

预算管理与减税效能提升*

——来自小微企业动态群聚的证据

刘 蓉 陈凌庆 王睿霆 邓兴华

内容提要:现代预算管理的要义在于通过预算约束,规范政府收支行为,提升政府治理效能。大规模减税降费政策推行以来,多次出现实际减税额超过预计减税额的现象。对此,本文采用群聚方法,研究门槛性减税政策引发企业动态群聚响应所导致的预算偏差问题,并提出了纳入群聚因素的预算精准化治理方案。本文发现:(1)减半征收政策实施过程中小微企业进行群聚响应,导致税基变化并产生预算偏差。(2)随着减半征收政策四次扩围,新门槛处的小微企业也呈现出群聚现象,而原减税门槛处的企业发生群聚反弹行为,导致税基再次变化。(3)综合考虑企业群聚响应和群聚反弹等因素对税基分布的影响,可以实现减税政策预算精准度达78%以上。(4)本文进一步讨论了不同征收方式、行业类型、知悉程度、涉税服务等方面的企业群聚响应行为。最后,本文提出优化减税政策预算管理是提升减税政策效能的重要路径。

关键词:预算管理 减税政策 小微企业 群聚

一、引言

党的二十大做出了“健全现代预算制度”的战略部署,对现代预算管理制度建设提出方向和思路。大规模减税降费政策作为实施逆周期调节的积极财政政策,在2012—2022年间已经累计实现减税降费和退税缓税缓费超过12万亿元,该政策在发挥“稳就业、稳金融、稳外贸、稳外资、稳预期”作用的同时,也对政府预算管理带来挑战。据此,2023年政府工作报告提出“积极的财政政策要加力提效”,这意味着减税降费政策也亟需提升效能。已有研究认为,通过评估预算编制的科学性、执行的精准性、预算是否超支等角度,可以优化绩效管理水平和提升政策效能(樊丽明等,2022)。对减税政策实施预算管理,旨在明确减税政策实施金额和结构,提升预算有效性和精准性,这不仅体现了公共财政对现代化预算体制的根本要求(白重恩和毛捷,2011),而且是落实减税政策,进行绩效管理的制度保障(高培勇和汪德华,2016)。然而,我国政府预算管理制度的基础相对薄弱,长期存在预算超收、超支的问题(高培勇,2023)。表1展示了2016—2021年我国大规模减税降费政策执行过程中,持续出现实际减税额高于预计减税额的现象,这不仅干扰了预算管理的科学性,也不利于大规模减税降费政策的精准施行。因此,基于现代化预算管理体系,精准预算减税政策实施规模,对提升减税政策效能具有重大的理论与现实意义。

* 刘蓉、陈凌庆(通讯作者)、王睿霆,西南财经大学财政税务学院,邮政编码:611130,电子信箱:liurong@swufe.edu.cn, chenling@foxmail.com, wangrt@swufe.edu.cn; 邓兴华,西南财经大学国际商学院,邮政编码:611130,电子信箱:dengxinghua@swufe.edu.cn。本研究是国家社会科学基金重大项目(20&ZD078)、教育部人文社会科学基金项目(20YJA790046)的阶段性成果。作者感谢第八届中国财政学论坛、第三届现代财税金融与全球治理高峰论坛与会学者的宝贵建议,感谢匿名审稿专家的建设性意见。当然,文责自负。

表 1 2016—2021 年全国预计新增减税额度与实际新增减税额度对比 单位:千亿元

年份	预计减税	实际减税	超额减税差额	减税预算偏离程度
	(1)	(2)	(3)	(4)
2016	5	5.70	0.70	14.00%
2017	8.50	10	1.50	17.60%
2018	11	13	2	18.19%
2019	20	23	3	15.00%
2020	25	26	1	4.00%
2021	7	10	3	42.86%

注:数据来自国家统计局官网、中国政府网与《政府工作报告》。

在大规模减税降费政策背景下,预算管理直接体现了政府的行为导向(吕炜和王伟同,2022)。中央经济工作会议中指出“提升财政政策效能,就是要统筹财政资源,强化预算编制、审核、支出和绩效管理”,与此同时,政府部门多次提出要将减税降费政策落实在预算安排中,山东等省份在2013年已率先开展税式支出预算管理。鉴于小型微利企业(以下简称“小微企业”)是国民经济的重要支柱,各地税务机关根据《税总发【2019】13号》文件对小微企业减税政策实施过程进行数据统计核算,以客观反映减税政策的运行情况和经济效应。^①在众多小微企业减税政策中,所得税减半征收政策(以下简称“减半征收政策”)是小微企业受益甚广且扩围频繁的典型门槛式减税政策,该政策规定企业应纳税所得额一旦超过减税政策门槛值,将导致其所得税税率从10%升至20%,这容易引起纳税人在政策门槛值处产生群聚响应,以获得更低税率优惠(王伟同等,2020;李昊楠,2021)。在信息不对称的情况下,政府无法区分某项减税政策的受益对象是否来自群聚响应,对政策门槛值内的小微企业一律实施税收优惠,则导致减税政策的实际执行值超过预算值,造成预算偏差。在政府绩效管理中,通常将消除实际值与预算值之间的偏差作为绩效工作的核心(何文盛等,2013)。因此,科学预判减税政策对企业造成的应税行为反应,有效消除企业应税行为所产生的预算偏差,并且发现纳税人的群聚反应行为的动态规律,这将有助于实现预算管理水平的精进与资源统筹能力的提升,最终有助于提升减税政策效能。

科学预判纳税人应税行为反应的关键在于计算应税反应弹性,应税反应弹性能够直观捕捉不同减税政策规则下纳税人的应税行为变化,是研究最优税制理论、优化减税政策工具、刻画企业行为特征的核心参数,已受到国际税收学界的长期高度关注(Saez,2001;Devereux et al.,2014)。近年来,群聚方法(bunching)逐渐被用于测算非线性减税政策实施后纳税人的应税反应弹性(Kleven & Waseem,2013),与断点方法(RD)的不同之处在于,后者假定纳税人无法预判政策实施内容或时间,无法提前做出应税行为反应,只能“被动接受”新政策,这使得断点回归设计能够排除外生干扰,对新政策产生的净效应进行实证分析。然而,群聚方法适用于政策在实施前给予纳税人充分时间进行“主动反应”的研究场景,在刻画纳税人应税行为变化的基础上,评估其政策实施效应(张航和范子英,2019;Kleven,2016),即通过计算非线性税率政策下纳税人的群聚程度、摩擦系数、门槛值变化等参数,预判门槛性减税政策的应税反应弹性及其政策效应(Chetty,2012),若将其纳入事前预算管理,则可以有效提升减税政策的预算精准性,也为提升减税政策效能提供了实证研究方案。

本文基于2008—2016年全国税收调查数据,发现减半征收政策实施后小微企业出现了显著的群聚现象。在构建理论模型基础上,进一步分析小微企业群聚响应的动机及其预算影响,论证了群

① 国家税务总局:《关于深入贯彻落实减税降费政策措施的通知》。

聚响应和群聚反弹对减税预算存在影响。在实证分析部分,采用未受减税政策影响的企业真实分布频数构建反事实估计框架,发现企业群聚响应会改变减税政策的税基分布,且减税门槛值越高群聚响应越强烈,预算偏离越大;本文还发现减税政策扩围后,在原政策门槛附近发生群聚响应的企业的群聚动机消失,这些企业将“反弹”到发生凹陷(missing mass)的应纳税所得额箱体中,进一步扩充了新增减税区间的税基,这意味着如果将群聚响应和群聚反弹纳入预算管理方案中,则会大幅提升预算管理的精准化程度。随后,本文讨论了不同征收方式、不同行业类型、不同政策知悉程度、不同涉税服务规范性对企业应税行为反应的异质性影响,以及小微企业实现群聚响应的可能路径。最后,本文就减税政策的预算管理提出政策建议,为减税降费政策提质增效提供理论支撑与实证依据。

本文可能的边际贡献有以下四个方面:第一,在预算管理这一财政宏观治理体系中引入了群聚方法,丰富了预算管理的实证研究视角及方法,为减税政策纳入全口径预算管理体系提供了研究经验;第二,将群聚方法应用于预判企业应税行为反应,测算不同减税政策门槛设计下的企业应税反应弹性,将减税政策的精准化治理由事后评估向事前评估转变,为减税政策评估补充事前研究场景;第三,不仅分析了减税政策施行后纳税人的行为反应,而且刻画了政策消失后纳税人的行为反应,从理论和实证上论证了企业应税行为动态变化对政府预算管理的影响,补充了群聚方法在动态财税政策评估中的应用;第四,本文亦关注小微企业的减税政策实施状况,探讨小微企业进行群聚响应的实现路径,这对特定群体的减税政策提质增效具有重要作用。

