零基预算：从方法到理念演进的要件分析

Zero-based Budgeting: Analysis of the Important Elements among Its Evolution from Method to Budgeting Concept

李成威 杜崇珊

LI Cheng-wei DU Chong-shan

［摘 要］从20世纪70年代起，美国等国和我国一些地方政府尝试开展零基预算实践，但由于预算编制程序复杂、工作量及技术难度大、预算监督力度不足等原因，国内外对零基预算的探索并未取得理想的效果。之所以出现上述情况，是因为他们都将零基预算作为一种预算方法而非预算理念加以运用。这种做法存在难以克服的缺陷，在现实中必然碰壁。而如果将零基预算作为一种预算理念，则其具有强大的生命力。2008金融危机后OECD国家带有零基预算性质的支出审查得以迅速推广并取得明显成效就证明了这一点。在零基预算从方法到理念的演进过程中，其繁琐的编制程序被舍弃，但其理念和核心策略却通过支出审查、绩效预算、以风险为导向、聚焦关键部门、实施程序控制的方式融入预算环节，从而实现了既定预算目标。我国应借鉴OECD国家的经验，运用零基预算理念实施以支出审查为核心的措施，并在实践中推动相关预算改革，以促进财政支出结构的全面优化。

［关键词］零基预算理念 支出审查 优化支出结构

［中图分类号］F810.2 ［文献标识码］A ［文章编号］1000-1549（2020）10-0003-07

Abstract: Since the 1970s, countries like US and some local governments in China have tried to apply zero-based budgeting into practice. However, due to the complicated budgeting procedure, heavy workload, technical difficulties, and insufficient budget supervision, the exploration at home and abroad has not achieved ideal results. The main reason for their failure is that they all treat zero-based budgeting as a budgeting method rather than a budgeting concept. Thus, their practice is trapped in the defects of such a method and bound to fail. However, if zero-based budgeting is applied as a budgeting concept, it will show great vitality. Spending review is a method containing the concept of zero-based budgeting and has been widely and successfully adopted by OECD member countries after the financial crisis. This is a testament to the idea that zero-based budgeting should be applied as a budgeting concept. In the evolution from a method to a budgeting concept, zero-based budgeting's cumbersome procedures are abandoned, but its concept and core strategies are integrated into the budgeting process by means of spending reviews, performance budgeting, risk-orientation, focusing on key departments and implementing procedure control. And in this way, the established budgeting targets are achieved. China should learn from the experience of OECD countries, apply the zero-based budgeting concept by implementing measures with spending reviews as the core, and promote relevant budget reform, so as to promote the overall optimization of fiscal expenditure structures.

Key words: The zero-based budgeting concept Spending review Expenditure structure optimization

［收稿日期］ 2020-05-11

［作者简介］ 李成威，男，1977年11月生，中国财政科学研究院研究员，研究方向为财政基础理论、大国财政；杜崇珊，女，1989年2月生，中国财政科学研究院博士研究生，研究方向为国际财政。

［基金项目］ 国家社会科学基金重大项目“国家治理现代化视角下的财税体制研究”（项目编号：17VZL011）。

感谢匿名评审人提出的修改建议，笔者已做了相应修改，本文文责自负。

一、引言

我国财政规模在改革开放后实现了高速增长，其中 1990—2012 年间财政收入的平均增速为 18%，快于同期 GDP（马骏，2013[1]）。但随着经济进入新常态，财政形势日益严峻，中央和地方政府优化支出结构、提高资金使用效率的诉求也日益迫切。2013—2019 年，全国一般公共预算收入增速从 10.2%减缓至 3.8%，而支出除 2016 年外增速均在 7%以上（详见图 1），财政预算环境发生了极大的变化[1]。尤其是在新冠疫情爆发之后，收支矛盾更加突出。2020 年一季度全国一般公共预算收入下降 14.3%，其中税收下降 16.4%[2]，且预算的大部分已被部门的惯性支出“锁定”，少有活钱。在此背景下，零基预算以其全面优化支出的理念受到了中央和地方的广泛关注，《国民经济和社会发展报告》《预算报告》都强调要“运用零基预算理念，打破支出固化格局”；山东省将零基预算写入政府工作报告；上海市在其预算报告中提出要落实零基预算理念[3]；江西省表示要在省直行政机关全面实行零基预算[4]。各地都希望通过零基预算方式打破对预算安排中“基数”的依赖，把有限的资金管住用好。

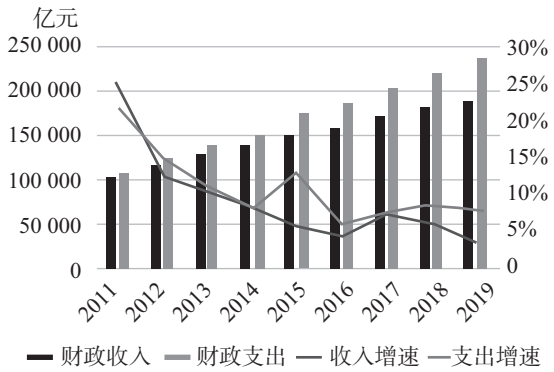

图 1　我国一般公共预算的收支矛盾

数据来源：国家统计局（2019）[5]、财政部（2020）[6]。

如何决定财政资金在各项目间的分配是预算的一个基本问题。在理论上，学者们对该问题展开了广泛的探讨，但并未得出令人满意的结论（牛美丽，2003[7]）。在实践上，从 20 世纪 30 年代起，美国等主要发达国家以支出控制、理性决策为目的，进行了一系列的理性预算改革，其中就包括零基预算（罗宏斌等，2006[8]）。零基预算作为一种预算方式诞生于 20 世纪 70 年代大萧条后的美国。它以提升资源配置效率为目的，并且更加注重预算的控制和管理功能（Austin 等，1979[9]）。然而，由于程序的繁琐性以及政治和法律上的难题，该预算方法在 20 世纪 80 年代后期逐步走入低谷（聂丽洁等，2004[10]）。

