

# 大数据时代的政府治理： 数字政府与企业研发操纵

孟元 杨蓉\*

**内容提要** 本文基于2014–2021年沪深A股上市公司样本,结合《高新技术企业认定管理办法》这一产业政策的实施,运用群聚估计检验企业是否借助研发操纵以获得高新技术企业认定。在此基础上,以各地区大数据管理机构改革作为数字政府建设的准自然实验,运用双重差分法考察数字政府对企业研发操纵的影响。研究发现,中国上市公司中存在较为明显的研发操纵现象。然而,数字政府建设显著抑制了企业研发操纵。原因在于,数字政府建设有效地提升了政府对企业研发操纵的监管能力。此外,在非国有企业、互联网和新闻媒体监督难以发挥作用的样本中,数字政府抑制企业研发操纵的作用更强。本文为政府在数字时代更好发挥产业政策的作用提供理论和经验证据。

**关键词** 数字政府 研发操纵 产业政策 大数据管理机构改革

## 一 引言

党的二十大报告提出要“坚持创新在我国现代化建设全局中的核心地位”。由于研发活动具有正外部性(Romer, 1990),并且研发活动的不确定性很高(Bloom *et al.*, 2019),企业的研发活动往往会出现投入不足的现象(Arrow, 1962; Jones and

\* 孟元、杨蓉(通讯作者):华东师范大学经济与管理学院 上海市中山北路3663号 200062 电子信箱:mengyuan\_ecnu@126.com(孟元);ryang@dbm.ecnu.edu.cn(杨蓉)。

作者感谢国家自然科学基金面上项目(71573090)的资助。感谢匿名审稿专家的宝贵意见和建议,文责自负。

Williams, 1998; Bloom *et al.*, 2013)。各国政府为激励企业创新,采取多种措施以缓解企业研发投入不足问题,其中税收优惠已被证明是一种激励研发活动的有效手段(Bloom *et al.*, 2002; Bloom *et al.*, 2019; Akcigit *et al.*, 2022)。因此,中国政府在2008年颁布《高新技术企业认定管理办法》,将高新技术企业的所得税税率降为15%,远低于一般企业25%的税率。

政府给予高新技术企业数量可观的政策优惠,但是政府受限于高昂的监管成本,很难发现企业为了达到高新技术企业资格认定而进行的研发操纵行为,导致一些研发能力较低的企业有动机通过研发操纵的方式将自己包装成高新技术企业,以此牟取政策套利(杨国超等,2017;杨国超和芮萌,2020;Chen *et al.*, 2021)。针对上述问题,本文重点关注随着数字政府的建设,特别是大数据等新一代数字技术应用于政府治理,帮助政府实现用数据说话、用数据管理、用数据决策,是否有助于减少企业研发操纵这一机会主义行为。

本文首先从理论层面探讨数字政府如何影响企业研发操纵行为。在Acemoglu *et al.*(2022)理论模型的基础上,本文首先推导出在政府没有出台《高新技术企业认定管理办法》的情况下,企业的最优研发强度与其研发能力呈正相关,可以解释该政策为何设置研发投入强度的门槛条件。当政府出台《高新技术企业认定管理办法》后,一些研发能力较低的企业有动机通过研发操纵的方式达到政策规定的门槛条件,导致获得高新技术企业认定的企业鱼龙混杂——不仅有政府希望支持的高新技术企业,也有研发能力较低的伪高新技术企业。然而,随着数字政府的建设,政府更有可能发现企业的研发操纵行为,致使研发能力较低的企业进行研发操纵的可能性下降,政策能够更好地发挥效果。

从经验分析层面,本文借助近年来最新发展的群聚估计(Cengiz, 2019;樊勇等, 2020;Chen *et al.*, 2021; 范子英等, 2022),发现研发操纵现象在中国上市公司中较为普遍。接下来,本文选取第八轮政府机构改革中各地区改革大数据管理机构作为准自然实验,运用双重差分法进行检验。研究结果表明,数字政府显著抑制了企业研发操纵。为了确认分析结果的可信性,本文进行了一系列稳健性检验,包括进行双重差分异质性检验、平行趋势检验和安慰剂检验、构造三重差分识别策略等,均得到一致的结论。机制检验结果显示,数字政府建设有效地提升了政府对企业研发操纵的监管能力,具体表现为:(1)政府可以更有效地发挥大数据等数字技术对企业研发操纵的治理作用;(2)企业更难通过研发操纵获得高新技术企业认定;(3)研发操纵企业更少将管理费用虚假标记为研发费用。本文最后的异质性分析发现,数字政府抑制企

业研发操纵的作用在非国有企业、互联网监督和新闻媒体监督难以发挥作用的企业中更强。

和以往文献相比,本文的边际贡献体现在:

第一,产业政策有效性的争论贯穿经济学的发展历程,特别是林毅夫和张维迎的产业政策之争,核心论点是一方认为市场会失灵,强调政府的作用,另一方则持相反的论点。然而,最近产业政策的讨论逐渐从关注“产业政策是否有效”转向“产业政策如何更有效”(戴小勇和成力为,2019;杨瑞龙和侯方宇,2019;杨国超和张李娜,2021)。本文在此基础上提出,随着数字政府的建设,政府可以更有效地挖掘海量数据背后蕴含的经济和社会运行规律,从而解决在产业政策实施过程中存在的政府失灵问题。这一发现有助于更好地理解大数据时代下政府与市场的关系。

第二,大数据时代政府治理的文献主要聚焦于数字政府提升政府治理能力(江小涓,2018;黄璜,2020)、数字政府建设存在的问题(王伟玲,2022)以及如何更好地建设数字政府(北京大学课题组,2020;米加宁等,2020;孟庆国等2020;黄璜等,2022)。然而,鲜有学者探讨数字政府的经济影响,特别是数字政府如何影响微观企业行为。本文重点关注数字政府对企业研发操纵行为的影响,帮助我们更好地理解数字政府的经济影响。更重要的是,通过深入理解数字政府的经济影响,又能反作用于更好地建设数字政府。

第三,数字化变革影响企业研发活动的文献主要集中在数字经济、数字金融、数据要素等方面(唐松等,2020;Cong *et al.*, 2021;唐要家等,2022;Bejara *et al.*, 2023a; Bejara *et al.*, 2023b)。然而,鲜有学者深入探讨数字政府对企业研发活动的影响。由于数字政府建设全面引领数字化变革,本文以数字政府作为切入点展开研究,有助于全面理解数字化变革对企业研发活动的影响。

## 二 制度背景与文献回顾

### (一)高新技术企业认定管理办法

为了更好地鼓励和支持高新技术企业的发展,引导高新技术企业从事研发创新,中国科学技术部、财政部和国家税务总局于2008年联合颁布《高新技术企业认定管理办法》(下文简称高企认定政策)。该政策规定,通过认定的高新技术企业(下文简称高企),所得税税率由25%减至15%。该政策同时规定,申请高新技术企业认定需要满足一系列门槛条件,其中最重要的是企业的研发投入占销售收入的比例需要达

到一定的门槛条件,具体为最近一年销售收入小于5000万元的企业,研发投入占销售收入的比例不得低于5%;最近一年销售收入在5000万元至2亿元的企业,比例不得低于4%;最近一年销售收入在2亿元以上的企业,比例不得低于3%。

## (二)大数据与数字政府建设

谷歌总裁施密特在2010年的美国加州科技经济会议上表示:“当今世界每两天产生的数据相当于2003年以前人类历史产生的数据总和。”<sup>①</sup>随着近年来深度学习算法的迅猛发展和计算机计算能力的巨大提升,人类社会逐渐能够有效处理和利用海量的数据,让大数据等数字技术越来越深入渗透到经济和社会运行的各个方面(Jones and Tonetti, 2020)。大数据等数字技术的典型应用甚至能够辅助乃至替代人类的工作(Acemoglu and Restrepo, 2018; Aghion *et al.*, 2019),例如自动驾驶(LeCun *et al.*, 2015)、大语言模型(如GPT)、气象大模型(Bi *et al.*, 2023)等。

数字政府建设强调将大数据等数字技术应用于政府治理。2017年底,习近平总书记论述国家大数据战略时,指出“要运用大数据提升国家治理现代化水平,要建立健全大数据辅助科学决策和社会治理的机制,实现政府决策科学化”。此后,数字政府迅速在广东、浙江等地方治理中推开,并且大数据一词成为新时代数字政府建设最主要的技术特征(黄璜, 2020)。2022年6月,国务院进一步颁布《关于加强数字政府建设的指导意见》,对数字政府建设做出了明确规范,指出建设数字政府要“强化经济运行大数据分析,提升经济调节能力”,“大力推行智慧监管,提升市场监管能力”,“建立健全大数据辅助科学决策机制,全面提升政府决策科学化水平”<sup>②</sup>。孟天广(2022)认为数字政府是政府运用数据和算法理解经济社会运行状态,并基于大数据为经济运行和社会治理提供决策辅助,推进精准化决策和靶向性治理。

## (三)数字政府建设提升政府治理能力

随着数字政府建设,特别是借助大数据等新一代数字技术,极大地提升了政府的治理能力(黄璜, 2020; 高奇琦, 2023)。具体而言,政府可以利用大数据技术挖掘有关市场运行的海量数据,更好地探究海量数据背后蕴含的经济、社会运行的规律(江小涓, 2018),从而起到辅助监管机构决策、提升监管精细化水平和监管效能的作用(北京大学课题组, 2020; 黄璜, 2020)。在社会治理方面,大数据帮助政府提升治理效率。Mastrobuoni(2020)发现政府利用大数据技术预测犯罪最有可能发生的地点,并

<sup>①</sup> 参见 <https://www.wsj.com/articles/googles-looming-battle-over-search-1430263075>。

<sup>②</sup> 参见 [http://www.gov.cn/zhengce/content/2022-06/23/content\\_5697299.htm](http://www.gov.cn/zhengce/content/2022-06/23/content_5697299.htm)。

将警力有针对性地部署在这些地点,提高政府治理犯罪的效率。在企业治理方面,大数据同样发挥关键作用。张克中等(2020)发现,大数据等新一代数字技术大幅提升了税务机关对涉税信息的监管能力,让税务机关更容易发现企业的逃税行为,降低了企业逃税的可能性。陈德球和胡晴(2022)发现,证监会等监管机构可以借助大数据等数字技术提升监管效率,精准识别财务造假的企业。例如,在獐子岛财务舞弊案中,监管机构使用卫星大数据对獐子岛公司进行监控和数据分析,通过大数据技术复原公司真实的扇贝捕捞养殖情况,发现真实捕捞区域和会计记录中记载的捕捞区域存在明显出入。

江小涓(2018)指出,传统监管受制于监管成本,政府面临着管与不管的两难境地,监管最后变成了“人在做,天在看”。然而,随着大数据等新一代数字技术应用于政府治理,监管能够实现“人在做,网在看,云在算”,政府可以更清楚、更全面地掌握被监管对象的行为。因此,随着数字政府的建设,政府可以运用大数据等数字技术,识别出最有可能进行研发操纵的企业,并将监管重点有针对性地放在这些企业上,避免了监管过程中“眉毛胡子一把抓”,进而提升政府对企业研发操纵行为的监管能力。例如,2022年福建省高企认定工作领导小组利用大数据排查,发现有226家企业可能是伪高新技术企业,并对其中58家企业进行了现场核查,结果只有7家企业通过了核查<sup>①</sup>。