本文余下的结构安排如下:第二部分为政策背景与文献综述;第三部分为理论模型与实证设计;第四部分是实证结果;第五部分为进一步分析,最后总结全文并提出政策建议。

二、政策背景与文献综述

(一)研究背景

一直以来,我国高度重视小微企业的生存发展。数据显示,“十三五”时期小微企业减免企业所得税政策累计为 2996 万户次,减免企业所得税超过 6910 亿元,覆盖九成以上的小微企业。本文所提到的小微企业,包含规模较小或者盈利微薄的法人企业以及个体工商户,其数量占全国市场主体总量的 96%,提供了 50% 以上的税收收入。从小微企业所得税减税政策沿革来看,《企业所得税法》中规定获得小微企业资质的企业,其所得税优惠税率为 20%,低于正常企业税率 25%;为了进一步降低小微企业税收负担,2010 年之后税务部门在小微企业所得税税率 20% 的基础上,实施应纳税所得额减半征收的政策(以下简称为“减半征收政策”),两项减税政策叠加使得小微企业的实际所得税税率降为 10%。具体而言,2010 年 1 月 1 日至 2011 年 12 月 31 日,减半征收的政策门槛为年应纳税所得额不高于 3 万元;2012 年 1 月 1 日至 2013 年 12 月 31 日,减半征收的政策门槛为年应纳税所得额不高于 6 万元;2014 年 1 月 1 日至 2014 年 12 月 31 日,政策门槛为年应纳税所得额不高于 10 万元;2015 年 1 月 1 日至 2015 年 9 月 30 日,政策门槛为年应纳税所得额不高于 20 万元;2015 年 10 月 1 日至 2016 年 12 月 31 日,政策门槛为年应纳税所得额不高于 30 万元。之后,政府不断扩增减半征收政策的受益范围,截至 2022 年已将减半征收门槛值扩围至 300 万元。^① 可见,该政策的受益面广、实施期限长,大量小微企业获此减税政策优惠。然而,鲜有研究从政府管理角度探讨该减税政策对预算管理的影响和挑战。

根据减半征收政策规定,当企业应纳税所得额低于减税政策门槛时,其所得税税率为 10%,超过减税门槛时,其所得税税率为 20%,如图 1 左侧所示,企业所得税税率在政策门槛值两侧跳跃上

① 因篇幅所限,相关内容详见本文附录 1。

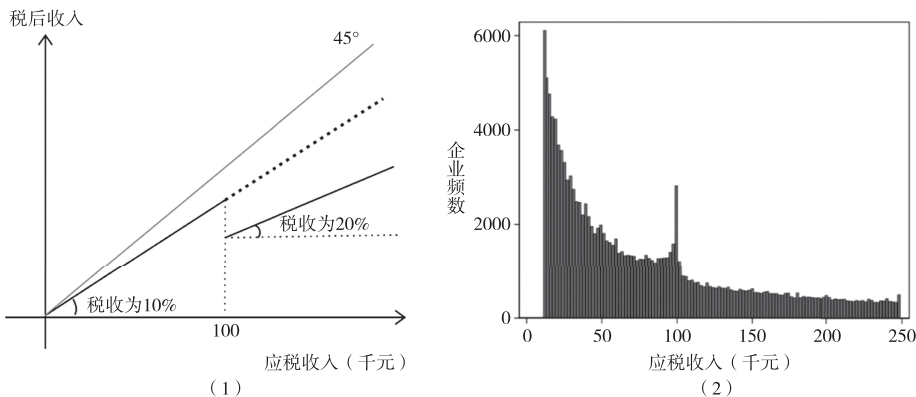


图 1 减半征收政策税率变化与企业群聚现象

升。图 1 右侧展示了 2014 年全国税收调查数据中小微企业的频数分布状况,可见,企业频数随着应纳税所得额上升而持续下降,仅在应纳税所得额为 10 万元处,企业频数异常高于两侧,这表明小微企业对减税政策产生了一定的群聚响应,改变了减税政策的税基分布状况。整体而言,该减半征收政策的主要特征有三:其一,处于减半征收政策门槛之外的企业无法获得减税红利,仅在门槛内的企业可以享受减税,该政策具有显著的门槛型特征;其二,减半征收政策频繁扩围,至今仍广泛使用,便于实证研究中及时捕捉企业的动态应税行为响应,减少因单一时点政策评估而引起的结论偏差;其三,小微企业对减半征收政策享受率逐年上升,受益企业数量和金额快速增加,几乎实现了“应享尽享”,导致减半征收政策的实施规模逐年上升,其预算管理问题值得各界关注。

传统的静态减税政策预算额的计算方法来自于减税政策预算中的收入放弃法 (revenue foregone method),即以基期的真实税收数据为基础,估算新减税政策导致的税收收入减少的金额,该方法适用于减税政策与纳税人挂钩的税式支出(白重恩和毛捷,2011),已被众多发达国家采用 (OECD,2010)。类似地,我国《预算法》中明确规定“要参考上一年预算执行情况、有关支出绩效评价结果和本年度收支预测,进行预算编制”。可见,减税政策的预算管理须对本年度的预算执行情况进行充分考量,再对未来进行预测,这与静态减税政策预算方法极为相似,即根据上一年度实际税基分布状况,加以预期减税政策规则后,编制出下一年度的减税预算规模。所有上述的减税政策预算方法都基于一个前提:纳税人的行为不会因减税政策的实施而发生重大改变,或者即使有所改变,也应在政府的正确预期范围内,如经济的总体增长等。然而,在门槛性减税政策实际执行中,大量实证研究证明纳税人或会根据减税政策门槛规定而产生行为响应 (Best et al.,2015;Harju et al.,2019;杨国超和芮萌,2020)。因此,原有的静态预算方法具有较大的优化空间,这为本文使用群聚方法研究减税政策预算管理提供了独特的契机。

具体而言,本文假设静态预算方法是在已知下一年新增减税政策的实施范围和减税税率的情况下,以本年度企业税基的分布作为预算基础,将预计新增扩围范围内的税基进行加总,再乘以税率降低幅度,即可得出下一年度新增减税政策的静态减税预算额,以 2012 年减税政策门槛由 3 万元扩围至 6 万元为例,其静态预算减税额以 2011 年应纳税所得额处于 3 万元至 6 万元的所得税税基进行汇总,再乘以税率降低幅度 10%,得出 2012 年预计新增减税额,属于预测减税额。^①而实际减税额则是来自于新增减税政策范围内的企业实际减税额进行的事后加总。表 2 中第(1)列表示新增减税范围内,满足减半征收政策要求并且享受该政策的小微企业数量占比,结果表明减半征收政策享受率在 2016 年已经上升至 99%,几乎实现了“应享尽享”;第(2)列展示了新增减税范围内,

① 减税税基的计算方法借鉴引自中国政府网《小微企业,减税红包有多大》。

所有小微企业中最终获得减半征收政策优惠的企业比重,可见,随着减税政策宣传与小微企业纳税水平的提升,减半征收政策的综合享受率逐步提高至67%。为排除各年政策享受率波动所导致的预算干扰,故本文的实证分析中始终假设该享受率为已知数。随后第(3)、(4)列展示了小微企业新增减税政策区间内的真实减税额与静态预算额,该真实减税额是指政策新增或扩围区间内的实际减税额,为事后汇总数据。由第(5)列可知小微企业减税政策实施过程中普遍存在实际减税额超过静态减税预算额的特征事实,表明传统静态预算方法已经不适用于企业频繁发生群聚响应的现实状况,这为前述的减税降费政策超预算提供了一项可能的解释,也为国内外减税政策预算管理优化提供理论参考。接下来,本文就以小微企业的应税行为反应对预算管理的影响为研究视角展开详细分析。

表 2 减税征收政策预算偏离 单位:千元

年份	减半征收 区间享受率 ^①	综合政策 区间享受率	真实减税额	静态预算额	初始偏离额	超预算程度
	(1)	(2)	(3)	(4)	(5)	(6)
2010	35.28%	7.03%	10826	7518	3308	44.00%
2012	82.58%	26.88%	24592	27103	2511	9.26%
2014	93.28%	58.40%	108179	65372	42807	65.48%
2015	98.74%	53.45%	198727	149020	49707	33.36%
2016	99.48%	66.14%	238377	181247	57130	31.52%

数据来源:数据来自于2010—2016年税收调查数据统计与测算^②

(二)文献综述

本文首次应用群聚方法研究小微企业应税行为反应对减税预算管理的影响,有可能对减税政策预算、小微企业的减税降费效能评估和市场主体的群聚行为等领域研究提供新的视角,对此,本文将从以下三个方面对文献进行系统梳理。

第一,关于对减税降费政策进行预算管理的相关研究。大规模减税降费政策实施期限长、规模大,不仅可以提升就业水平(甘犁等,2019)、影响企业投资决策(Chen et al.,2021)、鼓励企业创新发展(刘啟仁,2023),但这也给财政可持续性(郭庆旺,2019)、预算管理(汪德华等,2014)带来挑战。预算管理作为财政治理的重要支撑,对减税政策进行预算管理将会对财政制度的透明度、效率产生促进作用(郑方辉和费睿,2019)。因此,美国、加拿大、德国等众多发达国家都对减税降费政策进行了较为完善的预算管理制度(国家税务总局济南市税务局课题组等,2020;OECD,2010),其主要采用收入放弃法测算减税政策的预算金额,这将有效规范减税政策资金管理,提升减税政策的实施效率(白重恩和毛捷,2011;Bogacheva & Fokina,2017)。在我国现行预算管理制度中,不少学者呼吁建立税式预算制度(白重恩和毛捷,2011;汪德华等,2014),这有助于实现减税降费政策的科学监督、预算管理,并且控制减税政策实施金额、范围与结构,使减税政策的绩效评估更加精准(高培勇与汪德华,2016),促进减税政策提质增效(贾康和段爱群,2013)。已有不少学者论证了减税政策预算的必要性与实用性,却鲜有文献从实证角度讨论减税政策实施过程中的预算管理问题。

第二,关于小微企业的减税降费政策及其效能分析。小微企业是大众创业的摇篮(OECD,

① “区间”是指政策扩围后的新增减税区间,每年分别是(0,30],[30,60],[60,100],[100,200],[200,300]。该部分计算的是扩围区间内的减半征收企业占优惠税率企业的比重,综合享受率是指减半征收企业占全部区间内小微企业的比重。全区间企业的减税政策享受率见本文附录1。