我国在 20 世纪 90 年代开启了对零基预算的试点探索，当时它被视为部门预算编制的最佳模式（肖育才，2007[11]）。但与美国的早期实践类似，各地多将其作为一种具体的预算方法来使用，并且同样因为政治体制、技术等因素的制约，国内的探索也未能取得理想效果。截至目前，该预算方法在我国已有近 30 年的实践历史：1993 年河北等省份率先开启零基预算试点；1995 年《中华人民共和国预算法实施条例》确立了该方法的法律地位；2000 年中央部门预算正式采用该方法；当前多数省市都在积极推行零基预算，但尚未形成系统的编制方法（吕彩凤，2009[12]）。其应用范围也十分有限，主要用于专项支出，人员、公用经费等经常支出以及社保等权益支出一般不予采用，而实际上后几项开支占用了大部分预算（王辰，2011[13]）。零基预算方法在推广和实施过程中存在较多难点，主要包括政府职能划分不清晰导致难以确定决策主体（聂丽洁等，2004[10]）；预算编制程序复杂，工作量及技术难度大；预算公开不足，监督力度仍需加强；等等。另外，由于预算透明度不高，零基预算的实践效果究竟如何外界很难知晓。根据马骏（2005）[14]对某省的调查，尽管该省在 1993 年就开始推行零基预算，但其效果也仅限于精简项目的数量，对预算的改进则非常有限。零基预算变成了财政部门控制支出的工具而非改进资源配置效率的方式。综上所述，由于实践中的诸多困难，零基预算作为一种具体的预算方式很难在我国真正展开。

虽然零基预算作为一种预算方法已宣告失败，但其理念和核心策略却对后续的预算改革产生了深远的影响。在金融危机后的财政紧缩环境下，OECD 国家对零基预算进行了新的探索。它们舍弃了零基预算的具体形式，通过支出审查、绩效预算、风险导向、聚焦关键部门、程序控制等方式将其理念和核心策略融入预算程序，在削减财政支出、优化支出结构方面取得了很好的效果。从实践的成败经验可知，零基预算应被视为一种基于风险的预算理念而非具体预算方法，其最大优势是提供了一种全面优化支出结构的路径。基于此，我国政府应立足国情并借鉴国际经验，灵活运用零基预算理念，以此实现完善预算编制制度、缓和收支矛盾等目标。

二、零基预算：从方法到理念的演进

零基预算于 20 世纪 70 年代作为一种预算流程在美国等国得到较多的推广。从“零”开始制定预算意味着不再将过去的支出模式视为既定，但这种优化支出结构的预算方法并未充分考虑到实际执行中的难度，从而导致其实践效果远低于预期。虽然零基预算作为一种过度理想化的预算方法最终被抛弃，但其审视既有支出项目、对支出进行优先级排序的理念和核心策略却一直影响着预算改革的进程。

（一）作为预算方法的零基预算在现实中碰壁

自 20 世纪 30 年代起，美国进行了多次理性预算改革，70 年代出现的零基预算就是其中的一种。它从私人部门发展而来，并最早于 1973 年由时任州长卡特引入佐治亚州的政府部门，又在其当选总统后被引入联邦政府。此后，其他国家也陆续在改革中借鉴了该模式。

这种理性决策的预算模式采用了自下而上的方式，希冀以此融合基层部门的信息优势和决策部门的整体战略部署，从而让各项支出与政府的优先级相契合，达到优化支出结构、提高支出效率、削减预算规模的目标。其实施过程涉及多个层级的多个主体，按照程序先后可划分为预算编制单元、整合单元和决策单元。预算编制单元需要根据其职责和目标的优先顺序提供项目材料并对项目排序，形成基础决策包后交由整合单元统一整合。整合单元需根据与预算编制单元的沟通和自身对整体目标的理解将各预算编制基础单元的项目内容整合、排序，然后呈报给决策单元。决策单元则需根据整体目标对呈报上来的项目排序进行调整，并决定资金的分配。

然而，作为一种预算方法，零基预算缺乏对实际情况的考虑。理想与现实的差距是导致其难以执行的重要原因。抛开政治和法律因素，仅技术上的困难就已极大地限制了该模式的推广范围和实践效果。例如，在预算编制单元的确定方面，划分过细易造成预算的碎片化、重复化，从而致使工作量明显增加而效率却显著下降；划分过粗则会导致真正了解实际情况的人员无法参与其中，从而造成预算编制脱离实际，进而难以实现优化资源配置的目标。在项目排序方面，基层管理者可能会根据部门甚至个人利益安排项目优先顺序，而高层管理者对项目的重新排序也难免存在一定的任意性。此外，过大的工作量一直是其严重的问题，如 1979 年美国联邦政府实行零基预算时，产生了约两万五千个决策包（马骏和叶娟丽，2004[15]）。

这样的缺陷导致零基预算在执行过程中变形走样然后被抛弃。Schick（1978）[16] 等学者认为零基预算只是被嫁接到原有的渐进式预算过程中，并没有改变决策规则。且其倡导者卡特总统在 1979 年呈递的预算案与传统方法所形成的预算并无显著差别。而里根总统上任之后，零基预算被彻底搁置，此后佐治亚州也一度停用该预算方式。