### (四)大数据管理机构改革与数字政府建设

#### 1. 第八轮政府机构改革中的大数据管理机构改革

在十九届三中全会提出“深化党和国家机构改革”的基础上,2018年3月党中央印发《深化党和国家机构改革方案》明确提出“允许地方在规定限额内因地制宜设置机构和配置职能”。在该文件的指导下,省级政府开始了第八轮政府机构改革,其中大数据管理机构改革成为此轮省级政府机构改革最大的亮点之一——为了顺应数字技术迅速发展的浪潮,一些省份和直辖市进行了大数据管理机构改革,旨在实现数字技术与政府治理的深度融合(米加宁等,2020)。截至2021年末,已有超过20个省份和直辖市进行了大数据管理机构的改革,负责推进数字政府建设工作(王伟玲,2022)。在地方政府改革经验的基础上,中央政府在2023年成立了国家数据局,希望从国家层面解决数据孤岛和重复建设等瓶颈(高小平,2023)。

#### 2. 数字政府建设存在的问题

虽然数字政府建设提升了政府的治理能力,但是受制于数据孤岛和重复建设数

<sup>①</sup> 参见 [http://fujian.chinatax.gov.cn/fzsswj/zfxxgkzl/zfxxgkml/tajybl001/202303/t20230328\\_502209.htm](http://fujian.chinatax.gov.cn/fzsswj/zfxxgkzl/zfxxgkml/tajybl001/202303/t20230328_502209.htm)。



字政府等问题,数字政府发挥的作用大大降低。首先是数据孤岛的问题,数字化转型最大的挑战是数据孤岛问题,这同样是数字政府建设的主要挑战(姚期智和郁昱,2022)。尽管政府拥有80%的社会数据资源,是社会数据资源的主要拥有者(黄璜和孙学智,2018;王伟玲,2021),政府却无法充分利用如此大规模的数据资源——大量的数据分散在政府的不同部门之中,并且部门之间各自为政,导致数据难以实现跨部门的共享和利用,甚至同一部门内的数据有时都难以共享交换(王伟玲,2022)。正是受制于如此严重的数据孤岛瓶颈,数字政府发挥的作用大幅降低<sup>①</sup>。其次是数字政府重复建设的问题。大数据管理机构改革前,各个部门通常都要成立自己的信息化管理部门,导致严重的重复建设问题,降低了数字政府建设的效率,例如广东省在改革前,55个省直部门承担信息化工作的机构有44个,重复建设的问题十分严重(黄璜,2020)。

### 3. 大数据管理机构改革赋能数字政府建设

为了解决数据孤岛和数字政府重复建设等问题,全面提升数字政府发挥的作用,地方政府在第八轮政府机构改革中,重点进行了大数据管理机构改革(米加宁等,2020)。

大数据管理机构改革有助于解决数据孤岛的问题。经过改革后,大数据管理机构具有一定的统筹协调的权力,专门负责统筹管理部门间数据,从而实现数据在不同部门和地区之间的共享利用,提升数据的利用效率,打破数据孤岛瓶颈(米加宁等,2020;孟庆国等,2020;王伟玲,2022)。例如,上海市大数据管理机构的主要职能包括:(1)承担政务数据、行业数据、社会数据等各方数据归集和应用融合工作,开展大数据应用研究工作,为本市政府各部门管理、服务、决策提供数据支撑;(2)建设全市统一政务数据共享交换平台,开展跨地区、跨层级、跨部门的数据共享交换和利用。可以看出,这些职能与解决数据孤岛问题密切相关。

大数据管理机构改革有助于解决重复建设数字政府的问题。经过改革后,大数据管理机构负责统筹数字政府建设,实现专业的人干专业的事,从而减少数字政府的重复建设,提高数字政府的建设效率(张克,2019;黄璜,2020)。例如,广东省在改革后,由专门的大数据管理机构来统一负责其他部门的数字化工作。这样不仅减少了数字政府的重复建设,还能集中相关的资源、技术、人力和能力,产生协同效应,更好

---

<sup>①</sup> 限于篇幅,此处略去关于数据孤岛问题让数字技术难以发挥作用的相关解释,对此感兴趣的读者可访问《世界经济》网站([www.jweonline.cn](http://www.jweonline.cn))2024年第1期在线期刊中本文补充材料附录一。后文简称见网站。

地为其他部门提供统一的数字化管理和服务(黄璜,2020)。

可以看出,大数据管理机构改革全面提升了数字政府发挥的作用。具体到对企业研发操纵的监管,首先,数据孤岛问题极大地限制了科技局等政府部门利用大数据等数字技术对企业研发操纵的监管。大数据管理机构改革的主要目标是解决数据孤岛瓶颈,以便科技局等政府部门进行监管时能够充分利用其他部门的数据资源,例如,叶战备(2021)发现大数据管理机构通过整合市场监管局、税务局、发改委、人社局、环保局等政府部门的数据资源,以及建设数据共享平台来实现政务数据在不同政府部门之间的共享和利用。此时,科技局等政府部门可以借助不同的数据源进行交叉验证,更准确地还原真实情况(江小涓,2018),进而更容易发现企业的研发操纵行为。以福建省的高企申报推荐流程为例,科技局借助大数据技术检查申报企业的科技人员个税和社保缴纳情况,一旦发现异常,科技局将会深入展开调查,并且对不符合申报条件的企业不予以推荐认定为高企。由此可见,解决数据孤岛问题,实现政务数据共享利用之后,企业的研发操纵行为将更容易被发现。

其次,只有熟练掌握和运用数字技术,才能充分发挥数据的价值,让数据成为国家治理的重要力量(黄其松,2022)。大数据管理机构通过统筹数字政府建设,为科技局等政府部门提供更高效率的数字化服务,更好地发挥大数据等数字技术对企业研发操纵的治理作用。例如,福州市科技局在高企申报推荐流程以及向获得高企资格的企业拨付奖励资金时,委托大数据管理机构对企业的信用状况进行筛查,帮助监管企业研发操纵<sup>①</sup>。

### 三 理论模型

#### (一)模型设定

本文借鉴 Acemoglu *et al.* (2022) 的模型设定,构建数字政府降低企业研发操纵的理论模型。

首先假设某经济体只有一种最终消费品,对最终消费品的跨期效用可以表示为:

$$U_t = \rho^t \frac{C_t^{1-v} - 1}{1-v} \quad (1)$$

其中  $C_t$  是在时刻  $t$  该经济体对最终消费品的消费量,  $v$  是跨期替代弹性的倒数,

<sup>①</sup> 参见 [http://fujian.chinatax.gov.cn/fzsswj/zfxxgkzl/zfxxgkml/tajybl001/202303/t20230328\\_502209.htm](http://fujian.chinatax.gov.cn/fzsswj/zfxxgkzl/zfxxgkml/tajybl001/202303/t20230328_502209.htm)。

$\rho \in (0, 1)$  是折现系数。

假设使用劳动力和位于集合  $l$  的连续统的中间产品来生产最终消费品。对于任意  $j \in l$  的中间产品, 都由一个处于技术前沿的垄断企业  $j$  提供<sup>①</sup>。假设最终消费品的生产函数可以表示为:

$$Y_t = \frac{1}{1 - \alpha} \left( \int_l q_j^\alpha k_j^{1-\alpha} dj \right) L^\alpha \quad (2)$$

其中,  $k_j$  表示生产过程中使用中间产品  $j$  的数量。  $q_j$  表示中间产品  $j$  的生产效率, 反映中间产品  $j$  用于生产最终消费品时的生产效率。  $L$  是劳动力的数量, 假设该经济体中劳动力的数量保持不变并且劳动力的供给无弹性。该经济体在时刻  $t$  的消费可以表示为:

$$C_t = Y_t - \int_l p_j k_j dj \quad (3)$$

其中  $p_j$  表示中间产品  $j$  的价格, 该价格由企业  $j$  来制定, 使用最终消费品作为计价物。

根据(1)、(2)和(3)式, 在时刻  $t$ , 对于给定的价格向量  $P$ , 经济体的最优决策可以表示为:

$$\max_k \frac{\rho^t \left\{ \left[ \frac{1}{1 - \alpha} \left( \int_l q_j^\alpha k_j^{1-\alpha} dj \right) L^\alpha - \int_l p_j k_j dj \right]^{1-v} - 1 \right\}}{1 - v} \quad (4)$$

等价于:

$$\max_k \frac{1}{1 - \alpha} \left( \int_l q_j^\alpha k_j^{1-\alpha} dj \right) L^\alpha - \int_l p_j k_j dj \quad (5)$$

对(5)式进行逐点最优, 可得经济体对中间产品  $j$  的需求曲线为:

$$p_j = L^\alpha q_j^\alpha k_j^{-\alpha}, \quad \forall j \in l \quad (6)$$

接下来, 我们考虑企业  $j$  的决策。假设企业  $j$  以固定的边际成本  $\gamma$  生产中间产品  $j$ 。由于企业  $j$  将需求函数(6)式视为给定的, 并据此最大化其利润, 因此其最优化问题可以表示为:

$$\max_{k_j} p_j k_j - \gamma k_j = L^\alpha q_j^\alpha k_j^{1-\alpha} - \gamma k_j, \quad \forall j \in l \quad (7)$$

求解一阶条件, 可得企业  $j$  的产量为:

<sup>①</sup> 为了简化符号, 中间产品和生产该中间产品的垄断企业都使用相同的下标  $j$  来表示。



$$k_j = \left( \frac{1 - \alpha}{\gamma} \right)^{\frac{1}{\alpha}} Lq_j, \quad \forall j \in l \quad (8)$$

利用(6)、(7)和(8)式,可得企业 $j$ 的利润函数和收入函数为:

$$\Pi(q_j) = \alpha \left( \frac{1 - \alpha}{\gamma} \right)^{\frac{1-\alpha}{\alpha}} Lq_j = \pi q_j, \quad \forall j \in l \quad (9)$$

$$I(q_j) = \left( \frac{1 - \alpha}{\gamma} \right)^{\frac{1-\alpha}{\alpha}} Lq_j = \frac{\pi q_j}{\alpha}, \quad \forall j \in l \quad (10)$$

其中,  $\pi = \alpha \left[ (1 - \alpha) / \gamma \right]^{\frac{1-\alpha}{\alpha}} L$  为常数。由(9)和(10)式可以看出,企业 $j$ 的利润和收入可以表示为其生产效率 $q_j$ 的函数。

## (二)企业研发决策

假设模型一共持续两期。假设第一期所有企业的生产率是同质的,即 $q_{j1} \equiv \bar{q}_1$ 。根据(9)和(10)式,企业 $j$ 在第一期的利润为 $\Pi(q_{j1}) \equiv \pi \bar{q}_1$ ,在第一期的收入为 $I(q_{j1}) \equiv \frac{\pi \bar{q}_1}{\alpha}$ 。