② 由于2011年、2013年减税政策门槛值并未改变,不存在新增减税政策预算,因此本表未曾列出。本文后续将重点关注五次政策变化时点。

2010),可以有效促进市场竞争、激发经济活力、引导企业创新、提高企业全要素生产率(Saito et al., 2018),但是小微企业长期面临不确定性高、净资产较少、人力资源较缺乏的生存困境(甘犁等, 2019),因此各国为扶持小微企业发展出台了多项积极的财税措施(Decker et al., 2014),我国也不断扩围小微企业减税降费政策(朱武祥等, 2020)。在大规模减税降费政策实施的背景下,如何提升减税政策效能是学界关注的重点。政策效能中“效”是指政策结果的效率、效果、效益,而“能”指政策过程中的“治理能力”(范柏乃和林哲杨, 2022),提升效能不仅在于实现过程与结果的统一(吴建南等, 2011),而且需要提升政策的精准发力水平(白景明, 2022)。此外,樊丽明等(2022)认为可以从预算监督着手,优化预算管理制度,提升资源统筹能力,以推动政策效能提升。可见,提升减税政策效能意义重大,值得学界广泛关注。

第三,关于减税政策背景下企业群聚响应的研究。郑方辉和费睿(2019)认为减税政策绩效取决于动力机制、工具体系及其二者有机合成,而政府的各项经济政策亦会因为经济主体的行为响应而影响实施效果,减税政策亦是如此(Chen et al., 2021; 刘啟仁等, 2023)。已有研究证明,门槛型减税政策容易引起纳税人的群聚响应,而且这种群聚响应可以通过合法合规的纳税调整得以实现,这一现象在世界各国普遍存在(Kleven, 2016; Harju et al., 2019),例如美国企业所得税(Mortenson & Whitten, 2020)、个人所得税(Feldman et al., 2016)、社保费用(Gelber et al., 2020)、房地产税(Kopczuk & Munroe, 2015),巴基斯坦的个人所得税政策(Kleven & Waseem, 2013)、企业所得税减税政策(Best et al., 2015),瑞典个人所得税(Bastani & Selin, 2014),英国企业税收返还(Devereux et al., 2014)等门槛性政策无一例外均出现了群聚现象。在国内的研究中,王伟同等(2020)发现小微企业会根据减半征收政策进行群聚响应,且该减税政策能够有效缓解小微企业的融资约束;范子英等(2022)发现企业根据工业用地价格管制的门槛值进行群聚响应,而Chen et al. (2021)发现企业会根据减税政策门槛而操纵其研发费用投入占比,并获得减税红利。然而,上述文献使用群聚方法几乎都围绕着政策初次发生(Kleven, 2016; Diamond & Persson, 2016; 张航和范子英, 2019),鲜有文献关注政策门槛变化后的纳税人行为变化。

综上所述,本文可能在以下四个方面拓展与丰富现有文献的研究:第一,本研究通过搭建减税政策与预算管理的实证研究桥梁,试图丰富预算管理精准性的相关实证研究,为进一步提升减税政策效能提供实证经验;第二,本文将事前理论测算纳入减税政策的预算中,拟通过预算管理视角讨论如何提升政策效能,为精准化预算管理提供经验参考;第三,与单一政策评估不同的是,本文补充关于减税政策多次扩围后纳税人行为变化及其后果的经验研究,试图补充适应扩围式减税政策的实证研究场景,拓展减税政策动态评估的研究思路;第四,本文还分析了小微企业在动态减税政策下产生的差异化应税行为反应,为理解特定群体的应税行为反应、优化减税政策工具提供了现实依据。

三、理论模型与实证设计

本文首先建构一个减税政策和企业行为的模型刻画企业群聚响应的动机与减税政策预算偏离的成因,在此基础上进行实证设计。

(一) 企业行为

首先考虑减税政策门槛(notch point)为应纳税所得额 π_1^* 时的群聚情况。根据Kleven & Waseem(2013)可知,企业的效用函数 u 可以表示为:

$$\max_e u = (1 - t - \Delta t \cdot 1[\pi_1 > \pi_1^*]) \cdot \pi_1(e) - \frac{n}{1 + 1/\varepsilon} \cdot \left(\frac{\pi_1(e)}{n} \right)^{1 + 1/\varepsilon} \quad (1)$$

其中, e 表示企业经营的努力程度,努力程度会对企业的应纳税所得额 π_1 有正向的影响, π_1'

(e) > 0 且 $\pi''_1(e) < 0$ 。根据减半征收政策规定,当企业应纳税所得额大于减税门槛 π_1^* 时,其所得税税率为 $t + \Delta t$,反之则面临的税率是 t 。企业效用函数的第二项为努力程度的副作用。此外, n 为企业的能力系数, ε 为应纳税反应弹性。当企业应纳税所得额低于门槛时,企业不会对减税门槛发生行为响应,可以求得一阶条件为:

$$\pi_1(e) = n(1-t)^\varepsilon$$

对于能力系数为 $n^* + \Delta n^*$ 的企业,若不存在减税门槛,其创造的应纳税所得额为 $\pi^* + \Delta \pi^*$,该企业的能力系数和应纳税所得额满足条件 $\pi^* + \Delta \pi^* = (n^* + \Delta n^*)(1-t)^*$ 。若存在纳税门槛,该企业在面临存在减税门槛的税收约束条件下,可以选择将企业应纳税所得额调整到 π_1^* 或者 π'_1 。这里 $\pi'_1 > \pi_1^*$,同时满足 $u(\pi_1^*) = u(\pi'_1)$ 。此时,通过整理可以得到:

$$\frac{1}{1 + \Delta \pi_1^* / \pi_1^*} - \frac{1}{1 + 1/\varepsilon} \left(\frac{1}{1 + \Delta \pi_1^* / \pi_1^*} \right)^{1+1/\varepsilon} - \left(\frac{1}{1 + \varepsilon} \right) \left(1 - \frac{\Delta t}{1-t} \right)^{1+\varepsilon} = 0 \quad (2)$$

在此公式中只有 $\Delta \pi_1^*$ 和 ε 为未知数,当 $\varepsilon \rightarrow 0$ 时,可以得到 $\Delta \pi_1^* = \frac{\Delta \pi_1^*}{1-t+\Delta t} \equiv \Delta \pi_1^D$ 。当应纳税反应弹性趋近于零时,企业效用最大化决策中的无差异曲线呈里昂惕夫型。很显然,对于税前所得处在区间 $(\pi_1^*, \pi_1^* + \Delta \pi_1^D)$ 的企业,其税后所得小于税前所得为 π_1^* 的企业。因此,如果假设企业在调整税前利润的过程中不存在摩擦,应纳税所得额处于区间 $(\pi_1^*, \pi_1^* + \Delta \pi_1^D)$ 的所有企业都会群聚在减税门槛 π_1^* 。然而,由于现实生活中收入刚性和政策知悉程度低等因素,使得企业无法全部转移到减税门槛左侧,则可以定义企业在调整过程中的优化摩擦系数为:

$$a_1 = \frac{\int_{\pi_1^*}^{\pi_1^* + \Delta \pi_1^D} f(\pi_1) d\pi_1}{\int_{\pi_1^*}^{\pi_1^* + \Delta \pi_1^D} f_0(\pi_1) d\pi_1} \quad (3)$$

其中, $f_0(\pi_1)$ 为应纳税所得额的反事实频数分布,即不存在群聚行为时的分布, $f(\pi_1)$ 为应纳税所得额的实际频数分布。摩擦系数可以解释为,如果只考虑企业税后利润最大化,理论上应该群聚到门槛处而实际中没有群聚到门槛处的企业数量占比。在考虑应纳税反应弹性的基础上,通过摩擦系数可以估计群聚区间 $(\pi_1^*, \pi_1^* + \Delta \pi_1^*)$ 。在群聚区间 $(\pi_1^*, \pi_1^* + \Delta \pi_1^*)$ 上,企业群聚程度 b_1 可以表示为:

$$b_1 = (1 - a_1) \int_{\pi_1^*}^{\pi_1^* + \Delta \pi_1^*} f_0(\pi_1) d\pi_1 \approx (1 - a_1) f_0(\pi_1^*) \Delta \pi_1^* \quad (4)$$

由于通过实际分布与反事实分布可以求得 a_1 、 b_1 和 $f_0(\pi_1^*)$,所以通过此方程可估算出 $\Delta \pi_1^*$ 。当企业的减税政策扩围,其减税门槛从 π_1^* 移动到 π_2^* , (4) 式中的下角标由 1 变成 2,理论上企业在应纳税所得额为 π_2^* 处发生群聚响应的结论依然成立。

(二) 政府预算

假设没有减半征收政策,此时企业应当缴纳的所得税税率均为 $t + \Delta t$,政府对小微企业的企业所得税的预算收入为 $(t + \Delta t) \int f_0(\pi_i) d\pi_i, i = 1, 2$ 。当政府计划对应纳税所得额小于或等于 π_i^* 的企业进行减半政策时,如果企业不调整税前利润,该减税政策的静态预算额为 $\Delta t \cdot \int_0^{\pi_i^*} \pi_i f_0(\pi_i) d\pi_i$ 。考虑到企业税前利润对减税政策产生群聚响应,税前利润的频数分布会发生变化,此时企业的实际分布表示为 $f(\pi_i)$,政府减税总额为 $\Delta t \cdot \int_0^{\pi_i^*} \pi_i f(\pi_i) d\pi_i$,则实际减税额和静态减税预算的预算偏差(TCBD, tax cut budget deviation)为:

$$TCBD_i = \Delta t \left[\int_0^{\pi_i^*} \pi_i f(\pi_i) d\pi_i - \int_0^{\pi_i^*} \pi_i f_0(\pi_i) d\pi_i \right] \quad (5)$$

