^[46]研究,在双重差分之前采用逐年匹配方法为各年处理组找到

对照组,以避免试点的非随机性问题。具体步骤如下:本文选取了公司规模(*SIZE*)、现金持有(*CASH*)、固定资产比率(*FR*)、无形资产比率(*IR*)、账面市值比(*MB*)以及公司年龄(*AGE*)等作为匹配变量。然后,逐年按照半径为0.1的最近邻一对一匹配方法为处理组找到对照组,进行共同支撑假设检验和平衡性假设检验。为了考察位于试验区的公司和未处于试验区的公司在匹配前后是否存在系统性差异,本文分别绘制了匹配前后两者的倾向得分核密度分布。如图5所示,可以看到,匹配前两者的分离趋势明显,而在匹配后均值线的距离缩短,两条曲线更加接近,说明匹配在一定程度上有效,满足共同支撑假设。同时,本文还比较匹配前后不同年份Logit回归结果,发现匹配后各协变量的系数值有所减小、显著性降低以及伪R²减小,满足平衡性假设^①。

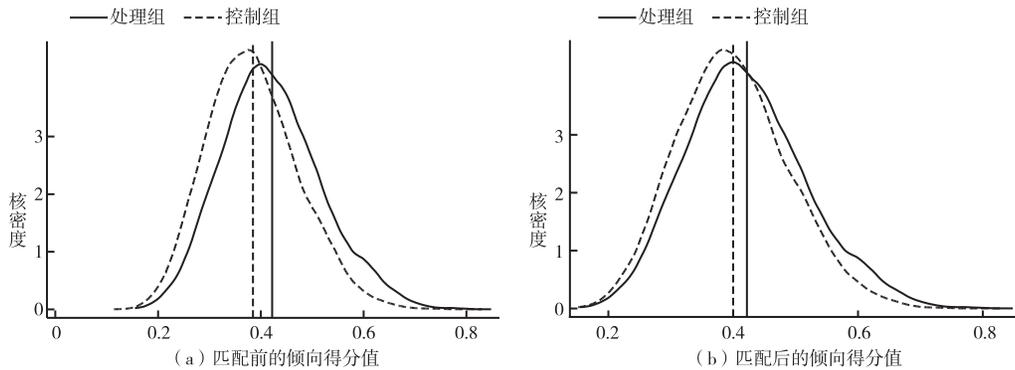


图5 匹配前后倾向得分核密度分布结果

利用匹配后样本进行双重差分检验。估计结果如表4列(1)所示,可以看出,大数据发展变量(*Dig × Post*)的回归系数依旧显著为负。为了提高匹配结果的稳健性,本文还按照半径为0.1的最近邻一对二匹配方法进行匹配。表4列(2)的结果显示,更换匹配方法后本文结论依旧稳健。

表4 考虑样本选择偏误的检验结果

变量	(1)	(2)	(3)	(4)
	近邻一对一	近邻一对二	三阶矩	二阶矩
<i>Dig × Post</i>	-0.0060** (0.0029)	-0.0061** (0.0029)	-0.0065** (0.0028)	-0.0065** (0.0027)
年份固定效应	是	是	是	是
公司固定效应	是	是	是	是
控制变量	控制	控制	控制	控制
观测值	21823	22774	28324	28324
R ²	0.0949	0.0990	0.4380	0.4400

其次,本文借鉴Hainmueller(2012)^[47]的方法,采用熵平衡法来减少估计时的选择偏误。与PSM不同,熵平衡法匹配处理组和对照组能够保证各协变量高阶矩分布的相似性,并且不会损失样本。选取与PSM同样的变量来进行匹配。表5列示了采用熵平衡法(三阶矩)进行匹配的效果,可以看到,在匹配前各协变量的均值、方差和偏度都有较大差异,在熵平衡匹配后各协变量的均值、方差和偏度的差异都明显缩小,表明匹配在一定程度上是有效的。然后,用熵平衡匹配后的样本进行双重差分检验,回归结果如表4列(3)和列(4)所示,可以看到,大数据发展变量(*Dig × Post*)的系数显著为负,表明进一步考虑样本选择偏误后,本文结论仍然成立。

① 平衡性假设检验:由于逐年PSM是一年一年进行匹配的,因此考察各协变量在两组间是否存在系统性偏差只能在同一年份中进行比较,不同年份的匹配样本没有可比性。限于篇幅,平衡性检验结果未列示,备索。

表 5 熵平衡法检验效果(三阶矩)

匹配前	处理组			控制组		
	均值	方差	偏度	均值	方差	偏度
SIZE	22.4000	2.0240	0.7280	22.1500	1.4530	0.5681
CASH	0.1708	0.0171	1.3970	0.1508	0.0133	1.5490
FR	0.1903	0.0248	1.0960	0.2293	0.0247	0.8255
IR	0.0449	0.0026	2.9150	0.0479	0.0026	3.1020
MB	0.6316	0.0666	0.0812	0.6254	0.0612	0.0072
AGE	19.8900	35.0700	0.1405	19.6000	31.7900	0.1265
匹配后	处理组			控制组		
	均值	方差	偏度	均值	方差	偏度
SIZE	22.4000	2.0240	0.7280	22.4000	2.0220	0.7288
CASH	0.1708	0.0171	1.3970	0.1706	0.0171	1.3990
FR	0.1903	0.0248	1.0960	0.1907	0.0248	1.0920
IR	0.0449	0.0026	2.9150	0.0450	0.0026	2.9170
MB	0.6316	0.0666	0.0812	0.6316	0.0666	0.0802
AGE	19.8900	35.0700	0.1405	19.8900	35.0400	0.1403

(5)安慰剂检验。为了排除试验区实施时间和试点地区的选择性,本文通过随机筛选试点地区并随机产生试点实施时间,然后使用构造的虚假实验估计基准模型。图 6 为随机处理 500 次后虚拟处理效应的概率分布。可以看到,随机模拟得到的回归系数分布在 0 附近,这进一步表明本文估计结果是稳健的。