假设第二期,企业 $j$ 的生产效率由下式决定:

$$q_{j2} = \tau \bar{q}_1 + \left[ \theta_j \ln(1 + R_{j1}) \right] \cdot e^{\varepsilon_2}, \quad \forall j \in l \quad (11)$$

其中 $R_{j1} > 0$ 表示企业 $j$ 在第一期的研发投入。随机变量 $\varepsilon_2$ 服从均值为0,方差为 $\sigma^2$ 的正态分布,用于刻画企业研发过程中的不确定性。 $\tau < 1$ 用于捕捉技术资本随着时间的折旧(Chen *et al.*, 2021)。 $\theta_j$ 是本文重点关注的参数,用于刻画企业研发能力的异质性。假设企业 $j$ 的研发能力 $\theta_j$ 是私有信息,也就是说,政府等外部人无法直接观测到 $\theta_j$ 的大小。

根据(9)式,我们写出企业 $j$ 在第二期的期望利润: $E[\Pi(q_{j2})] = \pi [\tau \bar{q}_1 + e^{\frac{\sigma^2}{2}} \theta_j \ln(1 + R_{j1})]$ 。我们首先考虑当所得税税率恒为 $t_H$ 时,企业 $j$ 的研发决策。在两期模型中,企业 $j$ 的跨期最优化问题可以表示为:

$$\max_{R_{j1}} (1 - t_H)(\pi \bar{q}_1 - R_{j1}) + (1 - t_H) \rho \pi \left[ \tau \bar{q}_1 + e^{\frac{\sigma^2}{2}} \theta_j \ln(1 + R_{j1}) \right] \quad (12)$$

求解一阶条件,可得:

$$R_{j1}^*(\theta_j) = e^{\frac{\sigma^2}{2}} \rho \pi \theta_j - 1, \quad \forall j \in l \quad (13)$$

由于高企认定政策针对企业研发投入强度(企业研发投入比销售收入)设置门槛条件,因此我们在(13)式的左右两端同时除以企业的收入,得到企业 $j$ 的最优研发

强度:

$$r_{j1}^*(\theta_j) = \frac{R_{j1}^*(\theta_j)}{\frac{\pi\bar{q}_1}{\alpha}} = e^{\frac{\beta}{\alpha}} \frac{\rho\alpha\theta_j}{\bar{q}_1} - \frac{\alpha}{\pi\bar{q}_1}, \quad \forall j \in l \quad (14)$$

为简化起见,重新表述(14)式为:

$$r_{j1}^*(\theta_j) = A\theta_j - B, \quad \forall j \in l \quad (15)$$

其中,  $A = e^{\frac{\beta}{\alpha}} \frac{\rho\alpha}{\bar{q}_1}$  和  $B = \frac{\alpha}{\pi\bar{q}_1}$  为固定常数。

由(15)式可以看出,企业的最优研发强度取决于其研发能力( $\theta_j$ ),企业的研发能力越强,企业的最优研发强度也越大。

### (三)高企认定政策与企业研发操纵

(15)式可以解释高企认定政策为何设置研发投入强度的门槛条件。政府希望筛选出研发能力强的企业,并激励这部分企业进行研发创新。然而,政府作为外部人,受制于信息不对称,无法观测到企业真实的研发能力。不过,由(15)式可以看出,在政府未出台高企认定政策的条件下, $r^*(\theta)$ 关于 $\theta$ 严格单调递增,也就是说,企业的最优研发强度和企业的研发能力正相关。因此,假如政府希望激励研发能力强的企业(例如, $\theta > \theta^s$ )进行研发创新,政府可以设置企业研发投入强度的门槛条件: $\bar{r} = A\theta^s + B$ ,并规定:(1)企业的研发投入强度只有达到高企认定政策规定的门槛条件( $r \geq \bar{r}$ ),才有资格申请高企认定;(2)当企业获得高企认定之后,在第二期可以享受税率为 $t_L$ 的低税率( $t_L < t_H$ )。

对于 $r^*(\theta) < \bar{r}$ 的企业来说,可以选择进行研发操纵以达到高企认定政策规定的门槛条件,也可以选择进行研发操纵。假设企业申请高企认定需要支付固定的认定成本 $c \frac{\pi\bar{q}_1}{\alpha}$ ,同时对于研发操纵企业来说,申请高企认定成功的概率为 $1 - P$ 。

假设存在研发能力为 $\theta^*$ 的企业,在进行研发操纵和不进行研发操纵之间无差异,即有:

$$\begin{aligned} & (1 - t_H) \left[ \pi\bar{q}_1 - \frac{\pi\bar{q}_1}{\alpha} (\bar{r} + c) \right] + (1 - P)(1 - t_L) \rho\pi \left[ \tau\bar{q}_1 + e^{\frac{\beta}{\alpha}} \theta^* \ln \left( 1 + \frac{\pi\bar{q}_1 \bar{r}}{\alpha} \right) \right] \\ & + P(1 - t_H) \rho\pi \left[ \tau\bar{q}_1 + e^{\frac{\beta}{\alpha}} \theta^* \ln \left( 1 + \frac{\pi\bar{q}_1 \bar{r}}{\alpha} \right) \right] \\ & = (1 - t_H) \left[ \pi\bar{q}_1 - R^*(\theta^*) \right] + (1 - t_H) \rho\pi \left\{ \tau\bar{q}_1 + e^{\frac{\beta}{\alpha}} \theta^* \ln [1 + R^*(\theta^*)] \right\} \end{aligned} \quad (16)$$

其中,根据(13)式可知: $R^*(\theta^*) = e^{\frac{\beta}{\alpha}} \beta \pi \theta^* - 1$ 。

接下来,我们证明研发能力  $\theta \in (\theta^*, \theta^s)$  的企业为了获得高企认定,将会进行研发操纵。注意到(16)式左端是企业进行研发操纵时享有的收益,(16)式右端是企业不进行研发操纵时享有的收益。因此只需要证明当  $\theta \in (\theta^*, \theta^s)$  时,(16)式左端严格大于(16)式右端;当  $\theta \in (0, \theta^*)$  时,(16)式左端严格小于(16)式右端即可。基于此,构造函数:

$$f(\theta) = (1 - t_H) \left[ \pi \bar{q}_1 - \frac{\pi \bar{q}_1}{\alpha} (\bar{r} + c) \right] + (1 - P)(1 - t_L) \rho \pi \left[ \tau \bar{q}_1 + e^{\frac{c}{\alpha}} \theta \ln \left( 1 + \frac{\pi \bar{q}_1 \bar{r}}{\alpha} \right) \right] \\ + P(1 - t_H) \rho \pi \left[ \tau \bar{q}_1 + e^{\frac{c}{\alpha}} \theta \ln \left( 1 + \frac{\pi \bar{q}_1 \bar{r}}{\alpha} \right) \right] - (1 - t_H) [\pi \bar{q}_1 - R^*(\theta)] \\ - (1 - t_H) \rho \pi \left\{ \tau \bar{q}_1 + e^{\frac{c}{\alpha}} \theta \ln [1 + R^*(\theta)] \right\} \quad (17)$$

并证明当  $\theta \in (\theta^*, \theta^s)$  时,  $f(\theta)$  严格大于零;当  $\theta \in (0, \theta^*)$  时,  $f(\theta)$  严格小于零。

对(17)式求导,可得:

$$f'(\theta) = e^{\frac{c}{\alpha}} (1 - P) \rho \pi \left\{ (1 - t_L) \ln \left( 1 + \frac{\pi \bar{q}_1 \bar{r}}{\alpha} \right) - (1 - t_H) \ln [1 + R^*(\theta)] \right\} \\ + e^{\frac{c}{\alpha}} P \rho \pi (1 - t_H) \left\{ \ln \left( 1 + \frac{\pi \bar{q}_1 \bar{r}}{\alpha} \right) - \ln [1 + R^*(\theta)] \right\} \quad (18)$$

当  $\theta \in (0, \theta^s)$  时,  $\frac{\pi \bar{q}_1 \bar{r}}{\alpha} = R^*(\theta^s) > R^*(\theta)$ , 因此当  $\theta \in (0, \theta^s)$  时,  $f'(\theta) > 0$ 。由于  $f(\theta^*) = 0$ , 因此当  $\theta \in (\theta^*, \theta^s)$  时,  $f(\theta) > 0$ ; 当  $\theta \in (0, \theta^*)$  时,  $f(\theta) < 0$ 。因此,研发能力  $\theta \in (\theta^*, \theta^s)$  的企业为了达到高企认定政策规定的门槛条件,将会进行研发操纵。

上述分析解释了高企认定政策如何诱使企业进行研发操纵。政府出台高企认定政策并设置研发投入强度的门槛条件,旨在筛选出研发能力强的企业并激励这些企业进行研发创新。但是,对于研发能力低的企业来说,也可以通过研发操纵的方式来达到研发投入强度的门槛条件(尽管在政府未出台高企认定政策时,其最优研发投入强度低于高企认定政策规定的门槛条件)。同时,政府受制于信息获取成本,很难准确识别出这些企业,因此部分研发操纵企业可以浑水摸鱼,成功获得高企认定并享受相关的政策优惠。在本文的理论模型中,研发能力  $\theta \in (\theta^*, \theta^s)$  的企业会进行研发操纵,使得获得高企认定的企业鱼龙混杂——既包括研发能力较强 ( $\theta \geq \theta^s$ ) 的企业,也包括研发能力较弱的企业 ( $\theta < \theta^s$ )。显然,研发能力较弱的企业的机会主义行为降低了高企认定政策的激励效果。

#### (四)数字政府与企业研发操纵

本文最后分析数字政府对企业研发操纵的影响。根据前文分析,随着数字政府

的建设,政府有更大的概率发现研发操纵企业,即(16)式中参数 $P$ 的值上升。为了便于分析,我们将(16)式中的参数 $P$ 替换为函数 $P(D)$ ,其中 $D$ 代表数字政府建设的程度,并且满足 $P'(D) > 0$ 。于是重新表述(16)式为:

$$\begin{aligned} & (1-t_H)\left[\pi\bar{q}_1 - \frac{\pi\bar{q}_1}{\alpha}(\bar{r}+c)\right] + [1-P(D)](1-t_L)\rho\pi\left[\tau\bar{q}_1 + e^{\frac{\alpha}{2}}\theta^*(D)\ln\left(1 + \frac{\pi\bar{q}_1\bar{r}}{\alpha}\right)\right] \\ & + P(D)(1-t_H)\rho\pi\left[\tau\bar{q}_1 + e^{\frac{\alpha}{2}}\theta^*(D)\ln\left(1 + \frac{\pi\bar{q}_1\bar{r}}{\alpha}\right)\right] \quad (19) \\ & = (1-t_H)\left[\pi\bar{q}_1 - R^*(\theta^*(D))\right] + (1-t_H)\rho\pi\left\{\tau\bar{q}_1 + e^{\frac{\alpha}{2}}\theta^*(D)\ln[1+R^*(\theta^*(D))]\right\} \end{aligned}$$