在图 2 左侧中,左边浅灰色曲线(l_1)表示上一年度减税税基分布状况,黑色曲线 l_2 表示本年度减税税基分布状况。如果不考虑优化摩擦,那么上一年度应纳税所得额处于 (π_1^*, π_1') 的企业将转移到 π_1^* 左侧,在减税税基分布上形成了左边凸起、右边凹陷的税基分布状况。当本年度减税政策扩围时,新增减税政策预算是基于 (π_1^*, π_2^*) 区间内的企业分布状况进行测算,此时需要考虑第一次变化:新门槛处的企业群聚响应对减税税基分布的影响,还要考虑第二次变化:原门槛值处企业不再群聚响应后对新增减税政策税基的补充,换言之,在原政策门槛值 π_1^* 左侧发生群聚响应的部分将反弹至右侧凹陷处,导致减税政策税基再次上升。图 2 右侧清晰展示了两部分预算偏差的来源,在新增减税税基 (π_1^*, π_2^*) 区域内,第一层偏差是在当年新增门槛后产生的新群聚反应导致的,第二层偏差是失效门槛处的企业发生群聚反弹所导致的,这可以更大程度地解释产生预算偏差的原因。

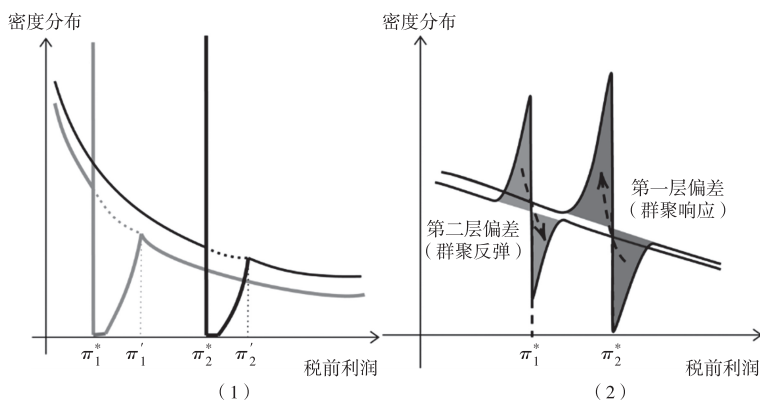


图 2 跨年度税基变化状况

此外,当新的减半征收门槛实施后,原处于 π_1^* 左边发生群聚响应的企业反弹回 π_1^* 右侧,恰好处于新增减税政策税基区间 (π_1^*, π_2^*) 中,填补了减税政策预算的税基,亦会产生减税偏差,表示为:

$$TCBD_{re} = \Delta t \left[\int_{\pi_1^*}^{\pi_1^* + \Delta \pi_1^*} \pi_2 f(\pi_2) d\pi_2 - \int_{\pi_1^*}^{\pi_1^* + \Delta \pi_1^*} \pi_1 f(\pi_1) d\pi_1 \right] \quad (6)$$

在现实中,并非全部小微企业皆能获得税收优惠,企业会出于政策知悉程度不足、纳税资料不全或适用其他税收优惠政策等原因,造成该减半征收政策的享受率并非百分之百。因此,本文假定每年小微企业对减半征收政策的享受率为系数 γ_i ,且该享受率系数能够由政府于预算编制中精准估计,因此:

$$real_TCBD_i = TCBD_i \cdot \gamma_i \quad (7)$$

综上所述,理论模型发现减半征收政策会导致纳税人产生群聚响应,而群聚响应会导致减税政策出现超预算额执行现象。当减税政策扩围之后,企业群聚响应随之变化,其群聚响应、群聚反弹行为可以同时扩大减税税基,则减税预算偏差进一步扩大。此外,由于每年度的经济状况差异导致减税政策税基分布不尽相同,则精准化预算管理还需要考虑跨年度间减税政策税基分布的正常波动。因此,在排除减税政策享受率的干扰后,减税预算偏差来自于三个部分,一是上一年度发生群聚响应的企业在本年度发生群聚反弹行为所增加的减税税基,二是本年度新增减税门槛引起的企业发生群聚响应行为使减税税基增加,三是本年度与上一年度由于其他原因造成的企业税基分布

曲线差异。^①

(三)实证设计与数据说明

1. 实证设计

群聚方法的核心思想是假设门槛性政策仅对门槛值附近的纳税人产生政策影响,而其余区域的纳税人不受政策影响,可以利用不受政策影响的分布曲线去估计受到政策影响区间的反事实分布,最后将反事实分布曲线与受到政策影响的真实分布曲线进行对比,评估企业群聚响应的影响。对反事实曲线进行估计,基本模型要满足两个条件:一是假设反事实分布曲线是连续且平滑可微的,即使政策门槛左右两侧的企业频数分布发生改变,仍可以从非群聚区域推断出反事实分布曲线;二是群聚响应本身不会使企业数量整体上升或下降,无论是否发生摩擦,必须满足左侧企业数量增长量与右侧企业数量减少量相等,即发生群聚响应的企业全部由门槛右侧转移而来,群聚响应不会引起企业总数的增减。借鉴 Kleven(2016)的估计思路,可以设定反事实曲线为:

$$c_j = \sum_{i=0}^p \beta_i \cdot (z_j)^i + \sum_{r \in R} \rho_r \cdot 1\left[\frac{z_j}{r} \in N\right] + \sum_{i=z_L}^{z_U} \gamma_i \cdot 1[z_j = i] + \nu_j \quad (8)$$

其中, c_j 是第 j 个税前利润区间中企业数量, π_j 是第 j 个区间上企业的税前利润水平, $[z_L, z_U]$ 是估计反事实曲线的排除区间。 $1[z_j = i]$ 为应纳税所得额是否在排除区间上的虚拟变量。同时,因为部分小微企业常以整百、500 或 1000 的倍数填报调查表,这会导致企业频数分布中呈现出明显的整数效应,会干扰反事实曲线估计的精准性(王伟同等,2020;李昊楠,2021)。因此,本文控制了整数效应对企业频数分布的影响。在估计反事实曲线的过程中集合 $R = \{500, 1000\}$, N 为正整数集合, $1[(z_j/r) \in N]$ 为是否存在整数效应的虚拟变量。在估计实际分布时,应纳税所得额考虑整数效应的实际分布曲线为:

$$\hat{c}_j = \sum_{i=0}^p \beta_i \cdot (\pi_j)^i + \sum_{r \in R} \rho_r \cdot 1\left[\frac{\pi_j}{r} \in N\right] \quad (9)$$

由上式可以推出,减税政策门槛左侧发生群聚的企业频数为 $\hat{B} = \sum_{j=\pi_L}^{\pi^*} (c_j - \hat{c}_j)$;同时,减税政策门槛右侧减少的企业频数为 $\hat{M} = \sum_{j=\pi^*}^{\pi_U} (\hat{c}_j - c_j)$ 。

估计反事实分布的核心在于最优排除区间和最优阶数,本文借鉴 Diamond & Persson(2016)最优区间与最优阶数的确定方法,引入 Bosch et al. (2021)与 Bergolo et al. (2021)所使用的数据驱动方法进行改进。本文以 $\hat{M} = \hat{B}$ 作为先验条件,对排除区间的上界、下界、多项式阶数同时迭代,直到反事实曲线与真实分布曲线出现交点,并且多项式拟合结果的均方误差最小,即得到估计反事实的最优参数组合;最后,对估计结果运用与 Diamond & Persson(2016)、王伟同等(2020)类似的 5 层交叉检验方法,验证多项式阶数、排除区间的可靠性,选择均方误差之和最小值对应的最优估计,再根据理论模型和 Kleven & Waseem(2013)计算弹性等统计量,并使用 500 次自助法获得标准误进行统计推断。为了进一步排除其他同门槛政策对实证结果可能造成的干扰,本文使用每年度企业群聚水平趋势减去上一年同区间内的群聚水平,得出因本年度新增减税门槛而产生净增群聚效应所改致的预算偏差:

$$\Delta TCBD = TCBD_2 - TCBD_1 \quad (10)$$

接下来,实证分析利用上述公式,通过群聚方法识别出了减税政策的税基变化情况。

2. 数据说明

本文使用 2008—2016 年全国税收调查数据、中国小微企业调查数据(CMES)。前者每年约

^① 第三种差异也可能是由于抽样波动和其他因素带来的总体分布差异,正文中主要分析了群聚响应和群聚反弹的影响。因篇幅所限,相关内容详见本文附录 5。

60—70 万家企业填报,涉及大中小规模企业的纳税数据和经营信息(高培勇和毛捷,2013),涵盖核定征收和查账征收的小微企业数据,是目前研究小微企业纳税信息的优良样本。在数据处理上,本文根据小微企业认定标准,从数据中筛选出就业人数低于 100 人并且平均资产低于 3000 万元的工业企业、就业人数低于 80 人并且平均资产低于 1000 万元的其他企业,剔除了应纳税所得额非正或者缺失的样本。CMES 来自于西南财经大学 2015 年对全国 28 个省市共 5601 户小微企业的调查问卷,重点关注家庭作坊式企业营商税收情况(甘犁等,2019),本文依托其抽样推断全国相关地区和企业政策知悉程度。