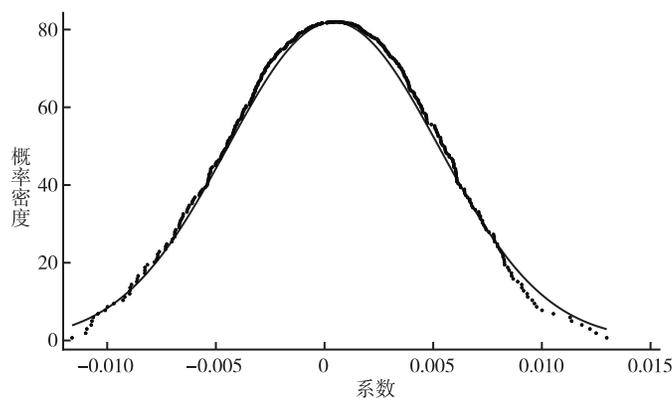


图 6 安慰剂检验

(6)考虑税收相关政策影响。第一,考虑“金税三期”工程建设的影响。“金税三期”工程以“以票控税”为指导思想,充分掌握和利用税收大数据,显著提高了税务机关的稽查能力,有效提升了公司税收遵从度(张克中等,2020)^[19]。为了排除该政策的影响,本文在基准模型中加入“金税三期”政策实施虚拟变量($Gtax$)。第二,考虑 2018 年国地税合并的影响。范子英等(2022)^[48]研究发现,国地税合并后,征管体系的去属地化有效压缩了企业逃税空间。为了排除该政策的干扰,本文构建了“国地税合并”政策虚拟变量($Com \times Time$)加入到基准模型中,其中,当公司的企业所得税原本是地税局征收, Com 取 1,否则为 0;2018 年及之后的样本, $Time$ 取 1,否则为 0。第三,考虑优化税收营商环境的影响。以往研究发现,国家税务总局于 2017 年起开展“税收营商环境试点”,显著提高了企业税法遵从度(范源源和李建军,2023)^[49]。因此,为了排除优化税收营商环境的影响,本文在基准模型中加入省市优化税收营商环境试点变量($Ytax$)。表 6 列(1)~列(3)列示了考虑税收

相关政策影响后的回归结果,可以看到,大数据发展变量($Dig \times Post$)的系数依旧显著为负。

(7)考虑其他数字经济相关政策的影响。第一,“宽带中国”战略实施的影响。本文将“宽带中国”实施试点变量($Inter$)纳入到基准模型中,以控制城市网络基础设施建设对当地税收征管效率的影响。第二,考虑信息惠民政策实施的影响。贺晓宇和储德银(2023)^[50]研究发现,信息惠民政策可以推动当地政府治理数字化转型。因此,本文将信息惠民试点城市变量($Infor$)加入到基准模型中,以控制政府治理数字化转型对当地税收征管效率的影响。表6列(4)和列(5)列示了考虑其他数字经济相关政策影响后的回归结果,可以看到,大数据发展变量($Dig \times Post$)的系数依旧显著为负。以上结果表明,在剔除其他政策影响后,本文结论依旧成立。

表6 排除同期政策的干扰

变量	(1)	(2)	(3)	(4)	(5)
	金税三期	国地税合并	优化税收营商环境	宽带中国	信息惠民
$Dig \times Post$	-0.0076*** (0.0027)	-0.0071*** (0.0027)	-0.0064** (0.0027)	-0.0073*** (0.0027)	-0.0067** (0.0027)
$Gtax$	-0.0085*** (0.0025)				
$Com \times Time$		0.0065** (0.0026)			
$Ytax$			-0.0105*** (0.0023)		
$Inter$				-0.0017 (0.0027)	
$Infor$					-0.0033 (0.0027)
年份固定效应	是	是	是	是	是
公司固定效应	是	是	是	是	是
控制变量	控制	控制	控制	控制	控制
观测值	28778	28778	28778	28330	28778
R ²	0.0953	0.0952	0.0962	0.0947	0.0949

(8)替换核心变量的衡量指标。为了避免单一度量指标给估计结果造成的偏误,借鉴以往研究(李青原等,2022^[33];范源源和李建军,2023^[49]),本文还采用扣除应计盈余影响的会计-税收差异($DDBTD$)和消除名义税率影响的所得税率(TA_ETR)来重新度量公司税收遵从水平。其中 $TA_ETR = 企业名义所得税率 - 实际所得税率$ 。此外,为了提高政策识别的准确性,尽管所有的试验区都是在2016年获批,但贵州省在2015年9月就已经开展相应建设,同时,其他试验区都是在2016年下半年才获批,因此,将贵州省的政策实施节点设定为2015年,其他则设定为2016年,据此设定新的多期双重差分变量 Dig_N 。表7列(1)~列(3)列示了采用新的衡量指标的回归结果,可以看到,更换解释变量和被解释变量的度量指标后,本文的结论仍旧成立。

(9)变换样本。首先,为了排除新设企业和退出企业的影响,本文采用平衡面板数据重新对基准模型进行回归;其次,为了排除时间跨度过长的数据可能带来的序列相关问题以及新冠疫情等的影响,本文仅保留2013—2018年的样本重新进行检验;最后,剔除贵州省的样本后进行估计,即只考虑第二批试验区建设的税收遵从效应。回归结果如表7列(4)~列(6)所示,可以看到,大数据发展变量($Dig \times Post$)的系数依旧显著为负,表明本文结论稳健。

表 7 替换核心变量和变换样本后的估计结果

变量	(1)	(2)	(3)	(4)	(5)	(6)
	<i>DDBTD</i>	<i>TA_ETR</i>	<i>BTD</i>	平衡面板	缩短样本	剔除贵州
<i>Dig</i> × <i>Post</i>	-0.0078*** (0.0029)	-0.0202*** (0.0068)		-0.0087*** (0.0030)	-0.0097*** (0.0029)	-0.0074*** (0.0027)
<i>Dig</i> _{<i>N</i>}			-0.0077*** (0.0027)			
年份固定效应	是	是	是	是	是	是
公司固定效应	是	是	是	是	是	是
控制变量	控制	控制	控制	控制	控制	控制
观测值	28778	28778	28778	16307	12564	28525
R ²	0.0841	0.0225	0.0949	0.0912	0.0799	0.0944

五、机制分析

1. 大数据发展与税收征管数字化升级

前文分析认为,大数据发展可以推动当地税收征管的数字化升级,进而提升公司的税收遵从水平。税收征管的数字化升级最直接的后果就是税务机关的征管水平提高。因此,本文首先检验大数据发展对当地税收征管水平的影响。为此,构建如下模型:

$$TE_{it} = \vartheta_0 + \vartheta_1 Dig_{it} \times Post_t + \beta \cdot X_{it} + \mu_s + \gamma_t + \varepsilon_{it} \quad (5)$$