接下来,我们进行比较静态分析,探讨数字政府建设(参数 $D$ 的值上升)对 $\theta^*$ 的影响。应用隐函数定理,(19)式的左右两端同时对 $D$ 求导:

$$\begin{aligned} & e^{\frac{\alpha}{2}}[1-P(D)](1-t_L)\rho\pi\ln\left(1 + \frac{\pi\bar{q}_1\bar{r}}{\alpha}\right)\frac{d\theta^*(D)}{dD} + e^{\frac{\alpha}{2}}P(D)(1-t_H)\rho\pi\ln\left(1 + \frac{\pi\bar{q}_1\bar{r}}{\alpha}\right)\frac{d\theta^*(D)}{dD} \\ & + (1-t_H)\rho\pi\left[\tau\bar{q}_1 + e^{\frac{\alpha}{2}}\theta^*(D)\ln\left(1 + \frac{\pi\bar{q}_1\bar{r}}{\alpha}\right)\right]\frac{dP(D)}{dD} - (1-t_L)\rho\pi \\ & \left[\tau\bar{q}_1 + e^{\frac{\alpha}{2}}\theta^*(D)\ln\left(1 + \frac{\pi\bar{q}_1\bar{r}}{\alpha}\right)\right]\frac{dP(D)}{dD} = e^{\frac{\alpha}{2}}(1-t_H)\rho\pi\ln[1+R^*(\theta^*(D))]\frac{d\theta^*(D)}{dD} \end{aligned}$$

化简得:

$$\begin{aligned} & \rho\pi[(1-t_L)-(1-t_H)]\left[\tau\bar{q}_1 + e^{\frac{\alpha}{2}}\theta^*(D)\ln\left(1 + \frac{\pi\bar{q}_1\bar{r}}{\alpha}\right)\right]\frac{dP(D)}{dD} \\ & = \left\{e^{\frac{\alpha}{2}}\rho\pi(1-t_H)P(D)\left[\ln\left(1 + \frac{\pi\bar{q}_1\bar{r}}{\alpha}\right) - \ln[1+R^*(\theta^*(D))]\right]\right\} \quad (20) \\ & + e^{\frac{\alpha}{2}}\rho\pi[1-P(D)]\left[(1-t_L)\ln\left(1 + \frac{\pi\bar{q}_1\bar{r}}{\alpha}\right) - (1-t_H)\ln[1+R^*(\theta^*(D))]\right]\frac{d\theta^*(D)}{dD} \end{aligned}$$

我们注意到,在(20)式中,由于 $\frac{dP(D)}{dD} > 0$ ;  $(1-t_L) > (1-t_H)$ ; 对于 $\theta \in (0, \theta^s)$ ,  $\frac{\pi\bar{q}_1\bar{r}}{\alpha} = R^*(\theta^s) > R^*(\theta)$ 。因此得到 $\frac{d\theta^*(D)}{dD} > 0$ 。

根据 $\frac{d\theta^*(D)}{dD} > 0$ ,我们得出结论:数字政府降低了企业进行研发操纵的概率。使用图1可以更清楚地展示上述结论,图1横坐标表示企业研发能力( $\theta$ ),纵坐标表示概率密度,曲线表示企业研发能力( $\theta$ )在实数空间上的概率密度函数。根据理论分析,政府希望激励研发能力强的企业( $\theta \geq \theta^s$ )进行研发创新,因此,政府出台了高企认定政策,并在其中设置了研发投入强度的门槛条件( $r \geq \bar{r}$ )。在数字政府建设程度为 $D_1$ 条

件下,根据(18)和(19)式,对于研发能力属于区间 $[\theta^*(D_1), \theta^s]$ 的企业来说,进行研发操纵是有利可图的。然而,当数字政府的建设程度由 $D_1$ 上升至 $D_2$  ( $D_2 > D_1$ ),对于研发能力属于区间 $[\theta^*(D_2), \theta^s]$ 的企业来说,进行研发操纵是有利可图的。由于 $\frac{d\theta^*(D)}{dD} > 0$ ,因此 $\theta^*(D_2) > \theta^*(D_1)$ ,也就是说,随着数字政府建设,研发能力属于区间 $[\theta^*(D_1), \theta^*(D_2)]$ 企业的最优决策将从进行研发操纵转为不进行研发操纵<sup>①</sup>。

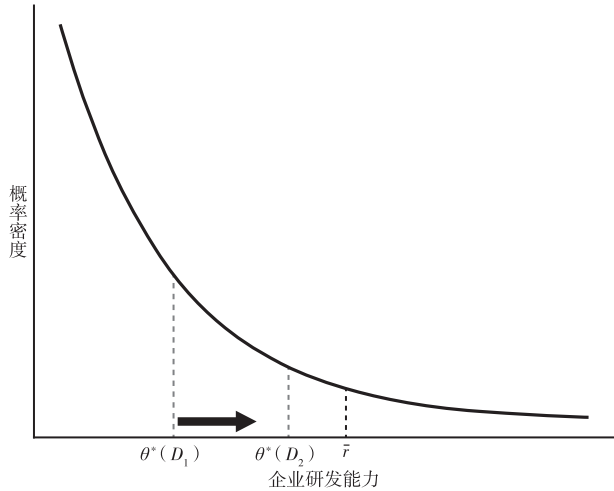


图1 数字政府抑制企业研发操纵

综上,我们提出:

**命题1:** 数字政府建设能够抑制企业研发操纵。

命题1背后的经济学直觉为,企业会权衡研发操纵的期望收益和期望成本,决策是否进行研发操纵以达到高企认定政策的门槛条件。随着数字政府的建设,企业进行研发操纵将会有更大的概率被政府发现,导致无法获得高企认定。因此,在其他条件不变的情况下,由于企业进行研发操纵的期望收益下降,企业进行研发操纵的动机也随之下降。

## 四 研究设计

### (一)样本

本文使用2014–2021年沪深A股上市公司为研究样本,选择2014年作为研究的起始样本,是因为中国自2018年开启了第八轮政府机构改革,省级政府随之进行大数据管理机构改革,这样最早一批接受处理的企业在处理前有4年的观测期,在处理当年及处理之后也有4年的观测期,保证了在政策实施前后有合理的观测长度。本

<sup>①</sup> 在数字政府建设程度为 $D_1$ 的条件下,这部分企业恰好是获得高企认定的企业中研发能力最低的企业。换言之,数字政府建设让高企认定政策更好地发挥作用。

文采用下列方式处理样本:(1)剔除金融类上市公司;(2)剔除资产负债率大于1的上市公司;(3)剔除样本数据缺失的公司;(4)剔除营业收入小于20 000万元的公司<sup>①</sup>;(5)为了控制离群值对回归结果的影响,对主要连续变量在上下1%分位进行缩尾处理。经过上述处理,本文得到观测值25 456个,本文的大数据管理机构的改革时间为手工收集,财务数据来自国泰安数据库(CSMAR),互联网和新闻媒体监督数据来自中国研究数据服务平台数据库(CNRDS)。

## (二)计量模型

为了利用大数据管理机构改革政策冲击识别数字政府对企业研发操纵的影响,我们构造计量模型如下:

$$RDM1_{igt}(RDM2_{igt}) = \alpha_1 + \beta_1 DG_{gt} + \Gamma X_{igt} + \Lambda_{gt} + \kappa_h + \eta_g + \lambda_t + \varepsilon_{igt} \quad (21)$$

其中,下标*i*、*g*、*t*、*h*分别表示企业、省份、年份和行业。

$RDM1_{igt}$ 或 $RDM2_{igt}$ 表示企业研发操纵。本文在理论分析中证明,在政府未出台高企认定政策的条件下,企业的最优研发投入强度与其研发能力呈正相关。因此,为了筛选出研发能力强的企业并激励其进行研发创新,政府出台了高企认定政策,并在其中设置了研发投入强度的门槛条件。然而,当高企认定政策出台后,一些研发能力较弱的企业也有动机通过研发操纵的方式达到高企认定政策规定的门槛条件,也就是说,研发操纵企业的研发投入强度恰好聚集在高企认定政策规定的门槛点处。因此,借鉴杨国超等(2017)、杨国超和芮萌(2020)的研究,将研发投入强度恰好超过法定门槛1%和0.5%的企业定义为研发操纵企业。具体来说,如果企业的研发投入占销售收入的比例属于区间[3%, 4%]或[3%, 3.5%], $RDM1_{igt}$ 或 $RDM2_{igt}$ 取1,否则为0。

$DG_{gt}$ 为双重差分项,用以捕捉数字政府的建设对企业研发操纵的影响,其数值根据各省份大数据管理机构改革的时间确定。具体来说,如果企业*i*所在的省份*g*在第*t*年进行了大数据管理机构改革,则在第*t*年及之后, $DG_{gt}$ 取1,否则为0<sup>②</sup>。

$X_{igt}$ 和 $\Lambda_{gt}$ 分别表示企业层面和地区层面的控制变量向量。首先,借鉴已有文献,控制公司层面的影响因素,具体包括:资产收益率(*roa*)、企业规模(*size*)、资产负债率(*lev*)、企业年龄(*age*)、托宾Q值(*tobin*)、第一大股东持股比例(*top*)、董事长是否兼任总经理(*dual*)、董事会规模(*board*)、独立董事比例(*indpt*)、股权制衡度(*balance*)、是否为“四大”审计(*audit*)、是否为亏损企业(*loss*)。其次,控制可能影响数字政府建设的

<sup>①</sup> 本文的样本中,营业收入小于20 000万元的企业仅有906家,在总体中仅占极小的比例,为了让研发操纵企业更具可比性,本文将这部分企业剔除。

<sup>②</sup> 限于篇幅,各地区大数据管理机构改革情况见网站附录二。



地区层面影响因素,具体包括:经济发展水平( $gdp$ )、财政状况( $fiscal$ )、数字经济发展( $de$ )、产业结构( $tertiary$ )、市场化程度( $market$ )。

$\kappa_h$ 、 $\eta_g$  和  $\lambda_i$  分别控制行业、省份和年度层面固定效应。 $\varepsilon_{igt}$  表示在省份层面做聚类调整的标准误。

限于篇幅,变量具体定义和描述性统计见网站附录三。

## 五 回归结果分析

### (一)企业是否存在研发操纵

为了检验数字政府建设是否可以抑制企业进行研发操纵,首先需要验证一个前提条件:高企认定政策出台后,企业会进行研发操纵以达到政策规定的门槛条件。接下来,本文使用近年来新兴的群聚分析法( $bunching$ )来验证这个前提条件。

群聚分析法的核心思想是找到未受到高企认定政策影响的企业,然后利用这些企业的分布来估计受到高企认定政策影响的企业的反事实分布<sup>①</sup>。我们使用群聚分析法得到操纵区间为 $[0.5\%, 4.3\%]$ ,图2绘制了企业研发投入强度反事实分布的估计