（二）作为预算理念的零基预算影响深远

尽管作为预算方法的零基预算在美国等西方国家已基本消失，但它所倡导的理念并没有被遗忘。零基预算要求政府每年都为它的全部预算进行辩护，即前一个预算年度中各部门获得的拨款不再理所当然地成为下个年度预算决策的基础，其初衷是为了找到一种能够在财力紧张情况下有效配置资源的方法。这种有效配置通过定期（如每个财年）对支出项目的重新排序来实现，而排序又需要建立在较为完善的项目审查和评估基础上。由此可知，对支出内容的定期评估和重新排序是零基预算实现支出结构优化目标的核心策略，也是其影响深远的主要理念。

金融危机后，随着收支矛盾的日益激化，一些国家对零基预算理念进行了新的探索。最近十多年来，具有零基预算理念的带有削减支出功能的预算方式明显增多。除了美国佐治亚州等州政府对相关理念进行的创新式运用外，英国等 OECD 国家在中央政府层面也通过支出审查等方式将零基预算理念应用于实践，且取得了较好的成效。

三、金融危机后 OECD 国家运用零基预算理念的成功实践

在金融危机后的财政紧缩环境下，OECD 国家对零基预算进行了新的探索，它们摒弃了零基预算过于繁琐的预算形式，保留了其内在理念和核心策略，并通过支出审查等形式，实现了优化支出结构的目的。

（一）受财政赤字和债务困扰，带有零基预算性质的支出审查在 OECD 国家迅速推广

2008 年金融危机之后，由于许多成员国深受赤字和债务的困扰，OECD 开始推广支出审查并得到快速响应。支出审查与零基预算的理念和核心策略相契合，强调对基准支出进行系统和深入的审查，以发掘并削减无效和非优先支出，进而节约预算资金、提升

支出效率。根据 OECD 的调查，2011 年开展支出审查的成员国有 16 个，而至 2018 年，当时的 35 个成员国中已有 27 个采用了该模式，还有 3 个在考虑中（详见表 1）。各国政府利用支出审查工具整合、削减预算，并调整支出以使其与关键政策优先次序保持一致，从而达到扩大政府可用财政空间、为重点政策提供资金的目的。

表 1　　2011—2018 年采用支出审查的 OECD 成员国数量

	已采用的国家数	考虑采用的国家数	未采用的国家数
2011 年	16	0	16
2016 年	23	5	4
2018 年	27	3	3

注：2018 年美国和以色列的数据未搜集到。
资料来源：OECD（2019）[17]。

由于预算制度和国情大不相同，各国对支出审查的具体运用方式也存在很大差别，但总结来看主要有两种模式：一是有针对性的年度审查，二是周期性的全面审查（OECD，2016[18]）。前者一般依照事前制定的主题列表展开，而后者则不受任何此类事先列出的审查主题的限制。就当前的实践来看，丹麦、荷兰等选择的是前一种模式，即每年对选定的支出方案进行有针对性的审查，并调整有限范围内的支出优先次序，以提高支出和行政效率。而英国则采用了每隔几年进行一次的全面审查，即根据政策重点和财政目标在整个政府范围内调整预算，以将资源重新分配给优先项目。通常，这样的审查会根据选举周期来安排，以将经济和政策环境的变化考虑在内。

随着开展支出审查的 OECD 成员国不断增加，其内部呈现出两种趋势。一是审查范围不断扩大。调查数据显示，越来越多的国家转向了范围更广的综合性审查，其中 2008 年为 5 个国家，而至 2015 年曾一度增加到 8 个国家；二是审查频率逐步提高，例如爱尔兰已从 2009—2014 年预算整顿和恢复期间的定期全面审查过渡到纳入年度预算程序的滚动审查，加拿大的支出审查也更为常规化（OECD，2019[17](114)），而且截至 2018 年采用频率相对更高的年度审查的国家已经达到了 17 个（详见表 2）。

（二）在一些 OECD 国家，支出审查已成为资源重新分配和支出规划的基础

无论是年度、定期、全面还是有针对性地小规模开展，各种形式的支出审查都已较为成功地融入了 OECD 成员国的预算编制过程。对丹麦、瑞典、英国

表 2　　2018 年 OECD 国家中支出审查的采用情况

已采用的国家		未采用的国家	
年度审查	定期审查	考虑中	暂未考虑
瑞典、斯洛伐克、葡萄牙、波兰、挪威、荷兰、拉脱维亚、日本、意大利、爱尔兰、匈牙利、希腊、爱沙尼亚、丹麦、德国、瑞士、奥地利	澳大利亚、加拿大、西班牙、芬兰、法国、英国、卢森堡、墨西哥、新西兰、斯洛文尼亚	冰岛、韩国、土耳其	比利时、智利、捷克

注：2018 年美国和以色列的数据未搜集到。
资料来源：OECD（2019）[17](113)。

等国来说，支出审查已成为财政资源重新分配和支出规划的基础，并为预算谈判、绩效评价和监督等提供了信息。

丹麦和瑞典是开展年度审查的典型国家。其中，丹麦已将支出审查纳入年度预算程序，其财政部通常每年都会确定几项审查任务，范围覆盖警察和后勤部门等。目前该国已开展了超过 50 次的审查。而在瑞典，支出审查主要集中在资本效率（如高速公路和铁路）和服务提供（如移民机构）方面（Assi，2019[19]）。

英国是开展定期全面审查的代表性国家。其支出审查类似于中长期财政规划，每 2 到 5 年进行一次。届时财政部会根据政府项目优先次序开展审查工作，并据此为每个部门设定规定年度（支出审查年限）内的最高支出限额，以确保财政资金的可持续性和成本效益的最优化。之后，在每年的预算审核及拨款评估中，部门支出限额可能会做出调整以反映公共风险及政策优先级的变化。在支出限额的框架下，主要政府部门将制定“单一部门计划”，其中包括工作计划及用于反映年度表现的业绩指标等。对于涉及多部门的领域，计划中需划定每个部门的责任范围，以及与其他部门的联合责任。每年上述计划都会根据部门职责或优先事项的变化进行修订。为强化监督，各部门还会将该计划细化，以便部门领导进行管理，且其执行情况也会体现在财年末的年度报告中。此外，政府网站还会定期发布用于绩效监控的官方统计数据等，以增加财政绩效信息的透明度。