其中,被解释变量 *TE* 为当地税收征管水平,借鉴江轩宇(2013)^[51]的做法,以地区实际税负和预测税负衡量的税收努力程度(*TE*)作为税收征管水平的代理变量^①,该比值越大,说明该地区税收征管水平越高;选取地区生产总值(*GDP*)的自然对数、财政支出占 *GDP* 的比例、进出口总额占 *GDP* 的比例、城镇人口占总人口的比例以及第一产业和第二产业的比值等作为控制变量,其余变量与上文定义相同^②。基于模型(5)的回归结果如表 8 列(1)所示,可以看到,大数据发展变量(*Dig* × *Post*) 在 1% 水平显著为正,说明大数据发展的确提高了当地税收征管水平。

表 8 机制分析结果

变量	(1)	(2)	(3)	(4)	(5)	(6)
	税收征管	征管数字化	<i>YN</i>	数字收入隐匿	<i>PRT</i> ₁	<i>PRT</i> ₂
<i>Dig</i> × <i>Post</i>	0.0655** (0.0269)	-0.0125*** (0.0032)	0.0029** (0.0015)	-0.0053* (0.0027)	-0.0097** (0.0044)	-0.0112** (0.0049)
<i>Dig</i> × <i>Post</i> × <i>Gtax</i> ₂		0.0108*** (0.0041)				
<i>Dig</i> × <i>Post</i> × <i>Inf</i>				-0.0252*** (0.0076)		
<i>Gtax</i> ₂		0.0228 (0.0149)				
<i>Inf</i>				-0.0166 (0.0117)		
年份固定效应	是	是	是	是	是	是

① 通过各地区第一产业产值占 *GDP* 的比重、第二产业产值占 *GDP* 的比重、进出口总额占 *GDP* 的比重等回归在地区税收收入占 *GDP* 的比重上,其拟合值为地区预测税负。

② 由于未获取到 2022 年部分宏观层面数据,因此这里只采用 2010—2021 年数据。

续表 8

变量	(1)	(2)	(3)	(4)	(5)	(6)
	税收征管	征管数字化	YN	数字收入隐匿	PRT_1	PRT_2
公司固定效应	是	是	是	是	是	是
控制变量	控制	控制	控制	控制	控制	控制
观测值	354	28778	21479	28778	20953	20953
R ²	0.6726	0.0957	0.0016	0.0963	0.0359	0.0409

注:除第(1)列,其余各列中括号内为公司层面的聚类稳健标准误

此外,如果税收征管的数字化升级是试验区设立提升公司税收遵从水平的原因,那么试验区设立的税收遵从效应应该主要存在于试点前位于税收征管数字化水平较低地区的公司,而不是那些位于税收征管数字化水平本身较高地区的公司。由于很难对地区税收征管数字化水平进行直接度量,本文以“金税三期”工程的实施作为当地税收征管数字化水平的替代衡量指标,如果试点前(2016年以前)当地实施了“金税三期”工程,认定为税收征管数字化水平较高的地区,相反,则为税收征管数字化水平较低的地区。主要理由在于,“金税三期”以智能化为特征,运用大数据技术充分整合涉税信息,推动了税收征管的数字化升级(刘慧龙等,2022)^[30]。为此,本文在基准模型中纳入试点前地区税收征管数字化变量($Gtax_2$)与大数据发展变量($Dig \times Post$)的交乘项进行回归。估计结果如表 8 列(2)所示,可以看到,大数据发展对公司税收遵从水平的提升效应在试点前税收征管数字化水平较低地区的公司中更强。这一结果也从侧面印证了试验区设立可以推动当地税收征管的数字化水平提升,进而提高公司税收遵从水平。

事实上,上述检验实际上只提供了“大数据发展—税收征管数字化水平提升—公司税收遵从”的间接证据。遗憾的是,本文无法将税收征管数字化水平匹配到试验区进行检验。即便如此,这里依然可以通过一些新闻事件来进行补充分析。例如,2021年贵州省在进一步深化税收征管改革实施方案中提出,“发挥国家大数据(贵州)综合试验区先行优势,依托贵州省数据共享交换平台和电子政务外网,有序有力推进涉税数据共享”^①。在该方案的指引下,贵州省积极建立省市县三级税收大数据共享机制,与有关部门充分沟通编制了《涉税涉费业务协同责任清单》和《政务涉税涉费数据清单》,累计梳理涉及 20 余个相关部门的业务协同事项 31 项,梳理涉及 50 个部门的政务涉税涉费数据达 226 个表单 2592 个数据项^②。因此,上述影响机制分析进一步说明了试验区建设推动了税收征管的数字化升级,进而提高了公司税收遵从水平。

2. 具体路径分析

上文的机制分析表明,大数据发展可以通过提升税收征管的数字化水平进而提高公司税收遵从水平。本部分从公司收入隐匿和关联交易的视角切入,进一步研究税收征管数字化水平提升影响公司税收决策的具体路径。

第一,大数据发展可以提升税收征管数字化水平,使得税务机关更容易查实公司有无隐匿销售收入的行为,进而提高公司税收遵从水平。理论上,在企业的货币资金会计科目总体上账实相符的前提下,收入隐匿即企业实际获得了相应收入但未体现在税负上。为此,本文借鉴孙鲲鹏和石丽娜(2022)^[52]研究,通过公司营运现金流增长和税负增长之间的匹配关系来衡量公司的收入隐匿状况。因此,本文构建如下模型:

$$YN_{sit} = \mu_0 + \mu_1 Dig_s \times Post_t + \beta X_{sit} + \mu_i + \gamma_t + \varepsilon_{sit} \quad (6)$$

① 省人民政府办公厅关于印发贵州省进一步深化税收征管改革的实施方案的通知(黔府办发[2021]18号[EB/OL]. https://www.guizhou.gov.cn/ztzl/gzszgzhgfwjksj/gfwjksj/srmzfbgt/202201/t20220128_72438903.html, 2021-07-30.

② 余昌旭.贵州省深化税收征管改革实现六个创新[N].贵州日报,2021-10-15.