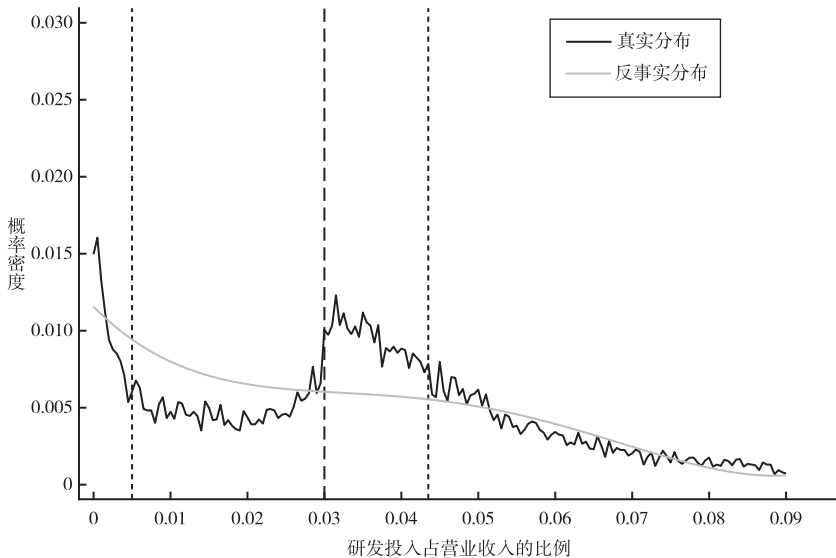


图2 企业研发投入强度的反事实频率分布

<sup>①</sup> 限于篇幅,群聚分析法的技术细节见网站附录四。

结果。可以看出,在高企认定政策规定的门槛点处,企业研发投入强度的分布呈现出明显的左侧样本缺失和右侧样本群聚。具体而言,在区间 $[0.5\%, 3\%]$ 中,真实分布相较于反事实分布缺失的样本(missing mass)约占总样本的10.9%(标准误为3.7%),在区间 $[3\%, 4.3\%]$ 中,真实分布相较于反事实分布多出的样本(bunching mass)约占总样本的9.8%(标准误为1.3%),表明在上市公司中,企业为获得高企认定而进行研发操纵是比较普遍的。

## (二)基准回归

表1列示了基准模型的回归结果。在第(1)列,我们未加入控制变量;在第(2)列,加入(21)式设定的一系列控制变量,回归结果显示,数字政府(DG)的系数显著为负,表明数字政府建设能够抑制企业进行研发操纵。为了看清数字政府建设对企业研发操纵的平均处理效应,在第(3)列,我们使用OLS重新进行估计,结果显示,RDM1的变化程度相当于RDM1样本均值的16.5%(2.8%/17%)<sup>①</sup>,表明从经济意义上讲,数字政府的建设对企业研发操纵行为的抑制作用同样是比较明显的。第(4)-(6)列,我们进一步使用RDM2衡量研发操纵,重新估计模型并得到一致的结论。综上,无论从统计意义还是从经济意义上讲,数字政府建设都会抑制企业研发操纵。

表1 数字政府和企业研发操纵:基准回归

	(1)	(2)	(3)	(4)	(5)	(6)
	Logit 模型	Logit 模型	OLS 模型	Logit 模型	Logit 模型	OLS 模型
	RDM1	RDM1	RDM1	RDM2	RDM2	RDM2
DG	-0.219*** (-3.526)	-0.208*** (-2.958)	-0.028*** (-3.182)	-0.266*** (-3.182)	-0.240*** (-2.810)	-0.024*** (-3.212)
roa		0.864 (1.321)	0.101 (1.305)		0.524 (0.646)	0.040 (0.618)
dual		0.029 (0.777)	0.005 (1.033)		0.042 (0.825)	0.005 (1.043)
board		-0.009 (-0.428)	-0.001 (-0.311)		-0.004 (-0.146)	-0.000 (-0.058)
indpt		-0.058 (-0.094)	0.008 (0.106)		-0.093 (-0.128)	0.001 (0.017)

<sup>①</sup> 虽然可以使用非线性双重差分模型估计处理效应(Wooldridge, 2023),但是Ai and Norton(2003)指出,使用非线性模型估计交互项的系数会出现偏误,因此我们使用OLS模型重新估计。此外,与OLS模型相比,非线性模型使用的样本数量有所下降,这和非线性模型使用的估计方法有关。关于非线性模型的估计问题,见网站附录五。

(续表)

	(1)	(2)	(3)	(4)	(5)	(6)
	Logit 模型 <i>RDM1</i>	Logit 模型 <i>RDM1</i>	OLS 模型 <i>RDM1</i>	Logit 模型 <i>RDM2</i>	Logit 模型 <i>RDM2</i>	OLS 模型 <i>RDM2</i>
<i>top</i>		0.002 (0.725)	0.000 (1.054)		0.003 (1.088)	0.000 (1.465)
<i>lev</i>		0.319* (1.793)	0.033 (1.499)		0.299 (1.631)	0.021 (1.315)
<i>size</i>		-0.158*** (-4.872)	-0.020*** (-4.993)		-0.146*** (-3.678)	-0.013*** (-3.616)
<i>age</i>		-0.004 (-0.900)	-0.000 (-0.844)		-0.003 (-0.402)	-0.000 (-0.309)
<i>loss</i>		-0.287*** (-2.799)	-0.033*** (-3.043)		-0.361** (-2.476)	-0.028** (-2.692)
<i>audit</i>		-0.063 (-0.475)	-0.006 (-0.495)		-0.110 (-0.718)	-0.007 (-0.726)
<i>tobin</i>		-0.147*** (-5.289)	-0.017*** (-5.064)		-0.117*** (-3.478)	-0.009*** (-3.466)
<i>balance</i>		-0.003 (-1.034)	-0.000 (-1.606)		-0.002 (-0.706)	-0.000 (-1.251)
<i>gdp</i>		-0.201 (-0.439)	-0.013 (-0.238)		-0.639 (-1.324)	-0.044 (-1.074)
<i>market</i>		-0.027 (-0.469)	-0.008 (-1.280)		-0.058 (-0.908)	-0.009* (-1.726)
<i>de</i>		-0.004 (-1.108)	-0.001 (-1.373)		-0.002 (-0.481)	-0.000 (-0.677)
<i>fiscal</i>		0.833 (1.300)	0.106 (1.467)		1.880** (2.029)	0.177* (1.950)
<i>tertiary</i>		-0.031* (-1.820)	-0.003* (-1.780)		-0.037** (-2.313)	-0.003** (-2.068)
省份固定效应	控制	控制	控制	控制	控制	控制
行业固定效应	控制	控制	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制	控制	控制
观测值	25 398	25 398	25 456	24 532	24 532	25 456
R <sup>2</sup>	0.115	0.122	0.095	0.088	0.094	0.061

说明:括号内为t值,\*、\*\*、\*\*\*分别表示在10%、5%、1%水平上显著,标准误聚类到省份层面。后表同。

### (三)稳健性检验

#### 1. 平行趋势检验

运用双重差分法识别数字政府对企业研发操纵的因果效应,需要满足平行趋势

假定的前提条件。图3展示了双重差分模型平行趋势假设的检验结果,可以看出,实验组受到大数据管理机构改革政策冲击之前,研发操纵企业占比随时间的变动趋势与控制组并无显著差异,满足平行趋势假定。但是,当受到大数据管理机构改革政策冲击之后,实验组的研发操纵企业占比相较于控制组有明显向下的趋势,并且这种趋势直至政策实施之后的第3年仍然是显著的,表明数字政府建设对企业研发操纵的影响是长期的,进一步佐证了数字政府建设抑制企业研发操纵的作用在经济意义上是比较明显的。

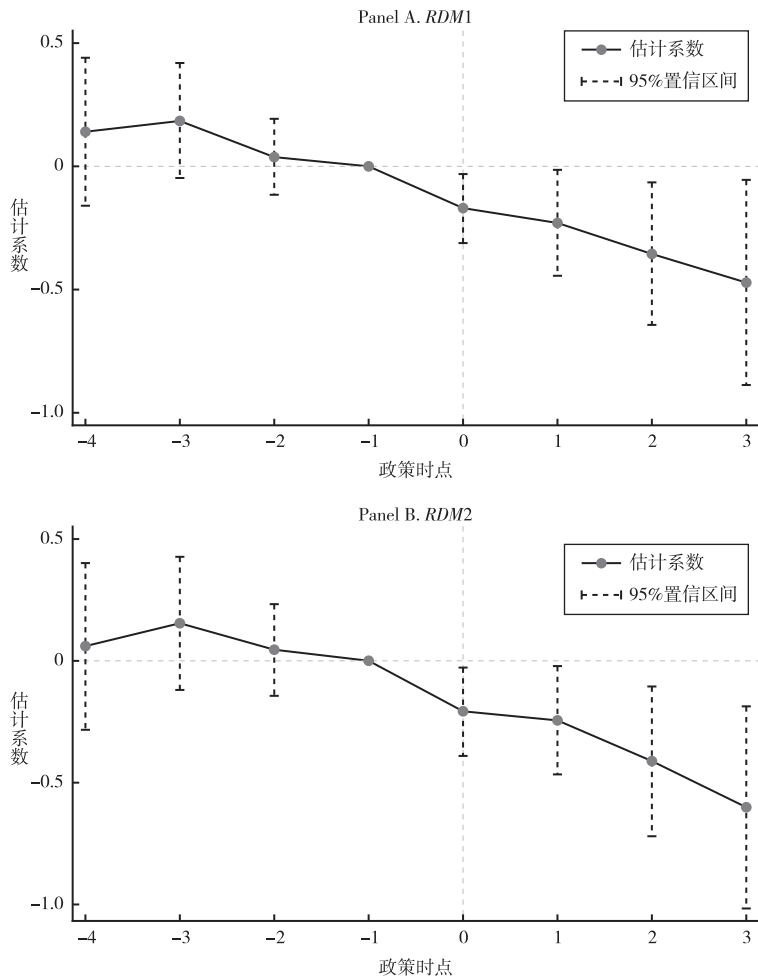


图3 双重差分平行趋势检验

说明:每条虚线的中间点表示参数估计值,虚线两端表示参数估计值的95%置信区间。

## 2. 双重差分异质性处理效应

现有研究发现,交叠型双重差分的估计可能存在偏误,原因在于该方式使用较早接受政策处理的个体估计较晚接受政策处理个体的反事实结果,导致处理效应的估计出现偏误(Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Borusyak *et al.*, 2021)。针对上述问题, *Borusyak et al.* (2021) 提出了插补估计量法(imputation estimation),使用尚未接受政策处理的个体估计接受政策处理个体的反事实结果,然后将接受处理个体的真实结果与反事实结果相减,得到处理效应的正确估计结果。*Wooldridge* (2023)进一步发现在非线性双重差分模型中,插补估计量的偏误更小。我们使用 *Wooldridge* (2023)的方式重新进行检验并发现结论不变(见网站附录六)。

## 3. 三重差分识别策略

三重差分的识别策略可以更精准地识别数字政府建设和企业研发操纵之间的因果关系。考虑到高企认定政策主要针对高科技企业,获得高企认定的企业通常属于高科技行业;相比之下,传统行业的企业则很难获得高企认定,因此它们进行研发操纵的动机相对较弱,受到数字政府建设的影响也相对较小。这给我们提供了识别策略的思路:选择难以获得高企认定的行业中的企业作为新的控制组。借鉴 *Beraja et al.* (2023b),构建三重差分模型:

$$RDM_{igt} = \alpha_2 + \beta_2 DG_{gt} \times HT_{ih} + \chi_2 DG_{gt} + \delta_2 HT_{ih} + \Gamma X_{igt} + \mathbf{I}A_{gt} + \kappa_h + \eta_g + \lambda_t + \varepsilon_{igt} \quad (22)$$