（三）OECD 国家实施支出审查的效果明显

2018 年 OECD 的调查显示，其成员国已将支出审查视为进行战略预算的有效工具。其中，绝大多数国家表示已基本实现了支出审查的目标，只有极少数

国家表示未实现相关目标。但各国在效果评估方面存在差异，有一些国家并未对预算结果进行评估，这导致上述调查难以反映支出审查效果的全貌。但根据已有数据可知，支出审查在改善中长期支出效率、削减开支、重新分配支出以契合政策重点这三方面具有较为突出的成效。在 27 个已实施支出审查的国家中有 12 个国家在调查中表示基本实现了改善支出效率的目标，11 个国家表示基本甚至完全实现了支出的削减，15 个国家表示基本甚至完全实现了资源的重新配置（详见表 3）。

从具体国家的实践效果来看，丹麦利用 5 年时间对跨部门的项目进行了支出审查，每年节省资金约 10 亿欧元（Assi，2019[19]）。英国通过对超过 200 亿英镑的年度支出项目进行审查，提高了 25 个部门的支出透明度，并实现了对资金分配的严格把控，从而将支出总额控制在政府所能负担的范围之内。而且即使经历了金融危机的考验，2000—2015 年间，与欧洲许多国家相比，英国的实际支出相较于计划也未发生过大的偏离（详见图 2）。

表 3　2018 年 OECD 成员国对支出审查有效性的报告

目标实现情况	国家数量			
	项目削减	支出效率的中长期改善	削减开支	重新分配支出以契合政策重点
完全实现目标	0	0	2	1
基本实现目标	3	12	9	14
基本没有实现目标	4	0	2	1
未评估结果	18	14	12	10

注：2018 年美国和以色列的数据未搜集到。

资料来源：OECD（2019）[17](115)。

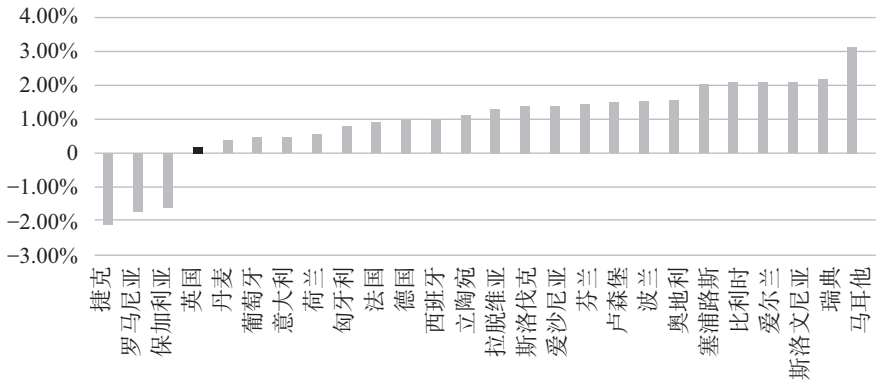

图 2　2000—2015 年欧洲公布支出计划的国家中其实际支出与计划偏差的 GDP 占比

注：数据显示了在支出计划公布后两年内实际支出与计划偏差的 GDP 占比。与其他大多数国家相比，英国的实际支出严格遵循计划安排。

资料来源：Wheatley 等（2018）[20]。

四、运用零基预算理念的要件分析

从各国实践经验来看，零基预算作为一种理想的预算模式败给了执行过程中的繁琐程序和巨大成本，而当前 OECD 国家实施的零基预算更多的是一种基于风险的预算理念，它依托于其他具体的预算形式而存在并发挥作用。这种模式有以下要件：

（一）支出审查：将零基预算的理念付诸现实

支出审查将零基预算所倡导的支出优先性这一核心策略与现实执行进行了较好的融合，从而能够有效地协助政府将有限的资源重新分配给社会效益大和公共风险较为集中的项目。

OECD 国家在中央政府层面运用了支出审查方式并较好地实现了优化支出结构的目标。例如，英国财政部根据政府关键政策的优先次序对支出进行审查，并据此为各部门设定规定年度（支出审查年限）内可支出的最高限额。在此基础上，各部门负责组织、分配和管理其资源。该程序相当于确定了一个支出框架，并在该框架下鼓励各部门在做出支出决策时将未来的公共风险和资金的时间价值考虑在内，防止短视行为，从而保证了财政资金的可持续性和成本效益。

支出审查的使用不仅限于中央政府层面，美国佐治亚州的新零基预算也特别强调项目的审查。这种新的零基预算方式于 2012 年成为州法，并将在 2020 年终止生效（除非重新授权）。它要求州众议院预算办公室、参议院预算评估办公室和州政府预算规划办公室一起，每年抽查年度预算一定比例的项目（约为当年项目总数的 10%）。这种抽查的方式在减少工作量的同时也实现了支出审查的主要目标[21]。

（二）绩效预算：为零基预算从理念到实践提供了平台

早期的零基预算并没有绩效预算作为基础，而从

当前的实践来看，零基预算理念都是嫁接于绩效预算之上的。OECD 国家普遍采用的支出审查实际上就是绩效预算的重要组成部分。从二者的协同作用来讲，支出审查的结果是预算编制的重要考虑因素，而绩效预算所提供的信息（包括对方案的定义、绩效指标和目标的确定等）对开展支出审查、锁定低效率和价值不大方案也是非常有帮助的。另外，考虑到政治等因素，有效的支出审查需要足够的证据支持，以压倒来自利益相关部门和组织的反对，达到减少或重新分配支出的目的，而绩效数据就是其有力的支撑。二者的联合使用更利于审视现有支出方案的目的和价值，并可通过对原有项目的重新排序和精简来为新的优先支出事项腾出财政空间。