其中, YN 为衡量公司收入隐匿状况的指标, 具体衡量方法为: 将公司的所得税增长率回归在营运环节现金流入增长率以及行业、年份等一系列控制特征上, 获得其残差, 并将该残差在年份和行业内部标准化为 0~1 之间的分值作为公司收入隐匿概率的反向指标。 YN 越大, 代表公司隐匿收入概率越小。此外, 由于公司预收账款计入现金流量表“销售商品、提供劳务收到的现金”, 但在当期不一定缴纳企业所得税。公司其他应付款虽不计入营运环节现金流入, 但其他应付款确实无法偿付时应计入其他收入缴纳企业所得税。为此, 本文还控制了公司预收账款的对数和其他应付款的对数, 其余控制变量与上文一致。基于模型(6)的回归结果如表 8 列(3)所示, 可以看到, 大数据发展变量 ($Dig \times Post$) 的系数显著为正。这表明, 大数据发展显著抑制公司通过收入隐匿的方式进行避税。

此外, 数字经济背景下, 数字化产品交易标的具有无形性、隐秘性和开放性等特征, 而大数据发展推动了税源信息的多样性和丰富性, 为加强对数字经济活动的税收监管提供了契机。为此, 本文还进一步考察大数据发展对公司隐匿数字化产品销售收入的影响。理论上, 如果减少公司隐匿数字化产品销售收入是大数据与税收征管融合发展提高公司税收遵从的具体路径之一的话, 那么大数据发展对公司税收遵从水平的提升效应应该对于以销售数字化产品为主的公司更为显著。为此, 本文在基准模型中引入公司是否属于信息传输、软件和信息技术服务业 (Inf) 与大数据发展变量 ($Dig \times Post$) 的交乘项进行回归。估计结果如表 8 列(4)所示, 可以看到, 大数据发展对公司税收遵从水平的提升效应对于销售数字化产品为主的公司更强, 表明大数据和税收征管融合发展提升了稽查信息化水平, 抑制了公司隐匿收入(尤其是数字化产品的销售收入), 进而提高了公司税收遵从水平。

第二, 大数据发展还可以提高税收征管数字化水平, 进而抑制公司通过关联交易的方式进行避税。为此, 本文构建如下模型:

$$PRT_{sit} = \rho_0 + \rho_1 Dig_s \times Post_t + \beta X_{sit} + \mu_i + \gamma_t + \varepsilon_{sit} \quad (7)$$

其中, PRT 为公司异常关联交易的衡量指标。借鉴宫晓云等(2022)^[53]研究, 本文采用公司关联交易总额超出行业一年度均值的部分来衡量公司异常关联交易 (PRT_1)。同时, 为了提高估计结果的稳健性, 本文还通过回归模型计算得到的残差部分, 作为公司异常关联交易的另一衡量指标 (PRT_2)^①。其余变量定义与基准模型一致。基于模型(7)的回归结果如表 8 列(5)和列(6)所示, 可以看到, 大数据发展变量 ($Dig \times Post$) 的系数显著为负。这表明, 大数据和税收征管融合发展通过减少异常关联交易进而提高了公司税收遵从水平。

六、异质性分析

1. 基于信息不对称理论的异质性分析

理论上, 信息不对称是现代税收征管的核心问题(Pomeranz, 2015)^[13]。企业信息越透明, 企业逃避税被税务机关查获的概率就越大, 其逃避税成本也越高(范源源和李建军, 2022)^[54]。因此, 当公司信息透明度较低时, 公司的税收遵从水平也相对较低。此时, 大数据发展对公司税收遵从水平的提升效应就应该越强。为此, 本文试图从盈余管理和分析师跟踪两个角度, 探究公司内部信息质量和外部信息环境对大数据发展与公司税收遵从二者关系影响的异质性。具体而言, 本文采用分行业分年度的修正 Jones 模型估计出的可操控性总应计利润的绝对值来测度公司内部信息质量。对于外部信息环境, 现有研究表明, 分析师作为重要的信息中介, 可以提高企业信息透明度(张纯和吕伟, 2009)^[55]。因此, 本文参考 Clement(1999)^[56]做法, 计算分析师预测准确度^②, 分析师

① 通过行业和公司规模、杠杆率和成长性等公司特征回归在关联交易比(上市公司与关联方的关联交易总额占资产总额的比例)上计算得到的残差部分, 定义为异常关联交易 (PRT_2)。

② 具体计算方法为: 某分析师某次对于某股票预测误差的绝对值减去其他所有分析师的误差的绝对值的平均值, 两者之差除以其他所有分析师误差的绝对值的平均数。

预测准确度越高,表明公司外部信息环境越好。

基于此,本文分别按照上述指标的年度中位数将样本进行分组回归。估计结果如表9所示,可以看到,大数据发展对公司税收遵从水平的提升效应只在内部信息质量和外部信息环境较差的样本组中显著。

表9 基于信息不对称理论的异质性分析结果

变量	(1)	(2)	(3)	(4)
	内部信息质量较好	内部信息质量较差	外部信息环境较好	外部信息环境较差
<i>Dig × Post</i>	-0.0018 (0.0026)	-0.0139*** (0.0042)	-0.0105 (0.0072)	-0.0153*** (0.0056)
年份固定效应	是	是	是	是
公司固定效应	是	是	是	是
控制变量	控制	控制	控制	控制
观测值	14139	13888	4718	4653
R ²	0.0964	0.1171	0.1622	0.0873

2. 基于税收分类监管的异质性分析

2021年中共中央办公厅和国务院办公厅印发的《关于进一步深化税收征管改革的意见》中明确提出,加强重点领域风险防控和监管,精准实施税务监管。对逃避税问题多发的行业、地区和人群加强税收风险管控与分类监管,可以有效节约税收征管资源,提升税收征管质效。首先,对于大企业来说,其作为“纳税大户”,在税收征管过程中可能有着更强的议价能力。同时,大企业可以聘请更多财务人员和税收中介,税收筹划能力也相对较强。孙鲲鹏和石丽娜(2022)^[52]的研究也表明,与规模较小的企业相比,大企业有着更多的逃税机会。此外,大企业具有集团化管理和规模化经营的特点,以往基于“点对点”税收管理模式主要针对单个成员企业的税收风险事项,难以解决大企业税收聚集性和税收管理零散性之间的矛盾。现有研究发现,上市公司集团可以通过高税率地区母公司向低税率地区子公司转移利润的方式进行避税(赵仁杰和周小昶,2022)^[57]。因此,大数据发展对公司税收遵从水平的提升效应对于规模较大的公司更为显著。

基于此,本文试图从公司资产规模和集团规模两个角度探究大数据发展对于不同税收稽查重点企业的异质性影响,其中集团规模用上市公司子公司数量来衡量。估计结果如表10所示,可以看到,大数据发展对公司税收遵从水平的提升效应只在公司资产规模较大的样本组和集团规模较大的样本组中显著。

表10 基于税收分类监管的异质性分析结果

变量	(1)	(2)	(3)	(4)
	资产规模较大	资产规模较小	集团规模较大	集团规模较小
<i>Dig × Post</i>	-0.0104*** (0.0033)	-0.0028 (0.0044)	-0.0081** (0.0036)	-0.0039 (0.0045)
年份固定效应	是	是	是	是
公司固定效应	是	是	是	是
控制变量	控制	控制	控制	控制
观测值	14465	14313	14393	13931
R ²	0.0944	0.0981	0.0961	0.0780