四、实证分析

(一)群聚响应与预算偏离

1. 基准结果

首先,本文通过群聚方法验证企业群聚响应的逐年变化及其对预算的影响。根据 2010—2016 年税收调查数据中企业分布的特征,假定企业频数分布整体上满足类似于 Pareto 的连续频数分布,企业频数随着应纳税所得额提高而降低;随着减半征收政策实施与扩围,在每一年的减税门槛处均出现了明显的企业群聚现象,展现为企业频数在减税政策门槛两侧“左高右低”,并且这一现象随着政策门槛扩围而愈发显著。当减税门槛变化后,原减税门槛附近的群聚现象消失,在新的减税门槛处又出现了显著的群聚响应现象,这与 Best et al. (2015)的发现类似,企业的群聚行为会随着减税门槛扩围而变化,这正适用于本文研究扩围式门槛性减税政策带来的动态群聚响应的研究主题。^①

表 3 2010—2016 年减税政策对小微企业群聚响应的影响:主结果 单位:千元

年份	减税门槛	操纵区间	群聚程度	应税反应弹性	摩擦系数	预算偏差(TCBD)	预算偏差解释度	改进后预算精度度
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2010	30	[26,36]	0.536 *** (0.117)	0.152 *** (0.0331)	0.964 *** (0.0310)	1917.575 *** (632.706)	4.08%	71.60%
2012	60	[50,64]	1.163 *** (0.323)	0.165 *** (0.0457)	0.710 *** (0.0313)	3276.445 *** (1001.566)	35.06%	87.04%
2014	100	[91,133]	5.430 *** (0.694)	0.4610 *** (0.0589)	0.633 *** (0.0342)	22210.988 *** (3112.728)	30.30%	83.54%
2015	200	[187,269]	11.298 *** (1.143)	0.480 *** (0.0485)	0.678 *** (0.0288)	39750.094 *** (3975.998)	42.75%	88.36%
2016	300	[263,373]	19.971 *** (5.0555)	0.565 *** (0.143)	0.528 *** (0.0550)	62178.844 *** (14694.347)	71.98%	91.88%

注:(1)括号内为标准误,标准误采用自助法重复抽样 500 次获得;(2)*、**和***分别表示在 10%、5%和 1%水平下显著,下同。

随着减半征收政策的门槛不断扩围,小微企业的群聚程度逐年上升,由 2010 年 0.54 上升到 2016 年的 19.97,意味着 2010—2016 年间减税门槛右侧来自群聚响应的企业比重显著上升,企业的应税反映弹性也从 0.152 上升到 0.565,可见,减税政策扩围带来了更高的群聚程度与应税反应

① 因篇幅所限,相关内容详见本文附录 3。

弹性。^① 在减半征收政策实施初期,发生群聚响应的企业较少,主要原因有二,一是由于小微企业财务信息不健全、税务监管成本高等原因,所得税征收主要采用核定方式,该部分企业仅需缴纳固定税额,无法享受减半征收政策,因此早期小微企业整体群聚程度低;2014 年以后,采用核定征收的小微企业亦可同等享受减半征收政策,并且减税政策惠及企业数量不断上升,发生群聚响应的企业显著增加;二是小微企业认定门槛在税务、工信、统计部门尚未统一,小微企业对减税政策把握不足,缺乏实践操作经验,此情况随着“放管服”改革而得到缓解。因此,在小微企业减税政策多次扩围后群聚响应行为愈发显著,引起的减税政策预算偏差亦随之增加。而应纳税所得额越高的市场主体,其财务水平越高,即随着减半征收政策门槛的提高,减税门槛附近的小微企业财务水平也可能更高,更能够对减税政策进行知悉并且灵活运用,其摩擦系数会显著降低,实证结果也验证了随着减税门槛的提升,摩擦系数下降,这会给企业带来更多的减税优惠,也给减税政策预算带来了显著的偏差。为保证多年度的计算结果可比性,在主结果中统一采用多项式阶数为 7 阶、箱体宽度(Bin)为千元进行估计,测算出 2010 年减税政策的预算偏差为 191.76 万元,在减税门槛处的税基扩大与企业群聚响应逐年上升的双重效应下,减税预算偏差快速攀升,2016 年的减税政策预算偏差达到了 6217.88 万元,这显著降低了政府对减税政策进行事前预算的准确性。总体而言,主结果说明小微企业会利用门槛式减税政策进行群聚响应,并且响应程度随着减税政策门槛扩围而逐年上升,其群聚响应行为给政府带来了显著的减税预算偏差。

小微企业产生的群聚响应现象可以多大程度解释初始的减税预算偏差呢?根据实证设计,表 3 中第(7)列展示了群聚响应行为所产生的预算偏离对整体预算偏离的解释程度,可见,减半征收政策初次实施以及四次扩围后,小微企业产生的群聚响应行为可以解释预算偏离约 4%—72%,该解释程度随着企业群聚程度、应税反应弹性、预算偏差上升而增强,这也意味着如果将群聚因素纳入事前预算管理中,将大幅提高预算管理的精准度。可见,结构性减税政策契合供给侧改革的时代要求,不断扩围小微企业所得税减半征收政策,一方面小微企业整体税负逐年下降,实现更大程度的惠企利民;另一方面,门槛性减税政策也引发了企业的应税行为反应,这对政府预算管理、税收征管、政策设计等会产生更多挑战,应当引起政府和学界关注。

2. 稳健性测试^②

(1)更改多项式阶数

群聚方法的核心是反事实估计曲线是否可靠,而不同的多项式阶数对反事实曲线的拟合程度不同,直接影响反事实曲线估计的准确性(Kleven,2016)。借鉴 Kleven & Waseem(2013)使用变换多项式阶数的方法进行稳健性的检验,本文使用 6 阶多项式阶数进行稳健性测试。从检验结果来看,小微企业在 2010—2016 年间均发生了显著的群聚响应现象,这与主结果一致,一定程度上体现了反事实估计系数的稳定性,验证了小微企业的群聚程度、应税反应弹性随着减税政策扩围而愈发强烈,其摩擦系数逐年缩小,并且减税政策的预算偏差逐年上升。

(2)更改排除区间

群聚方法中排除区间的上下界是企业真实频数分布区间与反事实估计区间的分割点,该上下界会随着减税政策门槛的变化而变化,其范围即为反事实估计中需要拟合的范围,因此,准确估计该上下界是群聚方法的核心(Diamond & Persson,2016)。在稳健性测试中,为了减少样本选择的偏误,本文将排除区间的范围进行了改变,在多项式阶数不变的情况下,选择数据驱动方法结果中的

① 2011 年和 2013 年小微企业减半征收政策并未扩围,该年份的实证结果不影响主要结论。限于篇幅,相关结果未予提供,如需可向作者索取。

② 因篇幅所限,相关内容详见本文附录 4。

次优区间进行验证。从检验结果看,小微企业的群聚程度、应税反应弹性逐年增强,所带来的减税预算偏差也更加显著。这与前述的研究结论相似,说明小微企业在不同的排除区间下,都可以清晰识别群聚响应不断增强的现象,这给减税政策造成了不可忽视的预算偏差。

(3) 更换标准误计算方法

为了进一步保证估计结果的准确性,本文参照张航和范子英(2019)的标准误估算方法进行稳健性检验。与基准回归中使用五层抽样方法不同的是,先以全样本完成第一次群聚估计,然后将所提到的非排除区间的残差放回进行随机自抽样,形成一个新的残差序列,再将其加回非排除区间的反事实估计中,得出非排除区间的企业频数分布,最后加上真实企业频数分布的操纵区间部分,即可完整模拟企业频数分布曲线进行再次群聚估计。以上步骤随机重复 500 次,可得到多个估计值,并计算出其标准误。本次稳健性测试与主结果类似,表明小微企业随着减税政策门槛值的提升而产生群聚程度、应税反应程度上升,对预算管理形成了更强的影响,这也验证了本文主结果的可信性。

(4) 证伪检验(falsification)

本文借鉴双重差分中平行趋势检验的思路,一种方案是考虑变更更减税政策的实施时间,若样本企业仅在当年政策冲击后才发生响应,而对伪政策冲击年份,不会出现显著的政策响应,则表明主结果是稳健的;另一种方案是考虑变更更减税政策门槛值,对同一年度数据样本,观察小微企业是否会对伪门槛值发生显著的群聚响应。为了减少样本选择的偏差,本文将每一年度真实企业分布数据与其他年份、其他门槛值进行配比,以分析企业是否会在伪政策实施年份、伪政策门槛值发生群聚响应行为。检验发现,2008—2009 年减半征收政策尚未实施,小微企业未发生明显的群聚响应行为,而 2010 年之后,企业仅会对当年的减税门槛产生群聚响应行为,对于之前或者之后年份的减税门槛均未发生显著应税行为响应,即小微企业对伪政策实施时间、伪减税门槛值并未出现群聚响应,这一结论为后文分析群聚反弹现象奠定基础。