（三）风险导向：为支出审查提供了标准和依据

零基预算理念要求对支出价值进行判别，这就需要设立判断标准，即为何否定一项支出，又如何确定一项支出。在 OECD 国家，支出审查是以风险为导向的，即审查的重点落脚于未来重大公共风险问题的解决。

以英国为例。首先，其支出审查模式要求为某些公共领域制定全面的财务和业务基线，以明晰各部门职责并充分考虑其中存在的风险点。在此要求下，各部门都会制作一份“价值地图”作为构建财务基线的微观基础；支出审查小组则会以公共风险为导向制定明确的业务基线，以划定复杂的、易发生部门摩擦的风险领域（Assi，2019[19]）。其次，英国通过支出审查确定的中长期部门支出上限设置有调整窗口以应对风险的变化，例如在年度预算审核和拨款评估期间，支出上限都可以根据公共风险及政策优先级的变化加以调整，以确保短期政策的灵活性。

（四）聚焦关键部门：零基预算理念长期可持续推行的关键

支出审查在实施过程中会涉及到成本和效益的权衡问题。不同的审查范围对技术、能力以及资金投入等的要求会有较大差异，例如对运营成本的审查往往会关注劳动力成本、外部支出、后台流程等环节，而对转移支付的审查则通常会着眼于资格、欺诈等问题。显然，这两种审查的投入成本不同。另外，审查对象筛选标准的确定也十分关键。跨部门审查可以解决更为复杂的问题，但实施起来也更具挑战性；审查项目的支出规模如果太小，其结果或难以产生实质性影响，但规模太大又可能无法在可用时间内完成细致的审查。在英国，大多数审查都集中在年度支出至少 10 亿英镑的领域（Assi，2019[19]）。由此可知，基于成本效益的考虑，支出审查应聚焦于支出金额和不确定性均较大的关键部门。

（五）程序控制：以零基预算理念促进管理创新

相较于传统方法，当前零基预算的落地模式更加强调程序控制和部门协同，这都较大地促进了管理方式和理念的创新。以英国为例。首先，在其预算流程中，支出审查是一个类似于中长期规划的关键环节，财政部基于它对部门资金做出框定，各部门只能在与国家优先政策保持一致且不超出资金限额的情况下安排支出。其次，在单一部门计划、年度报告等中，各部门均需公布自身的绩效指标及执行情况，而且统计等部门也会追踪和发布政府支出绩效信息。这一整套预算程序，几乎使英国政府实现了对各部门预算绩效的全流程追踪、监督和控制。再次，支出审查在部分继承传统零基预算自下而上的编制办法的同时，又揉入了自上而下的项目抽查评估、优先目标及部门支出限额设定等内容，部门协同的理念更加凸显。

五、在我国运用零基预算理念的几点建议

由于实践中的诸多困难，零基预算作为一种具体的预算方式很难在我国真正展开。面对当前突出的收支矛盾，我国应充分总结各地实施零基预算的经验和教训，并借鉴 OECD 国家的有益经验，运用零基预算理念开展以支出审查为核心的预算措施，同时推动相关预算改革，以促进财政支出结构的全面优化。

第一，针对部门职能划分不清晰、工作量和技术难度大的问题，建议舍弃零基预算的繁琐程序，在部门预算绩效管理流程中加入支出审查条款，并将是否聚焦化解重大公共风险作为审查的基本依据。对部门支出开展全面审查，可减少部门之间任务的重叠和冲突，改进部门绩效目标并增强其约束性。而目前我国所实施的预算绩效管理只强调了支出要重点关注的内容，并没有支出审查方面的要求。建议在其中加入支出审查条款，并在建立全过程预算绩效管理链条中加入“开展支出审查”和“以聚焦化解重大公共风险作为支出审查依据”等内容。

第二，针对工作量大但效果不佳的问题，建议对不同金额、不同类型的项目有区别地进行审查。建议对中央部门 1 000 亿元以上的支出项目轮流开展支出审查，5~8 年实现全覆盖，1 000 亿元以下的项目则实行一定比例的抽查，并按照“聚焦重点先行试点、形成经验全面推开”的原则设计覆盖范围。在试点选择上，可先行聚焦社会关切的重要项目和资金，如一些法定支出以及产业发展、扶贫、公共卫生等领域

的支出，并重点对有可能妨碍公平竞争、拖累市场效率的企业创新和产业转型升级扶持资金开展跨部门支出审查。地方可比照中央的做法按各地情况实行。

第三，针对难以确定决策部门及监督力度不足等问题，建议由人大预算部门、财政和审计部门等联合设立支出审查工作协调机构，以切实推行零基预算理念、开展支出审查。协调机构办公室建议设立在财政部，其主要职能包括：确定和评价各部门的“关键活动”，建立和评估绩效指标，分析项目预算支出，提出预算安排建议以及牵头推动绩效公开等。同时，建议在财政部设立成本计算部门，以对公共支出开展定期的快速审查，为支出审查工作协调机构提供决策依据。

第四，为从根本上解决制约预算改革的难题，建议在开展支出审查的过程中，同步做好相关基础性工作。这些工作包括：解决预算分配权分散的问题；有序开展部门情况摸底，推动人员精简、部门职能聚焦和转变并规范部门支出；建立有效的监督体系和可问责的决策体系，将资金决策从一种行政权力转变为一种法律责任和道德义务，即变成一种“风险”，从而使决策主体小心谨慎地申请、分配、使用和管理资金，促使部门权力与责任相对称；加快推动大数据、云计算和区块链等技术在支出审查中的运用；总结实践经验形成规范性文件；等等。

参考文献

[1] 马骏. 中国公共预算面临的最大挑战：财政可持续 [J]. 国家行政学院学报，2013（5）：20.