七、拓展性分析

1. 大数据发展对公司税收遵从的政策效应

考虑到试验区建立后,大数据与税收征管融合发展需要一定时间,而且公司可能也会根据所处环境调整其税收筹划策略。为此,本文设定模型(8)来探究大数据发展对公司税收遵从的长期和短期影响。

$$BTD_{sit} = \gamma_0 + \gamma_1 Post_t^S + \gamma_2 Post_t^L + \gamma_3 Dig_s \times Post_t^S + \gamma_4 Dig_s \times Post_t^L + \beta \cdot X_{sit} + \mu_i + \gamma_t + \varepsilon_{sit} \quad (8)$$

其中, $Post_t^S$ 表示试验区设立的短期效应,当年份为2017—2019时,赋值为1,否则为0; $Post_t^L$ 表示试验区设立的长期效应,当年份为2020—2022时,赋值为1,否则为0。在模型(8)中,系数 γ_3 和 γ_4 分别表示大数据发展对公司税收遵从的短期和长期因果影响。由表11可知, $Dig_s \times Post_t^S$ 的系数为-0.0062,在5%的水平上显著; $Dig_s \times Post_t^L$ 的系数为-0.0076,在5%的水平上显著。以上结果说明,大数据发展对公司税收遵从的长期效应大于短期效应,即试验区设立后,随着时间推移,大数据技术和税收征管深度融合,对公司税收遵从水平的提升效应也越强。

表 11 大数据发展对公司税收遵从的政策效应

变量	改革的长期效应
$Dig_s \times Post_t^S$	-0.0062** (0.0027)
$Dig_s \times Post_t^L$	-0.0076** (0.0033)
年份固定效应	是
公司固定效应	是
控制变量	控制
观测值	28778
R ²	0.0948

2. 大数据发展的治税收益

上文分析验证了大数据发展影响公司税收遵从的作用机制及异质性效应,但大数据发展发挥的治税效应带来了多大的收益还不明确。为此,本文借鉴 Chaurey(2017)^[58]的研究,粗略计算大数据发展的治税收益。其原理在于,基于试验区建设后地区企业所得税收入的情况,利用双重差分的估计系数倒推出反事实情况下的纳税额,通过二者的差额来评估大数据发展的治税收益。

具体来说,本文以地区企业所得税收入(取自然对数)对政策交互项($Dig \times Post$)回归,可以得到系数 $\delta(0.0928)$ ^①,即试验区建立后地区企业所得税收入增加了9.28%。表12列示了本文估算的大数据发展的治税收益,以2017年为例,试验区的企业所得税收入约为6151亿元,按照试验区建立后上升9.28%推算,如果没有成立试验区,这些地区的企业所得税收入为5628.66亿元,由此可以得到,2017年企业所得税收入增加值为522.34亿元。按此逻辑,可以推算出2018—2021年所对应的收益。可见,大数据发展使得地区企业所得税收入增加2842.25亿元,有效提高了地方财政收入,为进一步深入实施地区数字产业政策提供了财政基础。

① 这里基于2010—2021年各省份的面板数据,选取了地区生产总值(GDP)的自然对数、财政支出占GDP的比例、第一产业占GDP的比例、第二产业占GDP的比例以及进出口总额占GDP的比例等作为控制变量,估计大数据发展对地区企业所得税收入的影响。

表 12 大数据发展的治税收益 单位:亿元

年份	(1)企业所得税	(2)政策效应系数	(3)反事实收入	(4)收益
2017	6151	0.0928	5628.66	522.34
2018	6680	0.0928	6112.74	567.26
2019	6755	0.0928	6181.37	573.63
2020	6469	0.0928	5919.66	549.34
2021	7415	0.0928	6785.32	629.68
总收益(亿元)				2842.25

3. 经济后果分析

本文研究发现,大数据发展可以提高公司税收遵从水平,但税务合规成本的提高可能会加剧公司经营风险。理论上,公司可以通过逃税的方式来进行内源性融资,节约公司现金流,进而降低公司风险。然而,大数据发展水平提高后,税收征管数字化水平的提升会增加公司通过避税来化解经营风险的成本。为了考察这一经济后果,本文建立如下模型:

$$Risk_{sit} = \sigma_0 + \sigma_1 Dig_s \times Post_t + \beta X_{sit} + \mu_i + \gamma_t + \varepsilon_{sit} \quad (9)$$

$$Risk_{sit} = \varphi_0 + \varphi_1 Dig_s \times Post_t + \varphi_2 BTD_{sit} + \beta X_{sit} + \mu_i + \gamma_t + \varepsilon_{sit} \quad (10)$$

其中, $Risk$ 表示公司经营风险,借鉴毛志宏等(2024)^[59]的研究,本文用 Z 指数来衡量公司的经营风险, $Z = 1.2 \times (\text{营运资金} \div \text{总资产}) + 1.4 \times (\text{留存收益} \div \text{总资产}) + 3.3 \times (\text{息税前利润} \div \text{总资产}) + 0.6 \times (\text{权益的市场价值} \div \text{总负债的账面价值}) + 0.999 \times (\text{营业收入} \div \text{总资产})$ 。该指数越大,经营风险越低,其余变量定义与上文一致。表13列示了基于上述模型的估计结果。由列(1)可知,以公司经营风险($Risk$)为解释变量时,大数据发展($Dig \times Post$)的系数显著为负,这表明,大数据发展显著提高了公司经营风险。由列(2)可知,公司税收遵从水平(BTD)的系数显著为正,表明公司避税可以降低经营风险。此外,表2列(3)的结果已经表明,大数据发展会显著抑制公司避税、提高公司税收遵从。以上结果表明,虽然大数据发展显著抑制了公司避税,但在一定程度上增加了公司经营风险。此外,为了提高估计结果的稳健性,本文借鉴林炳洪和李秉祥(2024)^[60]的做法,计算了 O 指数来衡量公司经营风险^①, O 指数越大代表经营风险越高。表13列(3)和列(4)列示了以 O 指数衡量公司经营风险的估计结果,可以看到,上述结论依旧成立。