其中  $HT_{ih}$  是第三重差分项,代表企业  $i$  是否属于容易获得高企认定的行业  $h$ 。构造方式如下:我们首先计算全样本企业在所有细分行业中的概率密度( $Freq1$ )<sup>①</sup>;然后计算获得高企认定的企业在所有细分行业中的概率密度( $Freq2$ );再将获得高企认定的企业在所有细分行业中的概率密度除以全样本企业在所有细分行业中的概率密度( $HTRatio = Freq2 / Freq1$ ),得到每个细分行业中的企业获得高企认定的相对难易程度;最后计算  $HTRatio$  的中位数,如果大于中位数,则  $HT$  取 1,否则取 0。

通过三重差分识别策略,我们无需假设各地区大数据管理机构改革是外生的,只需要假设第三重差分项  $HT$  是外生的(*Beraja et al.*, 2023b)。因为本文中的绝大多数样本企业在大数据管理机构改革之前已经成立,所以这个假设更容易满足。表 2 的回归结果显示,  $DG \times HT$  的系数显著为负,我们进一步使用连续变量  $HTRatio$  代替离散

① 我们使用证监会 2012 年修订的《上市公司行业分类指引》确定细分行业,证监会分类指引将上市公司经济活动划分为门类和大类两级,我们使用大类一级确定细分行业。例如, C39 表示计算机、通信和其他电子设备制造业。

变量  $HT$ , 重新进行回归,  $DG \times HTRaio$  的系数仍然显著为负。综上, 三重差分的回归结果进一步支持了数字政府建设和企业研发操纵之间的因果关系。

	三重差分模型			
	(1)	(2)	(3)	(4)
	<i>RDM1</i>	<i>RDM2</i>	<i>RDM1</i>	<i>RDM2</i>
$DG \times HT$	-0.463*** (-2.615)	-0.484*** (-2.863)		
$HT$	1.799*** (5.706)	1.428*** (4.972)		
$DG \times HTRaio$			-0.731*** (-6.487)	-0.741*** (-5.590)
$HTRaio$			1.414*** (5.844)	1.282*** (7.001)
$DG$	0.216 (1.457)	0.199 (1.262)	0.670*** (4.756)	0.637*** (3.751)
企业特征	控制	控制	控制	控制
地区特征	控制	控制	控制	控制
省份固定效应	控制	控制	控制	控制
行业固定效应	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
观测值	25 398	24 532	25 398	24 532
$R^2$	0.126	0.096	0.126	0.098

#### 4. 构造伪研发操纵企业

由于高企认定政策针对研发投入占销售收入的比例设置门槛条件, 并没有对研发投入占企业规模的比例做出规定, 因此企业没有动机操纵研发投入占企业规模的比例这一指标。我们据此构造伪研发操纵变量 ( $RDM\ Fake$ ), 除了使用企业研发投入占企业规模的比例衡量研发投入强度外, 构造方式和研发操纵变量 ( $RDM$ ) 完全一致。可以预期, 在安慰剂检验中, 数字政府建设不应当对伪研发操纵产生显著影响。表3回归结果

	构造伪研发操纵企业	
	(1)	(2)
	<i>RDM Fake1</i>	<i>RDM Fake2</i>
$DG$	-0.123 (-1.141)	0.014 (0.172)
企业特征	控制	控制
地区特征	控制	控制
省份固定效应	控制	控制
行业固定效应	控制	控制
年份固定效应	控制	控制
观测值	24 988	23 312
$R^2$	0.118	0.081



显示,数字政府建设对伪研发操纵的影响不显著,说明并非随机性因素造成基准回归的负向关系。

### 5. 数字政府安慰剂检验

在第八轮政府机构改革之前,广东、浙江和贵州等地虽然成立了大数据管理机构,但是这些机构尚未具备统筹管理部门之间的数据和数字政府建设等职能。在第八轮政府机构改革之后,大数据管理机构承担起统筹管理部门间的数据和数字政府建设等职能。根据理论分析,这一改革有助于解决数字政府建设过程中存在的数据孤岛和重复建设等问题,让数字政府更好地发挥治理企业研发操纵的作用。我们基于这一观察构造安慰剂检验:如果理论分析成立,那么在第八轮政府机构改革之前成立的大数据管理机构不会对企业研发操纵起到治理作用。

在安慰剂检验中,我们将样本期间限定在2014–2017年,并构造了数字政府安慰剂变量(*DG Placebo*)。数字政府安慰剂变量的构造方式和数字政府变量(*DG*)完全一致,唯一的区别是将广东、浙江和贵州三个省份的大数据管理机构改革时间替换为最早成立大数据管理机构的时间<sup>①</sup>。观察表4的回归结果我们发现,*DG Placebo*的系数不显著,暗示只有当大数据管理机构具备统筹管理部门间的数据和统筹数字政府建设等职能后,才能发挥抑制企业研发操纵的作用。

表4 数字政府安慰剂检验

	(1)	(2)
	<i>RDM1</i>	<i>RDM2</i>
<i>DG Placebo</i>	0.038 (0.424)	-0.001 (-0.006)
企业特征	控制	控制
地区特征	控制	控制
省份固定效应	控制	控制
行业固定效应	控制	控制
年份固定效应	控制	控制
观测值	10 052	10 052
R <sup>2</sup>	0.128	0.109

### 6. 其他稳健性检验

本文进行了其他稳健性检验,结论一致(相关检验详见网站附录七)。这些稳健性检验包括:(1)借鉴Li *et al.*(2016)和孙天阳等(2022)平衡性检验的方法,验证大数据管理机构改革的条件随机性。(2)排除金税三期、宽带中国、研发费用加计扣除新政等同期政策干扰。(3)借鉴La-Ferrara *et al.*(2012)随机置换的方法进行安慰剂检验。(4)使用数字政府相关政策出台数量作为数字政府的代理变量。(5)考虑双重差分无预期效应假设。(6)考虑企业操纵研发人员数量。(7)使用群聚估计得到的操纵区间定

① 广东省替换为2014年,浙江省和贵州省替换为2015年。

义研发操纵企业。(8)保留营业收入小于20 000万元的样本。(9)根据稳健性检验5,在第八轮政府机构改革之前,广东、浙江和贵州已经成立大数据管理机构。我们担心这一因素可能对本文回归结果造成混淆影响,因此删除企业所在省份为广东、浙江和贵州的样本。

## 六 影响机制分析

### (一)政府更好地发挥数字技术作用

本文的理论分析认为,大数据管理机构改革让政府能够更有效地发挥大数据等数字技术对企业研发操纵的治理作用。构造下列计量模型进行检验:

$$RDM_{igt} = \alpha_3 + \beta_3 BigData_{gt} + \Gamma X_{igt} + \mathbf{I}A_{gt} + \kappa_h + \eta_g + \lambda_t + \varepsilon_{igt} \quad (23)$$

$$RDM_{igt} = \alpha_4 + \beta_4 DG_{gt} \times BigData_{gt} + \chi_4 DG_{gt} + \delta_4 BigData_{gt} + \Gamma X_{igt} + \mathbf{I}A_{gt} + \kappa_h + \eta_g + \lambda_t + \varepsilon_{igt} \quad (24)$$

其中, $BigData_{gt}$ 表示大数据等数字技术在政府治理中的应用。政府工作报告的工作回顾部分总结了政府在过去一年取得的主要成绩,因此,我们计算政府数字化治理相关的词汇在工作回顾部分中出现的次数来进行量化分析<sup>①</sup>。如果随着数字政府的建设,政府可以更有效地发挥大数据等数字技术对企业研发操纵的治理作用,我们将观测到(24)式中交互项 $BigData_{gt} \times DG_{gt}$ 的系数 $\beta_4$ 显著为负。

回归结果见表5。在第(1)列中,我们对(23)式进行回归,并发现 $BigData$ 的系数显著为负,表明政府应用大数据等数字技术可以发挥对企业研发操纵的治理作用,在第(2)列,我们对(24)式进行回归,主要关注交互项 $BigData \times DG$ 的系数。我们发现交互项的系数显著为负,这表明随着数字政府的建设,政府可以更好地发挥大数据等数字技术的治理作用。由于非线性模型交互项的估计系数可能会出现偏误(Ai and Norton, 2003),我们使用OLS模型重新估计,回归结果见第(3)列<sup>②</sup>。在第(4)–(6)列,我们使用 $RDM2$ 衡量企业研发操纵,得到一致的结论。综上,回归结果表明,

① 我们基于国务院颁布的《关于加强数字政府建设的指导意见》,以及对政府工作报告的阅读和分析,我们筛选出以下关键词:数字监管、智慧监管、智慧治理、在线监管、智能感知、一网统管、互联网监管、新型监管、数字化监管、大数据监管、精准识别监管、智能监管、数字化治理、智能监管、精准监管、一网通管、精细化治理、智能化管理、双随机一公开监管、政务大数据、政务云等。

② Ai and Norton(2003)给出计算非线性模型交互项的边际效应的方法,我们使用该方法计算非线性模型交互项的边际效应,发现研究结果未发生改变,交互项的边际效应仍然显著为负。检验结果见网站附录八。

随着数字政府的建设,政府可以更好地发挥大数据等数字技术对企业研发操纵的治理作用。

表 5 政府更好发挥数字技术作用

	(1)	(2)	(3)	(4)	(5)	(6)
	Logit 模型	Logit 模型	OLS 模型	Logit 模型	Logit 模型	OLS 模型
	<i>RDM1</i>	<i>RDM1</i>	<i>RDM1</i>	<i>RDM2</i>	<i>RDM2</i>	<i>RDM2</i>
<i>DG</i> × <i>BigData</i>		-0.208** (-2.036)	-0.025* (-2.019)		-0.226** (-2.481)	-0.019** (-2.372)
<i>DG</i>		-0.211*** (-2.794)	-0.029*** (-3.026)		-0.238*** (-2.611)	-0.024*** (-3.041)
<i>BigData</i>	-0.065*** (-3.656)	0.128 (1.263)	0.015 (1.183)	-0.048** (-2.177)	0.160* (1.800)	0.013 (1.635)
企业特征	控制	控制	控制	控制	控制	控制
地区特征	控制	控制	控制	控制	控制	控制
省份固定效应	控制	控制	控制	控制	控制	控制
行业固定效应	控制	控制	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制	控制	控制
观测值	25 398	25 398	25 456	24 532	24 532	25 456
R <sup>2</sup>	0.122	0.122	0.095	0.094	0.094	0.061

## (二)研发操纵企业更难获取高企认定

由前文理论分析可知,数字政府建设提升了政府对企业研发操纵的监管能力,增大了政府发现企业研发操纵行为的可能性,导致企业更难通过研发操纵的方式获得高企认定。接下来,我们检验随着数字政府的建设,企业更难通过研发操纵的方式获得高企认定。构建下列计量模型进行检验:

$$HTQ_{igt} = \alpha_5 + \beta_5 DG_{gt} \times RDM_{igt} + \chi_5 DG_{gt} + \delta_5 RDM_{igt} + \Gamma X_{igt} + \mathbf{1}A_{gt} + \kappa_h + \eta_g + \lambda_t + \varepsilon_{igt} \quad (25)$$