总体上,上述多项检验表明企业发生群聚响应并非偶然,且群聚响应行为与政策实施年份、政策规定门槛值显著相关。

(二) 群聚响应对预算偏离的解释

上述实证分析展示了 2012—2016 年间,小微企业根据 4 次扩围后的减半征收减税政策门槛,产生不同程度的群聚响应行为,政策实施中出现了大量“原不该享受也要享受”的小微企业,造成减税政策的实际减税额超过预算额,这可以一定程度上解释预算偏差的成因。同时,经过前述的证伪检验可知,减税门槛值扩围让原门槛值处的企业不再进行群聚响应,“反弹”至原来的箱体中,而新增减税政策区间并未充分考虑这部分反弹带来的税基变化,产生了再次的预算偏离。鉴于企业频数分布存在一定的固定趋势,比如数据填报的整数效应和财务处理惯性等导致频数分布出现规律性起伏,为了保证测算结果的科学性,本文考虑去除上一年趋势后增量的群聚程度,在排除上一年度企业分布趋势的影响之后,就可以得出本年度新增减税政策的群聚响应行为对减税政策预算的净影响。

表 4 第(1)列展示了根据公式(10)得出当年度企业发生净群聚响应导致的预算偏差额,如表 4 第(3)列实证结果可知,如果将净群聚响应纳入预算管理中,将更好地解释预算偏离现象。与此同时,本文还讨论了企业发生群聚反弹所导致的预算偏差,第(2)列展示了根据公式(6)计算出的预算偏差,第(4)列则是考虑净群聚响应和群聚反弹后的预算精准度,意味着如果将群聚响应和群聚反弹纳入减税政策的预算管理,精准度将提升到 78.80%—97.51%。此外,本文还考虑了抽样波动的相关变化。结果表明,在控制其他因素不变的情况下,考虑企业在新门槛的群聚响应和旧门槛的群聚反弹,总体上可以实现减税政策预算管理精准度达 78% 以上。^①

① 因篇幅所限,相关内容详见本文附录 5。

表 4

群聚响应和群聚反弹对预算偏离的解释

单位:千元

年份	净群聚响应与群聚反弹 解释的预算偏差		改进后的 预算精准度	
	$\Delta TCBD$	$\Delta TCBD_{re}$	考虑净群聚响应	考虑净群聚响应 和群聚反弹
	(1)	(2)	(3)	(4)
2010	2754.512 ** (1117.670)	\	71.24%	\
2012	3025.972 *** (965.671)	3116.598 ** (1272.327)	86.48%	83.08%
2014	26072.861 *** (4682.398)	7965.653 *** (2095.580)	74.50%	78.80%
2015	37433.206 *** (5817.217)	34348.0375 *** (6769.840)	85.06%	94.30%
2016	46209.116 *** (6532.254)	49149.222 *** (7738.387)	88.85%	97.51%

注:因 2010 年减半征收政策首次施行,不存在群聚反弹,上表中使用“\”表示空缺值。

本文讨论了企业在减半征收政策下的群聚行为反应,导致减税政策金额超预算执行,该结论在减税政策多次扩围中得以反复验证。可见,相比于单一考虑企业群聚响应行为,将群聚响应和群聚反弹纳入预算管理方案可以进一步提升减税政策预算的精准化。尽管上述预算偏离额的解释无法 100% 精准预测减税政策的实际执行额度,但本文所讨论的企业群聚因素能够对减税政策预算偏离进行良好的解释,如果将企业群聚因素纳入减税政策事前预算管理体系中,不仅可以充分消除减税政策执行过程中因企业行为变化带来的预算偏差,也可以增强财政资源的统筹规划能力,提升减税政策效能,为解决我国大规模减税降费政策持续超预算执行问题提供了可能的思路。

五、进一步分析

2020 年起政府实施减税降费政策强调更加精准可持续,重点突出精准性与针对性。因此,实施减税政策不仅要关注预算管理的精准性,还需注重企业自身特点、行业、地域差别并进行分类施策。同时,减税政策设计不仅要预判企业在减税政策实施后的行为反应变化,更要关注企业实现群聚响应的具体路径,这对提升减税政策效能具有重要意义。接下来,本文从异质性维度和企业行为路径分析群聚响应行为。

(一)异质性分析^①

1. 征收方式不同

多年来,小微企业通过查账征收或核定征收的方式进行税款缴纳。在减半征收政策实施伊始,减税范围仅仅针对采用查账征收的小微企业,该部分小微企业数量较少,其优化摩擦系数接近 1,因此几乎不存在显著的群聚响应行为。随着企业对减税政策的进一步了解和应用,采用查账征收的小微企业逐步实现了群聚响应。另一部分小微企业应用核定征收方式进行纳税,该类企业按照成本或者收入核定纳税,在 2014 年之前,该类企业无法享受减半征收政策优惠,因此企业并无群聚响应动机。国家税务总局 2014 年第 23 号文件中规定采用核定征收方式的企业也可以享受减半征

^① 针对每一组异质性分析,本文还讨论了群聚程度、摩擦系数等指标的分组检验,限于篇幅,相关结果可向作者索取。

收政策,这部分小微企业产生群聚响应动机,并且造成群聚现象。随着减税门槛与征管水平不断提升,新增减税门槛值附近的小微企业采用核定征收的比例逐渐下降,采用查账征收方式的企业占比逐渐上升,群聚程度进一步增强。^① 因此,表 5 中第(1)、(3)列表明不同的征收方式下的小微企业的应税反应弹性具有明显差异,产生了差异化的预算偏差,即采用查账征收方式的企业应税反应弹性始终高于采用核定征收的企业,表明今后税务部门可以根据不同征收方式对小微企业进行差异化的减税设计、税收监管和税政宣传。

表 5 异质性检验:征收方式差异与不同类型的服务业 单位:千元

年份	查账征收		核定征收		生活性服务业		生产性服务业	
	弹性	预算偏差	弹性	预算偏差	弹性	预算偏差	弹性	预算偏差
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2010	-0.0315 (0.0365)	162.615 (503.0154)	0.0564 (0.0367)	53.106 (64.705)	0.230 *** (0.0368)	1534.973 *** (370.309)	0.0408 (0.0841)	41.786 (68.566)
2012	0.180 *** (0.0350)	3803.841 *** (696.889)	0.0214 (0.0282)	-37.214 (88.521)	0.222 *** (0.0249)	2248.169 *** (303.158)	0.115 *** (0.0322)	365.210 *** (93.849)
2014	0.577 *** (0.0830)	22833.586 *** (3444.169)	0.107 *** (0.0257)	866.00540 *** (224.511)	0.561 *** (0.0838)	13240.128 *** (2032.539)	0.201 *** (0.0636)	2576.07 *** (684.774)
2015	0.621 *** (0.0575)	41752.348 *** (3971.440)	0.108 *** (0.0311)	1303.990 *** (416.937)	0.492 *** (0.0503)	16353.147 *** (1665.508)	0.337 *** (0.0350)	5832.356 *** (624.526)
2016	0.6670 *** (0.151)	59907.746 *** (13029.107)	0.106 ** (0.0522)	1954.314 * (1087.348)	0.418 *** (0.0910)	17051.838 *** (4153.0396)	0.328 *** (0.0755)	7877.204 *** (1771.199)

2. 不同行业类型

本文还讨论了不同行业的小微企业应税行为反应。在税收调查数据中,小微企业中服务业占比较大,各细分行业的避税能力不同,隐匿真实交易数据以灵活调整纳税数据的可能性不同(田彬彬等,2022)。由于现金交易更容易隐匿收入,便于调整财务数据获得税收优惠政策。本文对不同服务业小微企业进行讨论,根据现金交易的难易程度将服务业企业分类为生活性服务业和生产性服务业,生活性服务业包括住宿餐饮业、文化娱乐业、零售业、其他居民服务业,属于末端消费,现金交易更加容易;而生产性服务业包括交通运输业、房地产业、建筑设计、租赁和商务服务业以及信息传输、软件和信息技术服务业,其现金交易更困难,企业隐匿收入可能性较低。表 5 第(5)、(7)列展示了 2010—2016 年两类服务业企业的应税反应弹性,第(6)、(8)列展示了由企业群聚响应产生的预算偏差额,结论认为,生活性服务业的群聚程度和应税反应弹性始终强于生产性服务业,说明现金交易越容易的企业发生群聚响应行为的可能性越高,这对于政府部门建立应收尽收、应享尽享的税收环境造成现实挑战。

3. 政策知悉程度不同

《2015 年小微企业调查数据》显示,不同省份的纳税人对减税政策知悉程度不尽相同,而纳税人对于减税政策的知悉程度对其是否产生应税行为反应具有重要影响。极端情况下,企业不知任何减税政策,不会根据减税政策门槛进行群聚响应,反事实分布曲线与真实分布曲线完全重合,此时不会产生预算偏差;如果企业充分知悉并使用减税政策,就会顺应减税政策设定而进行群聚响应,并产生显著的群聚响应现象。因此,考虑到不同地区企业对减税政策的知悉程度不同,本文使

① 使用核定征收方式的企业一直是避税逃税的重灾区,一旦企业满足查账征收的财务水平,税务机关则要求其转变为查账征收方式进行纳税,此转变过程不可逆转,使得减税门槛两侧采用核定征收方式的企业数量下降。

用调查问卷中税收政策知晓程度中位数以上的省份作为“高知悉度地区”,余下省份作为“低知悉度地区”。表6中第(1)、(3)列结果表示应税反应弹性与政策知悉程度相关,即处于高知悉度地区的小微企业,其群聚响应现象更为显著,应税反应弹性更高,会产生更大的预算偏差,而低知悉度地区的小微企业,应税反应弹性系数更低。整体而言,税收政策知悉程度越高则应税反应弹性越高,这表明税务部门在提高纳税人的政策知悉程度之后,纳税人产生了更多的应税行为反应。