表 13 大数据发展的经济后果

变量	(1)	(2)	(3)	(4)
	Z指数	Z指数	O指数	O指数
$Dig \times Post$	-0.2608* (0.1486)	-0.2132 (0.1472)	0.1464** (0.0643)	0.0582 (0.0534)
BTD		5.6893*** (0.3823)		-11.2477*** (0.2285)
年份固定效应	是	是	是	是
公司固定效应	是	是	是	是
控制变量	控制	控制	控制	控制
观测值	28228	28228	25858	25858
R^2	0.2629	0.2742	0.1431	0.4322

① $O = -1.32 - 0.407SIZE + 6.03TLTA - 1.43WCTA + 0.0757CLCA - 2.37NITA - 1.83FUTL + 0.285INTWO - 1.72OENEG - 0.521CHIN$ 。其中, $SIZE = \ln(\text{总资产})$; $TLTA = \text{总负债} \div \text{总资产}$; $WCTA = \text{营运资金} \div \text{总资产}$; $CLCA = \text{流动负债} \div \text{流动资产}$; $NITA = \text{净利润} \div \text{总资产}$; $FUTL = \text{经营性净现金流} \div \text{总负债}$; $INTWO$,若过去两年净利润均为负数则为1,否则为0; $OENEG$,如果总负债>总资产为1,否则为0; $CHIN = (NI_t - NI_{t-1}) \div (|NI_t| + |NI_{t-1}|)$, NI 表示净利润。

八、研究结论与讨论

1. 研究结论

本文以“国家大数据综合试验区”的建设为准自然实验,利用中国上市公司的数据,构建双重差分模型探究了数字经济背景下大数据发展的治税效应,研究发现:大数据发展显著提升了公司税收遵从水平。机制分析发现,大数据发展通过推动税收征管的数字化升级,抑制了公司隐匿数字化产品销售收入和异常关联交易,进而提高了公司税收遵从水平。异质性分析结果表明,大数据发展的治税效应对于内外部信息环境较差的公司和规模较大的公司更为显著。拓展性分析结果表明,大数据发展对公司税收遵从影响的长期效应大于短期效应。此外,大数据发展显著增加了地区企业所得税收入,提高了国家财政能力。经济后果检验发现,大数据发展还增加了公司通过避税来化解经营风险的成本。

2. 政策启示

本文丰富了试验区建设的经济后果和税收征管效率影响因素的相关理论文献,同时也进一步拓展了数字化技术和税收征管融合发展的相关研究。此外,本文的研究结论还有着重要的政策内涵:

第一,增强宏观政策取向一致性,推动数字产业政策与财税政策协调配合。本文研究发现数字经济发展可以为政府进行数字税收治理提供技术条件,而政府税收治理能力的提升又可以为继续推动数字经济发展提供物质基础。因此,应将数字化改革和财税改革纳入到同一框架中进行。一方面,在谋划新一轮财税体制改革时不能仅仅考虑财税本身,而是应立足于我国数字经济飞速发展的背景之下;另一方面,增强数字产业政策与财税政策目标和措施的一致性,强化利好政策协同效应,确保不同政策共同发力、形成合力。

第二,不断做强做优做大我国数字经济,释放数字经济的税收红利。首先,在总结现有八大试验区经验的基础上,可以继续扩大试点范围,因地制宜建立具有地区特色的大数据试验区,推动数字经济高质量发展。其次,深入推进试验区的建设,释放试验区的数据资源,以数据治理创新来推动税收治理和国家治理现代化。一方面,应加快数据资源化,继续依托试验区建立的数据共享交换和开放平台来打造规模大、类型多、价值高、颗粒度细的税收大数据,为建成“以数治税”分类监管模式打下坚实基础;另一方面,积极推动公共数据资源开发利用,加快公共数据资源安全有序开放,推动政务数据资源共享交换。

第三,强化数字经济税收监管体系,促进数字经济健康发展。数字经济模式多样性、业务复杂性以及交易的虚拟性等特点给税收征管造成了前所未有的挑战。同时,数字经济时代海量数据的生成以及互联网、大数据、云计算、人工智能和区块链等技术加速创新和广泛使用,也为税收征管的数字化升级和智能化改造创造了条件。因此,可以以税收大数据为驱动力,探索建立适应数字经济发展的税收监管和稽查新秩序。首先,应扩大税收治理情景下数据技术应用和再开发,以数字技术和数据共享“双轮”驱动,达成涉税大数据获取、共享、集成、处理、分析和有效运用,以建成强大的智慧税务。其次,结合数字经济涉税特征,创新税收征管模式。税务部门应全面建立基于数字经济的税务登记系统、针对不同的数字经济业态进行分类监管、加强税法宣传的广度和深度以及推动税收共治体系的建设。

第四,进一步加强重点领域税务风险防控与监管,精准实施税务监管。充分挖掘和利用大数据推进税收精准监管是节约税收征管资源、提升税收征管质效和营造公平公开公正税收营商环境的重要措施。本文研究发现公司信息不对称程度和公司规模是影响公司逃避税动机的重要因素。因此,税务机关可以将这两方面纳入到税收监管和稽查选案标准中,提高税收监管和执法的精确性。