其中 $HTQ_{igt}$ 是虚拟变量,表示企业是否获得高企认定,如果企业获得高企认定,则 $HTQ_{igt}$ 取1,反之取0。其余变量的设定和前文一致。如果理论分析成立,我们将观测到交互项 $DG_{gt} \times RDM_{igt}$ 的系数 $\beta_5$ 显著为负,表明随着数字政府的建设,企业更难通过研发操纵的方式获得高企认定。

以表6第(1)列为例, $RDM1$ 的系数显著为正,表明在未受到大数据管理机构改

革政策冲击的情况下,企业可以通过研发操纵的方式获得高企认定。我们主要关注交互项 $DG \times RDM1$ 的系数,我们发现,交互项的系数显著为负,表明当受到大数据管理机构改革政策冲击之后,企业更难通过研发操纵的方式获得高企认定。

表6 研发操纵企业更难获取高企认定

	(1)	(2)	(3)	(4)
	Logit 估计 <i>HTQ</i>	Logit 估计 <i>HTQ</i>	OLS 估计 <i>HTQ</i>	OLS 估计 <i>HTQ</i>
<i>DG \times RDM1</i>	-0.120** (-2.203)		-0.045*** (-5.001)	
<i>RDM1</i>	0.286*** (6.984)		0.058*** (8.376)	
<i>DG \times RDM2</i>		-0.176** (-2.120)		-0.051*** (-3.889)
<i>RDM2</i>		0.333*** (6.410)		0.063*** (7.715)
<i>DG</i>	-0.132 (-1.170)	-0.135 (-1.185)	-0.015 (-1.240)	-0.017 (-1.443)
企业特征	控制	控制	控制	控制
地区特征	控制	控制	控制	控制
省份固定效应	控制	控制	控制	控制
行业固定效应	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
观测值	25 315	25 315	25 456	25 456
$R^2$	0.227	0.227	0.179	0.179

### (三)研发操纵企业更少的虚假标记行为

Chen *et al.* (2021)发现,为了达到高企认定政策的门槛条件,企业可能会将管理费用中不属于研发费用的部分虚假标记为研发费用,以此虚增研发费用<sup>①</sup>。如果数字政府建设能够增加政府发现并处罚企业虚假标记的概率,我们预期随着数字政府的建设,研发操纵企业会更少地将管理费用虚假标记为研发费用。基于此,构建下列计量方程进行检验:

$$Admin_{igt} = \alpha_6 + \beta_6 RDM_{igt} + \Gamma X_{igt} + \mathbf{I}A_{gt} + \kappa_h + \eta_g + \lambda_t + \varepsilon_{igt} \quad (26)$$

$$Admin_{igt} = \alpha_7 + \beta_7 DG_{gt} \times RDM_{igt} + \chi_7 DG_{gt} + \delta_7 RDM_{igt} + \Gamma X_{igt} + \mathbf{I}A_{gt} + \kappa_h + \eta_g + \lambda_t + \varepsilon_{igt} \quad (27)$$

① 中国新会计准则规定,研究阶段发生的研发费用,应当在发生时计入管理费用。

其中  $Admin_{igt}$  表示管理费用减研发费用占销售收入的比例 (Chen *et al.*, 2021)。如果研发操纵企业将管理费用虚假标记为研发费用,我们将观测到(26)式中  $RDM_{igt}$  的系数  $\beta_6$  显著为负。如果随着数字政府的建设,研发操纵企业更少地将管理费用虚假标记为研发费用,我们将观测到(27)式中交互项  $DG_{igt} \times RDM_{igt}$  的系数  $\beta_7$  显著为正。

表7第(1)列为(26)式的回归结果。我们发现  $RDM1$  的系数显著为负,表明研发操纵企业管理费用减研发费用后占销售收入的比例明显低于非研发操纵企业,暗示研发操纵企业有可能将管理费用虚假标记为研发费用,以满足高企认定政策的门槛条件。第(2)列为(27)式的回归结果,交互项  $DG \times RDM1$  的系数显著为正,表明随着数字政府的建设,研发操纵企业更少地将管理费用虚假标记为研发费用。更重要的是,我们不能拒绝  $\beta_7 + \chi_7 = 0$  的原假设,表明当大数据管理机构改革后,研发操纵企业不再进行虚假标记。在第(3)-(4)列,我们使用  $RDM2$  衡量研发操纵,得到一致的结论。回归结果表明,随着数字政府的建设,研发操纵企业更少地将管理费用虚假标记为研发费用。

表7 研发操纵企业更少的虚假标记行为

	(1)	(2)	(3)	(4)
	<i>Admin</i>	<i>Admin</i>	<i>Admin</i>	<i>Admin</i>
<i>RDM1</i>	-0.007*** (-4.102)	-0.011*** (-3.906)		
<i>DG</i> × <i>RDM1</i>		0.009** (2.040)		
<i>RDM2</i>			-0.008*** (-3.690)	-0.012*** (-3.168)
<i>DG</i> × <i>RDM2</i>				0.009* (1.906)
<i>DG</i>		-0.011*** (-3.089)		-0.011*** (-3.019)
企业特征	控制	控制	控制	控制
地区特征	控制	控制	控制	控制
省份固定效应	控制	控制	控制	控制
行业固定效应	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
观测值	22 300	22 300	22 300	22 300
R <sup>2</sup>	0.426	0.427	0.426	0.427
$H_0: \beta_7 + \chi_7 = 0$		0.579		0.642

## 七 进一步研究

本文从产权、互联网监督和新闻媒体监督三个角度讨论数字政府对企业研发操纵的异质性影响。选择这三个角度是基于以下考虑:如果数字政府建设能够提升政府对企业研发操纵的监管能力,那么在研发操纵动机更强的企业中,数字政府将会更好地发挥作用。

### (一)产权异质性

相较于非国有企业,国有企业有更弱的动机通过研发操纵谋取税收优惠(杨国超等,2017),原因在于上缴税收是国企高管考核的一个重要指标。因此,如果数字政府建设提升了政府对企业研发操纵的监管能力,那么对非国有企业将产生更大影响。与上述分析一致,表8回归结果表明,数字政府建设对非国有企业研发操纵的抑制作用更加显著。

	(1)	(2)	(3)	(4)
	国企	非国企	国企	非国企
	<i>RDM1</i>	<i>RDM1</i>	<i>RDM2</i>	<i>RDM2</i>
<i>DG</i>	0.004 (0.025)	-0.244*** (-3.141)	0.092 (0.438)	-0.329*** (-3.297)
企业特征	控制	控制	控制	控制
地区特征	控制	控制	控制	控制
省份固定效应	控制	控制	控制	控制
行业固定效应	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
观测值	7676	16 719	7112	16 336
$R^2$	0.170	0.116	0.135	0.095
组间系数差异	0.057		0.023	

### (二)互联网监督

在数字经济时代,社会公众通过互联网对公司进行监督逐渐成为一种新的公司



治理机制(陈德球和胡晴,2022)。如果一家公司受到更强的互联网监督,一方面公司的研发操纵行为更有可能被曝光;另一方面当公司的研发操纵行为被曝光后,迫于舆论压力,政府也会实施更严厉的处罚措施。因此,受到更强互联网监督的公司,其研发操纵的动机相对较弱。基于这一观点,如果数字政府建设提升了政府对企业研发操纵的监管能力,那么对受到互联网监督较弱的企业将产生更大影响。

本文借鉴杨国超等(2021)的研究,使用上市公司的网络搜索指数来衡量互联网舆论监督。根据指数是否大于年度中位数,将样本划分为互联网监督强和互联网监督弱两组。回归结果见表9。与前述分析一致,研究发现,在互联网监督弱组中,数字政府对企业研发操纵的抑制作用更加显著。

表9 进一步研究:互联网监督

	(1)	(2)	(3)	(4)
	互联网监督弱	互联网监督强	互联网监督弱	互联网监督强
	<i>RDM1</i>	<i>RDM1</i>	<i>RDM2</i>	<i>RDM2</i>
<i>DG</i>	-0.372*** (-3.816)	-0.024 (-0.180)	-0.404*** (-3.357)	-0.071 (-0.489)
企业特征	控制	控制	控制	控制
地区特征	控制	控制	控制	控制
省份固定效应	控制	控制	控制	控制
行业固定效应	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
观测值	12 666	12 651	12 341	11 961
R <sup>2</sup>	0.120	0.136	0.098	0.100
组间系数差异		0.017		0.047

### (三)新闻媒体监督

新闻媒体通过挖掘和解读上市公司的负面信息,增大了企业研发操纵行为被曝光的可能性,从而起到抑制企业研发操纵的作用(杨国超和张李娜,2021)。因此,受到更强新闻媒体监督的公司,其研发操纵的动机相对较弱。基于这一观点,如果数字政府建设提升了政府对企业研发操纵的监管能力,那么对受到新闻媒体监督较弱的企业将产生更大影响。

借鉴古朴和翟士运(2020)的研究,使用公司的负面新闻报道数量衡量新闻媒体监督,并根据是否大于年度中位数,将样本划分为新闻媒体监督强组和新闻媒体监督弱两组。由表10所示的回归结果可见,与前述分析一致,在新闻媒体监督弱组中,数字政府对企业研发操纵的抑制作用更加显著。

表 10 进一步研究:新闻媒体监督

	(1)	(2)	(3)	(4)
	新闻媒体监督弱	新闻媒体监督强	新闻媒体监督弱	新闻媒体监督强
	<i>RDM1</i>	<i>RDM1</i>	<i>RDM2</i>	<i>RDM2</i>
<i>DG</i>	-0.242*** (-2.597)	-0.133 (-0.977)	-0.296** (-2.331)	-0.128 (-1.016)
企业特征	控制	控制	控制	控制
地区特征	控制	控制	控制	控制
省份固定效应	控制	控制	控制	控制
行业固定效应	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
观测值	13 804	11 578	13 445	11 030
$R^2$	0.104	0.155	0.085	0.123
组间系数差异	0.111		0.113	

## 八 研究结论

党的二十大报告提出要“建设数字中国”。本文探讨数字政府对企业研发操纵的影响。研究发现,数字政府建设显著抑制了中国上市企业中的研发操纵现象,在进行一系列稳健性检验之后该结论仍然成立。对数字政府建设抑制企业研发操纵的内在机制探讨发现:数字政府建设有效地提升了政府对企业研发操纵行为的监管能力。最后,本文发现在非国有企业、互联网监督和新闻媒体监督难以发挥作用的企业中,数字政府抑制企业研发操纵的作用更强。