[2] 财政部. 2020 年一季度财政收支情况 [EB/OL]. (2020-04-20)[2020-05-30]. http://gks.mof.gov.cn/tongjishuju/202004/t20200420_3501077.htm.

[3] 知秋. 过紧日子须有“制度保障”（财经观）[N/OL]. (2020-02-10)[2020-05-09]. http://finance.people.com.cn/n1/2020/0210/c1004-31578536.html.

[4] 江西省人民政府. 江西省直行政机关今年起全面实行零基预算 [EB/OL]. (2020-01-21)[2020-05-09]. http://www.jiangxi.gov.cn/art/2020/1/21/art_398_1492285.html.

[5] 国家统计局. 中国统计年鉴 [M]. 北京：中国统计出版社，2019.

[6] 财政部. 关于 2019 年中央和地方预算执行情况与 2020 年中央和地方预算草案的报告 [R/OL]. (2020-05-30)[2020-06-05]. http://www.mof.gov.cn/gkml/caizhengshuju/202005/t20200530_3523307.htm.

[7] 牛美丽. 美国公共预算改革：在实践中追求预算理性 [J]. 武汉大学学报（社会科学版），2003（6）：795.

[8] 罗宏斌，赵彬. 关于我国政府预算引入零基预算的一些思考 [J]. 甘肃农业，2006（4）：182.

[9] Austin L A，Cheek L M. Zero-base Budgeting：A Decision Package Manual [M]. New York：Amacom，1979：2.

[10] 聂丽洁，王俊梅. 关于绩效预算与零基预算相结合的预算方法体系的思考 [J]. 中央财经大学学报，2004（12）：1.

[11] 肖育才. 中美零基预算改革之比较 [J]. 财会月刊（综合），2007（7）：82.

[12] 吕彩凤. 新预算法背景下我国实施零基预算改革的思考 [J]. 时代金融，2009（9）：202.

[13] 王辰，李彤. 中美零基预算的比较及启示 [J]. 经济纵横，2011（8）：97-99.

[14] 马骏. 中国的零基预算改革：来自某财力紧张省份的调查 [J]. 中山大学学报（社会科学版），2005（1）：59-67.

[15] 马骏，叶娟丽. 零基预算：理论和实践 [J]. 中国人民大学学报，2004（2）：122-129.

[16] Schick A. The Road From ZBB [J]. Public Administration Review，1978（38）：178.

[17] OECD. Budgeting and Public Expenditures in OECD Countries 2019 [M]. Paris：OECD Publishing，2019：112.

[18] OECD. 2016 OECD Performance Budgeting Survey：Integrating Performance and Results in Budgeting [R/OL]. (2016)[2020-05-08]. http://www.oecd.org/gov/budgeting/Performance-Budgeting-Survey-Highlights.pdf.

[19] Assi R，et al. Spending Reviews：A More Powerful Approach to Ensuring Value in Public Finances，McKinsey & Company [Z/OL]. (2019-03-08)[2020-05-08]. https://www.mckinsey.com/industries/public-sector/our-insights/spending-reviews-a-more-powerful-approach-to-ensuring-value-in-public-finances#.

[20] Wheatley M，et al. The 2019 Spending Review How to Run it Well [R/OL]. (2018-09)[2020-05-08]. https://www.instituteforgovernment.org.uk/sites/default/files/publications/IfG_2019_%20spending_review_web.pdf.

[21] Office of the Independent Budget Analyst Report. Zero-Base Budgeting Concepts and Examples [R/OL]. (2016-6-23)[2020-05-08]. https://www.sandiego.gov/sites/default/files/ibareport16-16.pdf.

（责任编辑：孙亦军　张安平）

中国城市财政压力现状与风险识别

——基于新口径的测算

The Status of Chinese Urban Fiscal Pressure and the Identification of Urban Fiscal Risk: Based on the Calculation of New Caliber

秦士坤

QIN Shi-kun

［摘　要］防范与化解地方政府财政风险已成为当下研究的重要课题，但当前对财政压力的测算尚缺乏足够的现实支撑。本文基于对地方政府债务、PPP 财政支出责任、地方政府综合财力的数据整理，提出了新的测算口径，对 2014 至 2019 年我国 337 座地级以上城市的财政压力进行了度量，并使用 KMV 模型对违约风险进行评估。研究发现，地方政府债券兑付与 PPP 的财政支出责任是当下我国城市财政压力的主要来源，忽视 PPP 所带来的隐性负担将显著低估城市财政压力水平。此外，基于空间相关性分析，本文通过 Getis-Ord Gi* 统计量进一步对财政压力的热点与冷点城市进行了识别，发现城市财政压力呈现空间集聚性特征，并已形成持续性风险高发区域。上述研究为完善我国城市财政压力测度方式、精准识别风险高发区域提供了一定的启示。

［关键词］财政压力　地方政府债务　PPP　财政风险

［中图分类号］F812.7　［文献标识码］A　［文章编号］1000-1549（2020）10-0010-16

Abstract: Preventing and resolving local government fiscal risks has become an important topic in current research, but the current calculation of fiscal pressure still lacks sufficient realistic support. Based on the data collation of local government debt, PPP fiscal expenditure responsibility, and local government comprehensive financial resources, this paper proposes a new calculation caliber, measures the fiscal pressure of 337 prefecture-level cities in China from 2014 to 2019 and uses the KMV model to evaluate the default risk. The study found that the local government bond redemption and PPP fiscal expenditure responsibility are the main sources of fiscal pressure in China's cities. Ignoring the hidden burden of PPP will significantly underestimate the level of urban fiscal pressure. In addition, based on spatial correlation analysis, this article further identified hotspots and coldspot cities of fiscal pressure through Getis-Ord Gi* statistics, and finds that urban fiscal pressure exhibits spatial agglomeration characteristics and has formed areas of high persistent risk. The above research provides some inspiration for improving China's urban fiscal pressure measurement method and accurately identifying areas with high risk.