表6 异质性:知悉程度与涉税服务不同 单位:千元

年份	知悉程度不同				涉税服务规范性			
	高知悉度地区		低知悉度地区		规范性弱的地区		规范性强的地区	
	弹性	预算偏差	弹性	预算偏差	弹性	预算偏差	弹性	预算偏差
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2010	0.0292 (0.0588)	411.158 (497.387)	0.193*** (0.0365)	1185.252*** (302.870)	\	\	\	\
2012	0.2162*** (0.0344)	2371.649*** (426.999)	0.124*** (0.0288)	1595.918*** (340.599)	\	\	\	\
2014	0.456*** (0.0624)	11631.905*** (1724.608)	0.351*** (0.0351)	8464.0230*** (896.0836)	0.464*** (0.0592)	21830.205*** (3053.434)	0.230*** (0.0807)	267.354*** (91.469)
2015	0.515*** (0.0489)	31995.682*** (3051.623)	0.417*** (0.0402)	19290.074*** (1809.188)	0.443*** (0.0435)	37404.504*** (3643.816)	0.3678*** (0.0331)	1849.440*** (183.235)
2016	0.567*** (0.138)	28271.125*** (6663.990)	0.377*** (0.0867)	23911.0900*** (5213.0977)	0.562*** (0.141)	57490.227*** (13785.006)	0.220*** (0.0594)	2195.591*** (527.804)

注:因2010年、2012年末有相应分组,无法进行统计,故上表中使用“\”表示空缺值。

4. 涉税服务规范性不同

随着依法治税的全面推进,以涉税服务为主营业务的代理记账行业成为促进小微企业依法纳税和获得减税政策优惠的重要渠道。近年来,越来越多的小微企业选择代理记账获取税务代理服务,造成从事代理记账的企业数量不断增长,其地区级行业协会也随之成立。当某地区成立代理记账协会,一方面表明该地区从事专业涉税服务的机构较多,小微企业更容易获得涉税服务,提升小微企业财务资料的合规性,有助于小微企业提高纳税遵从度并获得减税政策优惠(Alm et al., 2010);另一方面,该行业协会将对本地从事涉税服务的机构进行监督约束,并根据财政部令第80号《代理记账管理办法》进行规范管理,对行业内的企业虚假、违规记账行为加以处罚,这将有效保障纳税资料的真实完整性并且减少纳税人的应税行为反应。因此,本文根据各地代理记账行业协会成立的时间,将行业协会已经成立的地区作为涉税服务强规范的地区;若行业协会未成立,则该地区作为涉税服务弱规范地区。表6的第(5)、(7)列实证结果表明,涉税服务弱规范地区的小微企业应税反应弹性显著高于涉税服务强规范的地区,即对涉税服务进行规范监督能够显著减少企业应税行为反应,促进小微企业依法合规纳税,降低企业群聚响应程度这也表明增强涉税服务行业规范性将有效增加企业的纳税遵从度。^①

(二)群聚响应的实现路径

群聚现象是指在恰低于政策门槛值处的企业数量异常上升,而高于门槛值处的企业数量减少,

^① 作者手动收集代理记账协会成立地区的相关数据,发现在2013年以前,仅有成都市、南通市、海安县等少量地区成立代理记账协会,其分组样本量无法满足群聚方法要求的大数据样本要求。因此,对于已经成立代理记账协会的地区,本文仅保留2014—2016年的实证结果作为对比。

在前述中已经表明了这一群聚现象对政府预算管理产生的影响,但未讨论群聚现象产生的路径,对此,本文进一步考虑了企业经营活动变化对应纳税所得额的影响。根据 Chen et al. (2021) 提出的结构性断点 (structural break) 的研究方法,再考虑到不同年份群聚的企业并非同一类企业,本文结合图形和结构性断点估计的统计推断群聚企业在门槛值两侧的操纵行为。从实证分析和稳健性测试结果来看,门槛值两侧的小微企业在上一年度的财务指标均未发生显著的跳跃,仅在政策实施的当年,门槛值左侧的小微企业利润出现了下降,可见,小微企业主要通过提高主营成本率、降低费用率的方式,调减营业利润和利润总额,最终减少了应纳税所得额。因此,本文认为发生群聚响应的企业在不进行收入调整的情况下,主要通过营业成本调整,而非调高费用比重方式实现群聚响应行为。^①

六、结论与政策启示

当前,我国预算管理制度改革仍处于深水期,将减税政策与现代化预算管理制度相结合,充分发挥预算制度的管理职能,将有利于加速推进减税政策“控成本、优工具、精管理”,提升政府治理效能。在三期压力叠加背景下,政府对财政部门的资源配置能力、预算编制和预算执行能力提出更高的要求,而减税政策的预算管理通过收支透明化、政策实施精准化可以提高政策有效性,为减税政策效能指明路径。本文通过群聚方法捕捉企业应税反应弹性并对减税政策实施规模进行事前预判,补充了群聚方法应用于动态减税政策评估的研究,可以通过厘清减税政策受益对象的结构和金额变化,践行“花钱必问效”的预算管理理念;将减税政策精准化管理由事后续效评估向事前预算评估转移,可以提升预算编制的可靠性与精准性,助力实现预算管理的监督约束职能,即通过预算管理的优化最终提升减税政策效能。

本文通过群聚方法识别出小微企业在门槛性减税政策实施后,其群聚程度、应税反应弹性变化及其预算影响,为预判减税政策实施结构和金额提供可靠依据。本文研究发现在减半征收政策实施后,小微企业的群聚程度、应税反应弹性、预算偏差均随减税政策扩围而持续上升,且直接估计的企业群聚响应对预算偏离具有一定的解释程度。接着,研究发现减税门槛消失后,企业会产生群聚反弹并引起税基分布变化,综合考虑群聚响应、群聚反弹因素可以更好地解释预算偏离,为提升减税政策的预算精准化提供了现实路径;进一步研究发现,查账征收方式、生活性服务业、政策知悉程度高、涉税服务规范性强的小微企业群聚程度、应税反应弹性更高,此外,本文还讨论了小微企业群聚响应的实现路径,为减税政策和预算治理效能提升提供研究参考。最后,本文提出将群聚因素纳入减税预算管理的量化设计,有助于更加全面的理解减税政策效能实现的内在机理和优化治理手段,为减税降费政策提质增效提供理论支撑与实证依据。

基于以上研究结论,本文提出以下三点政策建议。

第一,坚持数字税收引领,营造良好税收环境。首先,依托金税三期、四期工程,税务部门应当充分利用税收大数据,对小微企业加强税收监管,逐步实现“以数治税”的税收征管模式。其次,政府应对于获得减税政策优惠的纳税人加大审查力度,鼓励企业诚信纳税、诚信减税,加强税务稽查力度,减少纳税人逃避税的空间。最后,应当增强涉税服务的可得性,打造应收尽收、应享尽享的绿色税收环境,倡导企业提高纳税遵从度,并且加速推进《税收征管法》修改,在法治轨道上推进税收现代化行稳致远。

第二,优化减税政策设计,减少应税行为反应。首先,在减税政策工具设计上应当更加精准,充分考虑纳税人对税费政策进行策略性响应的复杂性和动态性,更要将纳税人应税行为反应纳入政

^① 因篇幅所限,相关内容详见本文附录 6。

策设计,避免“一刀切”的减税政策,谨防因税率跳跃上升所导致的纳税人相机抉择,可考虑使用累进式减税工具,缩短不同应税收入区间内纳税人的实际税负差距,减少纳税人的应税行为反应。其次,应进一步降低小微企业的纳税成本,进一步规范代理记账行业服务内容、收费标准与监管制度,及时为小微企业提供财务及纳税辅助,增强纳税人对减税降费政策的知悉程度。最后,可以考虑对于不同行业、不同征收方式、不同地区的企业进行差别化政策宣传和纳税辅导,提升减税降费政策的享受率。

第三,健全预算管理制度,提升减税政策效能。首先,从预算执行精准化的角度,厘清减税政策的受益范围和金额,对减税政策进行事前预算、事中监督、事后评估,实现减税政策的预算监督效能。其次,从预算资金管理的角度,探索建立税式支出预算制度,增强公众监督,实现规范减税政策实施过程管理与结果管理。与此同时,税务部门应重点关注减税政策实施状况,优化组合式税费支持政策工具,将减税资金花到“刀刃”上,实现减税政策的预算约束效能。最后,将现代预算制度与减税政策协调配合纳入宏观治理体系,重点关注减税政策的执行率、超预算率等,助力减税政策的绩效管理及保障财政可持续性,提升预算的资源统筹效能。如此,才能多位一体提升减税政策效能。

参考文献

白重恩、毛捷,2011:《公共财政视角下的税式支出管理与预算体制改革》,《中国财政》第2期。

樊丽明、史晓琴、石绍宾,2022:《我国地方人大预算监督评价理论、指标及应用》,《管理世界》第2期。

范柏乃、林哲杨,2022:《政府治理的“法治一效能”张力及其化解》,《中国社会科学》第2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。

甘犁、秦芳、吴雨,2019:《小微企业增值税起征点提高实施效果评估——来自中国小微企业调查(CMES)数据的分析》,《管理世界》第11期。

高培勇,2023:《论健全现代预算制度的基础工程》,《中国工业经济》第1期。

高培勇、毛捷,2013:《间接税税收优惠的规模、结构和效益:来自全国税收调查的经验证据》,《中国工业经济》第12期。

高培勇、汪德华,2016:《本轮财税体制改革进程评估 2013. 11—2016. 10(下)》,《财贸经济》第12期。

郭庆旺,2019:《减税降费的潜在财政影响与风险防范》,《管理世界》第6期。

国家税务总局济南市税务局课题组,2020:《减税降费背景下税式支出纳入地方政府预算的研究——基于国际借鉴与山东实践的加强税式支出治理的调研》,《财政监督》第3期。