本文研究具有重要的政策启示。首先,本文发现,数字政府有效地抑制了企业研发操纵行为,从而更好地发挥高企认定政策的效果。因此,政府应当加快建设数

字政府,在政府治理中广泛应用大数据等新一代数字技术;同时,政府应当相应转变治理理念,从经验驱动的治理模式转向数据驱动的治理模式,实现“用数据说话、用数据管理、用数据决策、用数据创新”。其次,企业进行研发操纵的原因在于高企认定政策在实施过程中存在政府失灵。本文发现,建设数字政府使政府更有可能发现进行研发操纵的企业,有效缓解了政府治理失灵问题。因此,随着大数据等数字技术广泛应用于政府治理,政府的治理理念和治理模式发生了深远变革,政府可以在市场失灵的领域更加充分地发挥有形之手的作用,实现有效市场和有为政府的更好结合。再次,数字政府建设过程中存在数据孤岛等问题,极大降低了数字政府发挥的作用。本文发现,大数据管理机构改革帮助解决数据孤岛等问题后,政府可以更有效地发挥大数据等数字技术对企业研发操纵的治理作用。因此,政府在应用大数据等数字技术时,应同时解决数据孤岛等问题,积极推动跨部门、跨地区、跨层级的数据充分共享和利用。最后,在高企认定政策的执行过程中,传统监管模式下政府受到监管成本的制约,在实际工作中可能根据企业的研发投入强度是否落在操纵区间来识别研发操纵企业。尽管这种方式在一定程度上符合成本效益原则,但也存在加大误判风险的问题,对初创企业尤为不利。在大数据时代,政府可以通过利用大数据等数字技术,更全面、清晰地洞察企业行为。因此,在对高新技术企业进行监管时,政府应当充分借助数字技术,尽可能地做到具体企业具体分析,减少对“真高新技术企业”的误判风险。

### 参考文献:

- 北京大学课题组、黄璜(2020):《平台驱动的数字政府:能力、转型与现代化》,《电子政务》第7期。
- 陈德球、胡晴(2022):《数字经济时代下的公司治理研究:范式创新与实践前沿》,《管理世界》第6期。
- 戴小勇、成力为(2019):《产业政策如何更有效:中国制造业生产率与加成率的证据》,《世界经济》第3期。
- 樊勇、李昊楠、管淳(2020):《小微企业所得税优惠间断点是否存在聚束效应》,《世界经济》第3期。
- 范子英、程可为、冯晨(2022):《用地价格管制与企业研发创新:来自群聚识别的证据》,《管理世界》第8期。
- 高奇琦(2023):《国家数字能力:数字革命中的国家治理能力建设》,《中国社会科学》第1期。
- 高小平(2023):《中国式现代化公共管理创新的重大探索——对2023年机构改革的理论分析》,《学海》第5期。
- 古朴、翟士运(2020):《监管不确定性与企业盈余质量——基于证监会换届的准自然实验》,《管理世界》第12期。
- 黄璜(2020):《数字政府:政策、特征与概念》,《治理研究》第3期。
- 黄璜、孙学智(2018):《中国地方政府数据治理机构的初步研究:现状与模式》,《中国行政管理》第12期。
- 黄璜、谢思娴、姚清晨(2022):《数字化赋能治理协同:数字政府建设的“下一步行动”》,《电子政务》第4期。

- 黄其松(2022):《数字时代的国家理论》,《中国社会科学》第10期。
- 江小涓(2018):《大数据时代的政府管理与服务:提升能力及应对挑战》,《中国行政管理》第9期。
- 孟庆国、林彤、乔元波、王理达(2020):《中国地方政府大数据管理机构建设与演变——基于第八次机构改革的对比分析》,《电子政务》第10期。
- 孟天广(2022):《数字治理生态:数字政府的理论迭代与模型演化》,《政治学研究》第5期。
- 米加宁、彭康珺、章昌平(2020):《大数据能驱动地方政府机构改革吗?》,《电子政务》第1期。
- 孙天阳、陆毅、成丽红(2022):《港口管理“放管服”改革与出口结构升级》,《世界经济》第3期。
- 唐松、伍旭川、祝佳(2020):《数字金融与企业技术创新——结构特征、机制识别与金融监管下的效应差异》,《管理世界》第5期。
- 唐要家、王钰、唐春晖(2022):《数字经济、市场结构与创新绩效》,《中国工业经济》第10期。
- 姚期智、郁昱(2022):《奋力攻坚关键核心技术 加快构建数据基础制度》,《光明日报》12月23日第10版。
- 王伟玲(2021):《我国数字政府顶层设计的理念辨析与实践指向》,《行政管理改革》第6期。
- 王伟玲(2022):《中国数字政府形态演进和发展瓶颈》,《行政管理改革》第5期。
- 杨国超、刘静、廉鹏、芮萌(2017):《减税激励、研发操纵与研发绩效》,《经济研究》第8期。
- 杨国超、芮萌(2020):《高新技术企业税收减免政策的激励效应与迎合效应》,《经济研究》第9期。
- 杨国超、张李娜(2021):《产业政策何以更有效?——基于海量媒体报道数据与研发操纵现象的证据》,《经济学(季刊)》第6期。
- 杨瑞龙、侯方宇(2019):《产业政策的有效性边界——基于不完全契约的视角》,《管理世界》第10期。
- 叶战备(2021):《政务数据治理的现实推进及其协同逻辑——以N市为例》,《中国行政管理》第6期。
- 张克(2019):《省级大数据局的机构设置与职能配置:基于新一轮机构改革的实证分析》,《电子政务》第6期。
- 张克中、欧阳洁、李文健(2020):《缘何“减税难降负”:信息技术、征税能力与企业逃税》,《经济研究》第3期。
- Acemoglu, D.; Kacigit, U. and Celik, M. A. “Radical and Incremental Innovation: The Roles of Firms, Managers, and Innovator.” *American Economic Journal: Macroeconomics*, 2022, 14(3), pp.199–249.
- Acemoglu, D. and Restrepo, P. “The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment.” *The American Economic Review*, 2018, 108(6), pp.1488–1542.
- Aghion, P.; Jones, B. F. and Jones, C. I. “Artificial Intelligence and Economic Growth,” in A. Agrawal; J. Gans and A. Goldfarb ed., *The Economics of Artificial Intelligence: An Agenda*. Chicago: University of Chicago Press, 2019.
- Ai, C. and Norton, E. C. “Interaction Terms in Logit and Probit Models.” *Economics Letters*, 2003, 80(1), pp.123–129.
- Kacigit, U.; Grigsby, J.; Nicholas, T. and Stantcheva, S. “Taxation and Innovation in the Twentieth Century.” *The Quarterly Journal of Economics*, 2022, 137(1), pp.329–385.
- Arrow, K. “Economic Welfare and the Allocation of Resources for Invention,” in R. Nelson eds., *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton: Princeton University Press, 1962.
- Beraja, M.; Kao, A.; Yang, D. Y. and Yuchtman, N. “AI-tocracy.” *The Quarterly Journal of Economics*, 2023a,

138(3), pp.1349–1402.

Beraja, M.; Yang, D. Y. and Yuchtman, N. “Data-Intensive Innovation and the State: Evidence from AI Firms in China.” *Review of Economic Studies*, 2023b, 90(4), pp.1701–1723.

Bi, K.; Xie, L.; Zhang, H.; Chen, X.; Gu, X. and Tian, Q. “Accurate Medium-Range Global Weather Forecasting with 3D Neural Networks.” *Nature*, 2023, 619(7090), pp.533–538.

Bloom, N.; Griffith, R. and Van Reenen, J. “Do R&D Tax Credits Work? Evidence from a Panel of Countries 1979–1997.” *Journal of Public Economics*, 2002, 85(1), pp.1–31.

Bloom, N.; Schankerman, M. and Van Reenen, J. “Identifying Technology Spillovers and Product Market Rivalry.” *Econometrica*, 2013, 81(4), pp.1347–1393.

Bloom, N.; Van Reenen, J. and Williams, H. “A Toolkit of Policies to Promote Innovation.” *Journal of Economic Perspectives*, 2019, 33(3), pp.163–184.

Borusyak, K.; Jaravel, X. and Spiess, J. “Revisiting Event Study Designs: Robust and Efficient Estimation.” SSRN working paper No. 2826228, 2021.

Callaway, B. and Sant’Anna, P. H. “Difference-in-Differences with Multiple Time Periods.” *Journal of Econometrics*, 2021, 225(2), pp.200–230.

Cengiz, D.; Dube, A.; Lindner, A. and Zipperer, B. “The Effect of Minimum Wages on Low-Wage Jobs.” *The Quarterly Journal of Economics*, 2019, 134(3), pp.1405–1454.

Chen, Z.; Liu, Z.; Suárez Serrato, J. C. and Xu, D. Y. “Notching R&D Investment with Corporate Income Tax Cuts in China.” *The American Economic Review*, 2021, 111(7), pp.2065–2100.

Cong, L. W.; Xie, D. and Zhang, L. “Knowledge Accumulation, Privacy, and Growth in a Data Economy.” *Management Science*, 2021, 67(10), pp.6480–6492.

Goodman-Bacon, A. “Difference-in-Differences with Variation in Treatment Timing.” *Journal of Econometrics*, 2021, 225(2), pp.254–277.

Jones, C. I. and Tonetti, C. “Nonrivalry and the Economics of Data.” *The American Economic Review*, 2020, 110(9), pp.2819–2858.

Jones, C. I. and Williams, J. C. “Measuring the Social Return to R&D.” *The Quarterly Journal of Economics*, 1998, 113(4), pp.1119–1135.

La Ferrara, E.; Chong, A. and Duryea, S. “Soap Operas and Fertility: Evidence from Brazil.” *American Economic Journal: Applied Economics*, 2012, 4(4), pp.1–31.

LeCun, Y.; Bengio, Y. and Hinton, G. “Deep Learning.” *Nature*, 2015, 521(7553), pp.436–444.

Li, P.; Lu Y. and Wang, J. “Does Flattening Government Improve Economic Performance? Evidence from China.” *Journal of Development Economics*, 2016, 123, pp.18–37.

Mastrobuoni, G. “Crime is Terribly Revealing: Information Technology and Police Productivity.” *The Review of Economic Studies*, 2020, 87(6), pp.2727–2753.

Romer, P. M. “Endogenous Technological Change.” *Journal of Political Economy*, 1990, 98(5, Part 2), pp.S71–S102.

---

Wooldridge, J. M. "Simple Approaches to Nonlinear Difference-in-Differences with Panel Data." *The Econometrics Journal*, 2023, 26(3), pp.C31–C66.

## Government Governance in the Big Data Era: Digital Government and Corporate R&D Manipulation

Meng Yuan; Yang Rong

**Abstract:** Using a sample of A-share listed companies between 2014 and 2021 and based on the specific industrial policy of "Management Measures for the Recognition of High-tech Enterprises", this paper employs the bunching method to examine whether companies have obtained high-tech enterprise status through R&D manipulation. Drawing on this, and using the reform of big data management institutions in various regions as a quasi-natural experiment, the paper employs the difference-in-differences (DID) model to explore the impact that digital government exerts on corporate R&D manipulation. The research reveals that the phenomenon of R&D manipulation is often prevalent among China's listed companies. However, the development of digital government has significantly suppressed corporate R&D manipulation activities. The key reason for all this lies in the fact that digital government development has been effectively enhancing the government's capacity to oversee corporate R&D manipulation. Furthermore, the influence of digital government on corporate R&D manipulation is particularly prominent in samples of non-state-owned enterprises and in some cases where internet and media supervision is weak. This paper offers theoretical and empirical evidence for the government to leverage industrial policies more effectively in the new digital era.

**Key words:** digital government, R&D manipulation, industrial policy, big data management institution reform

**JEL codes:** D21, L53, O32

(截稿:2023年12月 责任编辑:宋志刚)