Key words: Fiscal pressure　Local government debt　Public-Private Partnership（PPP）　Fiscal risk

［收稿日期］　2020-06-21

［作者简介］　秦士坤，男，1991 年 4 月生，中央财经大学中国财政发展协同创新中心博士研究生，研究方向为政府预算与绩效评价。

感谢匿名评审人提出的修改建议，笔者已做了相应修改，本文文责自负。

一、引言

自 2014 年新《预算法》修改以来，随着我国财政预算制度的不断完善、融资平台市场化转型的不断推进，地方政府的融资模式在逐步得以规范化、透明化，地方政府的无序举债行为得到了一定的约束。但与此同时，当前地方政府仍面临较大的财政压力：从收入端来看，经济下行与减税降费政策直接影响了地方政府财政资金的流入，此外，一系列重大公共危机的突发亦对地方财政形成了直接冲击，这种冲击既包括全局性的（例如 2020 年初爆发的新冠肺炎疫情导致 2020 年上半年我国地方一般公共预算收入下降 7.9%），也包括区域性的（如突发性的洪灾、地震等自然灾害使得受灾地区企业无法正常运转，继而影响当地财政收入），为地方政府财政收支平衡带来了极大的挑战。而从支出端来看，一方面由于前期累积的债务规模较为庞大，地方政府面临债券期限到期后的本息兑付压力；另一方面由于 2014 年以来我国大力推广政府和社会资本合作（Public-Private Partnership，PPP）模式用于公共基础设施建设，随着大量的 PPP 项目进入运营阶段，财政支出责任的履行也将成为当下面临的现实问题。

财政压力的累积直接导致财政收支矛盾的加剧，继而可能引发财政风险，应对财政压力是当前地方政府面临的重要议题。在相关研究中，合理的财政压力测度是所有讨论的首要前提，唯有真实有效地测度财政压力，才能真正厘清地方政府财政压力现状、识别区域财政风险、保障政策精准实施。然而在现有研究中，对财政压力的测度要么未能适应新时期地方政府面临的客观现实（如未考虑政府债务偿付与 PPP 支出责任），要么局限于某项政策对财政收支的短期冲击，无法有效地刻画地方政府的整体财政压力，极大地限制了应对财政压力的讨论空间。此外，限于数据的可得性，相关研究主要集中于省级以上层面，而城市财政压力未能得到充分的关注。本文选取城市作为财政压力的分析对象，其原因在于：第一，城市是经济活动的基本空间单元，彼此间既有竞争又相互依存，随着经济发展阶段的演变，城市已成为区域竞争力的核心，深刻改变了经济活动的空间形态（如近年来城市群的崛起），以城市为焦点进行财政压力的测度有助于揭示城市维度的财政状况，更为细致地刻画财政压力的空间分布特征。第二，由于我国各省份的城市间财政状况差异较大，仅关注省级层面的财政压力忽略了城市层面的异质性，导致区域性财政风险无法充分暴露，政策手段也难以实现因地制宜的调控。基于上述原因，本文基于新《预算法》修改以来地方政府面临的新的客观现实，构建了财政压力测度的新口径，并在此基础上，从违约风险、空间集聚等角度对城市层面的高风险区域进行识别，为财政压力的讨论提供了基础的指标支撑。

本文剩余部分安排如下：第二部分为文献综述；第三部分为财政压力值测算与风险评估；第四部分为财政压力的空间特征分析；第五部分为结论与政策含义。

二、文献综述

地方政府财政的收支矛盾影响了地方政府的行为活动，并可能引发财政风险，具有一定的负面效应。举例而言，地方政府为了解决资金问题，有动机进行土地的掠夺（罗必良，2010[1]），推高当地的房地产价格（王雅龄和王力结，2015[2]），扭曲行业发展重心（Han 和 Kung，2015[3]），影响税收负担的分配（Chen，2017[4]），并导致国企过度投资（曹春方等，2014[5]）。此外，财政压力的累积过程蕴含了财政风险，亟需进一步加以识别，如刘尚希（2004）[6]认为无论是地方政府内部，还是地方政府与国企之间，风险责任的界定都具有模糊性，这种“风险大锅饭”导致了财政兜底，从而使财政风险不断累积。

对于如何测度地方政府的财政压力，学界尚未形成普遍共识（陈晓光，2016[7]），随着地方政府投融资模式的变迁，其测度手段也在不断更新。梳理文献可知，当前学者对财政压力的测算主要有三种角度，第一种是从当年的收支缺口角度进行衡量，如孙开和孙磊（2019）[8]、毛捷等（2020）[9]使用当年一般公共预算收支差额与 GDP 的比重衡量财政压力，亓寿伟等（2020）[10]通过对收支缺口的比例进行回归构造了财政压力变量。这种方法的特点是较为直观，数据相对容易获取，也成为当前应用较为普遍的测度方式，但其缺陷也较为明显：由于地方政府当年的收支缺口包含转移支付收入，属于地方政府可以灵活支配的部分，因此收支缺口对地方政府财政压力的解释较为有限。第二种是从存量债务角度进行衡量，如朱军等（2019）[11]搜集了 2014—2017 年 333 座地级市的债务余额数据，以潜在债务违约率衡量财政压力。这种方法的特点是能够从广义政府收支角度考量地方政