参考文献

- [1] Wilson, C., and I. Mergel. Overcoming Barriers to Digital Government: Mapping the Strategies of Digital Champions [J]. *Government Information Quarterly*, 2022, 39, (2), 101681.
- [2] 陈艳利, 蒋琪. 数据生产要素视角下开放公共数据与企业创新——基于建立公共数据开放平台的准自然实验[J]. 北京: 经济管理, 2024, (1): 25-46.
- [3] 何雨可, 牛耕, 逯建, 赵国昌. 数字治理与城市创业活力——来自“信息惠民国家试点”政策的证据[J]. 北京: 数量经济技术经济研究, 2024, (1): 47-66.
- [4] Welch, E. W., C. C. Hinnant, and M. J. Moon. Linking Citizen Satisfaction with E-Government and Trust in Government [J]. *Journal of Public Administration Research and Theory*, 2005, 15, (3): 371-391.
- [5] 聂爱云, 潘孝虎. 数字经济能否提升地方政府治理能力? ——基于中国 275 个城市 2011—2019 年面板数据的实证检验 [J]. 北京: 经济社会体制比较, 2023, (6): 109-120.
- [6] 刘奥, 张双龙. 数字经济发展对财政透明度的影响研究——基于“宽带中国”战略的准自然实验[J]. 南京: 产业经济研究, 2022, (4): 46-58.
- [7] 余典范, 龙睿, 王超. 数字经济与边界地区污染治理[J]. 北京: 经济研究, 2023, (11): 172-189.
- [8] Neuman, E. L., and R. J. Sheu. Big Data Analytics in IRS Audit Procedures and Its Effects on Tax Compliance: A Moderated Mediation Analysis [J]. *The Journal of the American Taxation Association*, 2022, 44, (2): 97-113.
- [9] 李鑫钊. 数字经济背景下加强税收征管的建议[J]. 北京: 税务研究, 2023, (1): 129-132.
- [10] Wei, X., F. Jiang, and L. Yang. Does Digital Dividend Matter in China's Green Low-Carbon Development? Environmental Impact Assessment of the Big Data Comprehensive Pilot Zones Policy [J]. *Environmental Impact Assessment Review*, 2023, 101, (7), 107143.
- [11] 孙伟增, 毛宁, 兰峰, 王立. 政策赋能、数字生态与企业数字化转型——基于国家大数据综合试验区的准自然实验[J]. 北京: 中国工业经济, 2023, (9): 117-135.
- [12] 游家兴, 苏三妹. 打破数据“烟囱”: 大数据驱动的新质生产力与劳动投资效率[J]. 北京: 经济管理, 2024, (9): 5-29.
- [13] Pomeranz, D. No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax [J]. *American Economic Review*, 2015, 105, (8): 2539-2569.
- [14] 魏志华, 夏太彪. 社会保险缴费负担、财务压力与企业避税[J]. 北京: 中国工业经济, 2020, (7): 136-154.
- [15] 严若森, 钱晶晶, 祁浩. 公司治理水平、媒体关注与企业税收激进[J]. 北京: 经济管理, 2018, (7): 20-38.
- [16] Dečman, M., J. Stare, and M. Klun. E-Government and Cost-Effectiveness: E-Taxation in Slovenia [J]. *Transylvanian Review of Administrative Sciences*, 2010, 6, (31): 48-57.
- [17] Bhuasiri, W., H. Zo, and H. Lee et al. User Acceptance of E-government Services: Examining an E-tax Filing and Payment System in Thailand [J]. *Information Technology for Development*, 2016, 22, (4): 672-695.
- [18] Night, S., and J. Bananuka. The Mediating Role of Adoption of an Electronic Tax System in the Relationship Between Attitude Towards Electronic Tax System and Tax Compliance [J]. *World Journal of Economics, Finance and Administrative Science*, 2019, 25, (49): 73-88.
- [19] 张克中, 欧阳洁, 李文健. 缘何“减税难降负”: 信息技术、征税能力与企业逃税[J]. 北京: 经济研究, 2020, (3): 116-132.
- [20] Bellon, M., E. Dabla-Norris, and S. Khalid et al. Digitalization to Improve Tax Compliance: Evidence from VAT e-Invoicing in Peru [J]. *Journal of Public Economics*, 2022, 210, (6), 104661.
- [21] Okunogbe, O. M., and V. M. J. Pouliquen. Technology, Taxation, and Corruption: Evidence from the Introduction of Electronic Tax Filing [J]. *American Economic Journal: Economic Policy*, 2022, 14, (1): 341-372.
- [22] Anomah, S., B. Ayebofo, and M. Aduamoah et al. Blockchain Technology Integration in Tax Policy: Navigating Challenges and Unlocking Opportunities for Improving the Taxation of Ghana's Digital Economy [J]. *Scientific African*, 2024, 24, (2), e02210.
- [23] 冯绍伍, 江峰, 杨智曾. 大数据思维下实现精准税务监管的思考与探索[J]. 北京: 税务研究, 2022, (11): 83-88.
- [24] 姚东旭, 严亚雯. 数字时代的税收治理: 关联性、现实挑战与应对方案[J]. 长春: 税务与经济, 2023, (6): 19-25.
- [25] Goldfarb, A., and C. Tucker. Digital Economics [J]. *Journal of Economic Literature*, 2019, 57, (1): 3-43.
- [26] 高博楠, 王志刚. 关于完善涉税数据共享制度的几点思考[J]. 北京: 税务研究, 2023, (2): 141-144.
- [27] 陈思霞, 卢盛峰, 王冰璇. 跨部门信息共享与地方税收收入增长: 基于联合激励的视角[J]. 北京: 财政研究, 2021, (7): 98-113.
- [28] 洪剑峭, 薛皓. 股权制衡如何影响经营性应计的可靠性——关联交易视角[J]. 北京: 管理世界, 2009, (1): 153-161.