何文盛、王焱、蔡明君,2013:《政府绩效评估结果偏差探析:基于一种三维视角》,《中国行政管理》第1期。

贾康、段爱群,2013:《预算法修改中的创新突破与问题评析——关于〈预算法〉修改的意见和建议》,《财政研究》第6期。

李昊楠,2021:《减税效率的提升路径——来自结构性减税时期小微企业应税收入弹性的证据》,《管理世界》第11期。

刘啟仁、龙健雄、张展辉、赵灿,2023:《税收激励、研发支出与出口绩效——基于高新技术企业认定条件改革的聚束分析》,《中国工业经济》第4期。

吕炜、王伟同,2022:《党的十八大以来财政领域改革成就、内在逻辑与未来展望》,《财政研究》第9期。

田彬彬、杨健鹏、汪丹、叶菁菁,2021:《第三方信息获取与税收征管效率:来自有奖发票推行的证据》,《世界经济》第9期。

汪德华、任永美、周文,2014:《税式支出核算方法的国际经验与启示》,《国际税收》第10期。

王伟同、李秀华、陆毅,2020:《减税激励与企业债务负担——来自小微企业所得税减半征收政策的证据》,《经济研究》第8期。

吴建南、马亮、杨宇谦,2011:《比较视角下的效能建设:绩效改进、创新与服务型政府》,《中国行政管理》第3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。

张航、范子英,2019:《群聚分析法原理、争议及应用前景》,《数量经济技术经济研究》第9期。

郑方辉、费睿,2019:《财政收入绩效评价兑现减税降费政策目标的价值工具》,《中国社会科学》第6期。

朱武祥、张平、李鹏飞、王子阳,2020:《疫情冲击下中小微企业困境与政策效率提升——基于两次全国问卷调查的分析》,《管理世界》第4期。

Alm, J., T. Cherry, M. Jones, and M. McKee, 2010, “Taxpayer Information Assistance Services and Tax Compliance Behavior”, *Journal of Economic Psychology*, 31(4), 577—586.

Bastani, S., and H. Selin, 2014, “Bunching and Non-bunching at Kink Points of the Swedish Tax Schedule”, *Journal of Public Economics*, 109, 36—49.

- Bergolo, M., G. Burdin, M. De Rosa, M. Giacobasso, and M. Leites, 2021, “Digging into the Channels of Bunching: Evidence from the Uruguayan Income Tax”, *Economic Journal*, 131(639), 2726—2762.
- Best, M. C., A. Brockmeyer, H. J. Kleven, J. Spinnewijn, and M. Waseem, 2015, “Production Versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan”, *Journal of Political Economy*, 123(6), 1311—1355.
- Bogacheva, O. V., and T. V. Fokina, 2017, “Evaluation of Social Tax Expenditures Efficiency in OECD Countries”, *Financial Research Institute*, 3, 22—36.
- Bosch, N., V. Dekker, and K. Strohmaier, 2020, “A Data-driven Procedure to Determine the Bunching Window: An Application to the Netherlands”, *International Tax and Public Finance*, 27(4), 951—979.
- Chen, Z., Z. Liu, J. C. Suarez Serrato, and D. Y. Xu, 2021, “Notching R&D Investment with Corporate Income Tax Cuts in China”, *American Economic Review*, 111(7), 2065—2100.
- Chetty, R., 2012, “Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply”, *Econometrica*, 80(3), 969—1018.
- Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda, 2014, “The Role of Entrepreneurship in US Job Creation and Economic Dynamism”, *Journal of Economic Perspectives*, 28(3), 3—24.
- Devereux, M., L. Liu, and S. Loretz, 2014, “The Elasticity of Corporate Taxable Income: New Evidence from UK Tax Records”, *American Economic Journal: Economic Policy*, 6(2), 19—53.
- Diamond, R., and P. Persson, 2016, “The Long-term Consequences of Teacher Discretion In Grading of High-stakes Tests”, NBER Working Paper.
- Feldman, E. N., P. Katusčák, and L. Kawano., 2016, “Taxpayer Confusion: Evidence from the Child Tax Credit”, *American Economic Review*, 106(3), 807—835.
- Gelber, A. M., D. Jones, and D. W. Sacks, 2020, “Estimating Adjustment Frictions Using Nonlinear Budget Sets: Method and Evidence from the Earnings Test”, *American Economic Journal: Applied Economics*, 12(1), 1—31.
- Harju, J., T. Matikka, and T. Rauhanen, 2019, “Compliance Costs Vs. Tax Incentives: Why do Entrepreneurs Respond to Size-based Regulations?”, *Journal of Public Economics*, 173, 139—164.
- Kleven, H. J., 2016, “Bunching”, *Annual Review of Economics* (8), 435—464.
- Kleven, H. J., and M. Waseem, 2013, “Using Notches to Uncover Optimization Frictions and Structural Elasticities :Theory and Evidence from Pakistan”, *Quarterly Journal of Economics*, 128(2), 669—723.
- Kopczuk, W., and D. Munroe, 2015, “Mansion Tax: the Effect of Transfer Taxes on the Residential Real Estate Market”, *American Economic Journal: Economic Policy*, 7(2), 214—257.
- Mortenson, J. A., and A. Whitten, 2020, “Bunching to Maximize Tax Credits: Evidence from Kinks in the US Tax Schedule”, *American Economic Journal: Economic Policy*, 12(3), 402—32.
- OECD, 2010, “Tax Expenditures In OECD Countries”.
- Saez, E., 2001, “Using Elasticities to Derive Optimal Income Tax Rates”, *The Review of Economic Studies*, 68(1), 205—229.
- Saito, K., and D. Tsuruta, 2018, “Information Asymmetry in Small and Medium Enterprise Credit Guarantee Schemes: Evidence from Japan”, *Applied Economics*, 50(22), 2469—2485.

Increasing Tax Reduction Policy Effectiveness by Budgeting: Evidence from Dynamic Bunching of Micro and Small Enterprises

LIU Rong^a, CHEN Lingqing^a, WANG Ruiting^a and DENG Xinghua^b

(a: School of Public Finance and Taxation, Southwestern University of Finance and Economics;

b: School of International Business, Southwestern University of Finance and Economics)

Summary: Modern budget management primarily aims at controlling government revenue and expenditure through fiscal constraints, which ultimately enhances government effectiveness. A crucial component of this structure, the large-scale tax and fee reduction policies serve as a counter-cyclical adjustment tool and introduce considerable challenges to traditional budget management practices. Empirical data reveal that during the implementation of these policies, the actual tax reduction often surpasses the initial estimates. This discrepancy between the projected and realized figures can undermine the sustainability and precision of budget management, thereby impeding the accurate execution of tax reduction policies. Consequently, our research delves into this discrepancy, with a focus on taxpayers' behavioral responses to tax reduction thresholds. Using a dynamic bunching approach, we anticipate enterprises' behavioral adaptations following the implementation of tax reduction policies, with the ultimate goal of minimizing budget deviations and enhancing the precision of budget management in the context of tax reduction policies.

This study draws on the administrative enterprise income tax records from the surveys of the State Taxation Administration of China (STAC), collected between 2008 and 2016. Our analysis reveals a substantial bunching phenomenon among micro and small enterprises (MSEs) following the implementation of a half-taxation policy. This notable bunching response is further explored using a theoretical model to shed light on the underlying motivations of such bunching of enterprises and their impacts on budgetary governance.

The research proposes that both the initial bunching response and subsequent rebound significantly influence the tax reduction budget. We, by using the frequency distribution of enterprises unaffected by tax reduction policies, construct a counterfactual distribution and validate our hypotheses. Our findings show that the bunching mass significantly alters tax base distribution, and a stronger bunching response correlates with higher tax reduction thresholds, leading to larger budget deviations. Upon the expansion of the tax reduction policies, this paper observes that the initial bunching diminishes and shifts towards decreased taxable income or "missing mass." This shift further increases the tax base within the new cut range, providing insights into budget deviations. Investigation into the heterogeneity of bunching responses, considering different collection methods, industry types, tax policy knowledge levels, and tax-related services, underscores economic behaviors' complexity, requiring a subtle policy implementation approach. Our comprehensive policy recommendations aim to optimize budget management precision of tax reduction policies, enhancing policy effectiveness. Our theoretical and empirical contributions significantly advance the understanding of fiscal policies and budget management.

This study offers several key contributions. Firstly, we integrate the bunching approach within the macro-fiscal governance framework, revealing budget deviations in tax reduction policies and providing empirical evidence to support their integration into a comprehensive budget management system. Secondly, we use the bunching approach to anticipate enterprise behavior, promoting a transition from ex-post to ex-ante evaluations of tax reduction policies and enriching ex-ante research scenarios. Thirdly, we analyze taxpayers' behavioral responses to threshold-type tax cuts and the subsequent behavioral changes after policy implementation, providing first-time insights into the impact of such changes on the tax base following multiple policy expansions. It also applies the bunching approach for dynamic policy evaluations. Fourthly, we assess the implementation of tax reduction policies for MSEs, exploring the mechanisms underlying bunching responses, and thereby contributing to improving the effectiveness of tax reduction policies.

Keywords: Tax Reduction Policy; Budget Management; Micro and Small Enterprises; Bunching Approach

JEL Classification: H20, H32, H61

(责任编辑:陈小亮)(校对:王红梅)