- [29]张乾,葛国庆,薛健.数字经济促进了企业避税吗——基于电子商务示范城市创建的准自然实验[J].北京:会计研究,2022,(4):71-88.
- [30]刘慧龙,张玲玲,谢婧.税收征管数字化升级与企业关联交易治理[J].北京:管理世界,2022,(6):158-176.
- [31]李平.运用大数据推动税收监管创新的思考[J].北京:国际税收,2020,(12):54-58.
- [32]Desai, M. A., and D.Dharmapala. Corporate Tax Avoidance and High-Powered Incentives[J]. *Journal of Financial Economics*, 2006, 79, (1): 145-179.
- [33]李青原,邹秉辰,肖泽华.破产法律制度改革对企业避税的影响——来自《企业破产法》实施的经验证据[J].北京:经济管理,2022,(12):185-202.
- [34]Nguyen, M., and J.H.Nguyen. Economic Policy Uncertainty and Firm Tax Avoidance[J]. *Accounting & Finance*, 2020, 60, (4): 3935-3978.
- [35]于文超,殷华,梁平汉.税收征管、财政压力与企业融资约束[J].北京:中国工业经济,2018,(1):100-118.
- [36]Zimmerman, J.L. Taxes and Firm Size[J]. *Journal of Accounting and Economics*, 1983, (5): 119-149.
- [37]Hanlon, M., E.L.Maydew, and D.Saavedra. The Taxman Cometh: Does Tax Uncertainty Affect Corporate Cash Holdings? [J]. *Review of Accounting Studies*, 2017, 22, (3): 1198-1228.
- [38]陈作华,方红星.融资约束、内部控制与企业避税[J].哈尔滨:管理科学,2018,(3):125-139.
- [39]丁方飞,刘敏,乔紫薇.事务所规模、制度环境与企业税收激进[J].南京:审计与经济研究,2019,(1):14-23.
- [40]陈德球,陈运森,董志勇.政策不确定性、税收征管强度与企业税收规避[J].北京:管理世界,2016,(5):151-163.
- [41]Roth, J., P. H. C. Sant'Anna, and A. Bilinski et al. What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature[J]. *Journal of Econometrics*, 2023, 235, (2): 2218-2244.
- [42]许文立,孙磊.市场激励型环境规制与能源消费结构转型——来自中国碳排放权交易试点的经验证据[J].北京:数量经济技术经济研究,2023,(7):133-155.
- [43]Lyu, Y., X.Xiao, and J. Zhang. Does the Digital Economy Enhance Green Total Factor Productivity in China? The Evidence from a National Big Data Comprehensive Pilot Zone[J]. *Structural Change and Economic Dynamics*, 2024, 69, (2): 183-196.
- [44]Allingham, M. G., and A. Sandmo. Income Tax Evasion: A Theoretical Analysis[J]. *Journal of Public Economics*, 1972, 1, (3): 323-338.
- [45]Conley, T. G., C. B. Hansen, and P. E. Rossi. Plausibly Exogenous[J]. *The Review of Economics and Statistics*, 2012, 94, (1): 260-272.
- [46]Blundell, R., and M. Costa Dias. Evaluation Methods for Non-Experimental Data[J]. *Fiscal Studies*, 2005, 21, (4): 427-468.
- [47]Hainmueller, J. Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies[J]. *Political Analysis*, 2012, 20, (1): 25-46.
- [48]范子英,朱星姝,冯晨.去属地化与企业税负:基于国地税合并的研究[J].北京:财贸经济,2022,(10):23-39.
- [49]范源源,李建军.税收“放管服”改革能否提高企业税收遵从[J].北京:财贸经济,2023,(1):68-82.
- [50]贺晓宇,储德银.政府治理数字化转型与城市创业活跃度提升[J].上海经济研究,2023,(9):41-53.
- [51]江轩宇.税收征管、税收激进与股价崩盘风险[J].天津:南开管理评论,2013,(5):152-160.
- [52]孙鲲鹏,石丽娜.企业互联网使用与大数据治税的效应[J].北京:经济研究,2022,(5):176-191.
- [53]宫晓云,权小锋,刘希鹏.供应链透明度与公司避税[J].北京:中国工业经济,2022,(11):155-173.
- [54]范源源,李建军.资本市场开放与公司税收遵从——来自“陆港通”交易制度实施的经验证据[J].北京:经济管理,2022,(3):24-38.
- [55]张纯,吕伟.信息环境、融资约束与现金股利[J].北京:金融研究,2009,(7):81-94.
- [56]Clement, M. B. Analyst forecast accuracy: Do Ability, Resources, and Portfolio Complexity Matter? [J]. *Journal of Accounting and Economics*, 1999, 27, (3): 285-303.
- [57]赵仁杰,周小昶.区域性税收优惠与企业异地投资的避税效应[J].北京:财贸经济,2022,(11):59-74.
- [58]Chaurey, R. Location-Based Tax Incentives: Evidence from India[J]. *Journal of Public Economics*, 2017, 156, (12): 101-120.
- [59]毛志宏,李燕,金龙,哈斯乌兰.客户资产减值与供应商企业投资[J].天津:南开管理评论,2024,(1):179-191.
- [60]林炳洪,李秉祥.ESG责任履行对企业经营困境的影响:“雪中送炭”还是“雪上加霜”? [J].北京:中国软科学,2024,(6):121-130.

Digital Empowerment: The Tax Treatment Effect of Big Data Development

FAN Yuan-yuan¹, LI Jian-jun²

(1.School of Public Finance and Taxation, Dongbei University of Finance and Economics, Dalian, Liaoning, 116025, China;
2.School of Public Finance and Taxation, Southwestern University of Finance and Economics, Chengdu, Sichuan, 611130,
China)

Abstract: In recent years, the convergence and integration of information technology and the economy and society have triggered a rapid growth in data, which has become a fundamental strategic resource of the State, and big data are increasingly having an important impact on global production, circulation, distribution and consumption activities, as well as on the mechanism for the operation of the economy, the way of life of the society, and the ability of the State to govern. The rapid development of the digital economy not only provides new momentum for the high-quality development of real industries, but also provides new opportunities for building a digital government and achieving the government's macro-control goals. However, the impact on fiscal governance, which is an important part of the government's governance system, is not yet clear. Theoretically, the implementation of big data industrial policy can promote the integration and development of digital technology and tax administration, and thus enhance corporate tax compliance.

Therefore, based on the data of listed companies in China from 2010 to 2022, this paper examines the impact of location-oriented policies based on big data on corporate tax compliance in the context of the development of digital economy, using the construction of the "National Comprehensive Pilot Zone for Big Data" as a quasi-natural experiment. The study finds that the development of big data significantly improves corporate tax compliance, and this finding is still valid in a series of robustness tests. Mechanistic analyses suggest that big data development ultimately enhances corporate tax compliance by driving digital upgrades in tax collection and management, which in turn inhibits unusual connected transactions and the concealment of revenues from the sale of digitised products. Heterogeneity analyses based on information asymmetry theory and tax classification and regulation show that the effect is concentrated in firms with poor internal and external information environments, firms with large asset sizes, and firms with large group sizes. In addition, the long-term effects of big data development on corporate tax compliance are greater than the short-term effects. From the perspective of tax governance benefits, the development of big data significantly increases government tax revenues, improves the country's fiscal capacity, and provides impetus for the continued promotion of the digital economy. The economic consequences test also finds that big data development also increases the cost to firms of mitigating business risks through tax avoidance.

Compared with previous studies, this paper makes the following three contributions. First, this paper adds to the literature on the economic consequences of the construction of the National Big Data Comprehensive Pilot Zone. Previous literature on the economic consequences of the construction of pilot zones has mainly focused on corporate labor productivity, total factor productivity, and urban economic development. This paper verifies the impact of pilot zone construction on corporate tax compliance, expanding the research framework of the existing literature. Second, this paper expands the research related to the influencing factors of corporate tax compliance. Previous literature explaining the influencing factors of corporate tax compliance is mostly based on corporate financing constraints, information transparency and agency problems. This paper explores the specific ways in which the development of big data affects corporate tax decisions from the perspectives of unusual connected transactions and hidden income. Third, this paper also provides policy references to help modernize tax governance and modernize national governance. At present, there is a lack of systematic empirical analysis of the relationship between big data development and tax administration digitization. Through empirical research, this paper finds that big data development can promote the digitalization of tax collection and administration by facilitating the digitalization of the economy and society, which provides new evidence for the further construction of a tax collection and administration system that adapts to the development of the digital economy.

Key Words: new productive force; digital government; big data; taxation by numbers; corporate tax avoidance

JEL Classification: G38, H26

DOI: 10.19616/j.cnki.bmj.2024.12.003

(责任编辑:张任之)