府未来的偿付责任，相较于当年收支缺口包含的信息量更大，缺点则是地方政府债务数据相对较难获取。第三种侧重于财政压力的“变化”，通常以某个事件对地方政府财政收入带来的冲击进行衡量，如陈晓光（2016）[7]使用 2005 年的取消农业税改革、席鹏辉等（2017）[12]使用 2002 年前后的增值税分成变动、毛捷等（2020）[9]使用 2008 年的企业所得税改革来衡量财政压力变化。这些事件通常在某一个时刻对不同地区产生了不同影响，因此便于构建准自然实验进行因果推断，缺点则是未能衡量地方政府的整体财政压力水平。根据上述梳理可知，财政压力是广义概念，用于描述地方政府的财政收支状况，可根据不同的研究需要从不同角度加以衡量。

在以上三种方法中，基于存量债务角度对财政压力的测度最为全面，更适用于分析地方政府在一段时期内所面临的广义收支矛盾，但同时也存在以下问题：

第一，忽视了 PPP 财政支出责任所形成的财政压力。2014 年修订的《预算法》规定地方政府债券是地方政府举借债务的唯一形式，因此地方政府的传统筹资手段（如融资平台）在 2014 年以来面临合法性危机，而这一时期 PPP 作为创新的筹资手段在中央推动下迅速发展，截至 2019 年年末，其撬动的投资规模已逾 14.4 万亿元，在投资规模上已接近于地方政府债券。① 在 14.4 万亿元的 PPP 投资规模中，由于 PPP 大量采用政府付费与可行性缺口补助模式，地方政府面临较大的财政支出责任，这些支出责任被分散在未来 PPP 项目的运营期，容易为外界所忽视。尽管为了防止支出固化带来的财政风险，财政部规定财政支出责任水平应与绩效考核结果进行挂钩，② 但其支出在某种程度上仍是刚性的、无法回避的，为地方政府带来了沉重的财政压力。因此以 2014 年修订的新《预算法》作为分界点，在当前阶段，地方政府债券与 PPP 已成为我国地方政府的两大主要筹资渠道，分析财政压力不应忽视 PPP 所带来的财政支付责任。借鉴国外经验，可发现发达地区实施 PPP 更多以效率与服务为导向，而发展中地区更多考量经济与社会效益（Osei-Kyei 和 Chan，2018[13]），因此，基于财务困难的考量，政府部门可能会试图通过 PPP 模式缓解投资压力（Buso 等，2017[14]）。然而，并非所有类型的公共基础设施都适用于 PPP（Oliver，2003[15]），由于资金成本更高，PPP 的不当运用可能导致地区财务状况更加恶化（Hodge 和 Greve，2007[16]）。基于这一背景，马恩涛和李鑫（2018）[17]结合国际经验提出了管理 PPP 政府或有债务的建议，谭艳艳等（2019）[18]提出了 PPP 项目中的政府债务风险的识别思路，Wang 等（2019）[19]认为应建立风险评估和预防机制以规避 PPP 过度投资带来的债务负担。尽管 PPP 财政支出责任在形式上不能等同于地方政府债务，但学者们的观点反映了两者之间的模糊性。陈少强（2018）[20]认为当前我国的 PPP 财政风险管理属于流量管理，③ 缺乏对存量部分的重视，应当引入基于当前存量 PPP 数据的风险管理指标，从而避免“有意识的模糊”。近年来，已有学者开始关注 PPP 形成的财政压力，如张牧扬等（2019）[21]通过收集 2010—2018 年财政部 PPP 项目库的项目数据测算了 2019—2033 年地方政府的 PPP 财政支出责任，魏蓉蓉等（2020）[22]根据省级层面的 PPP 投资额占比估算了 PPP 地方政府隐性债务规模。然而相关领域的研究刚刚起步，且尚未有学者对地方政府债务与 PPP 所形成的综合财政压力进行过测算。其主要原因在于地方政府债务与 PPP 的数据较难获得，工作量较大，尤其是 PPP 财政支出责任的计算需要对全部 PPP 项目的相关数据进行加总，因此鲜见定量、综合的分析。

第二，未体现政府性基金收入在应对财政压力中的作用。在当前我国的政府预算体系中，政府性基金预算已成为重要的构成部分。根据地方预算数据可知，2019 年地方一般公共预算收入为 17.5 万亿元，而地方政府性基金收入达 8.0 万亿元。由于地方政府（特别是城市层面）政府性基金收入的数据较难获取，现有的财政压力测度通常未考虑这部分收入。在使用方面，政府性基金收入不仅可以部分调入一般公共预算使用，而且是专项债务偿还的直接收入来源，因此计算地方政府财力时忽视政府性基金收入将造成

① 根据 CPPPC 发布的《全国 PPP 综合信息平台管理库项目 2019 年报》，截至 2019 年年末，管理库 PPP 项目累计 9 440 个，累计投资额 14.4 万亿元。根据财政部发布的《2019 年 12 月地方政府债券发行和债务余额情况》，截至 2019 年 12 月末，全国地方政府债务余额 21.3 万亿元。

② 详见《财政部关于推进政府和社会资本合作规范发展的实施意见》（财金［2019］10 号）。

③ 《政府和社会资本合作项目财政承受能力论证指引》（财金〔2015〕21 号）提出，“每一年度全部 PPP 项目需要从预算中安排的支出责任，占一般公共预算支出比例应当不超过 10%。”

地方政府财力水平的低估。在债务风险的测算中，已有学者通过构造专项债务率指标纳入了政府性基金因素（刁伟涛和傅巾益，2019[23]），而财政压力的相关研究尚鲜有对政府性基金收入的考量。

本文可能的边际贡献在于：首次引入财政压力值这一指标，在支出端综合考虑了地方政府债务与 PPP 财政支出责任，在收入端综合考虑了一般公共预算收入与政府性基金收入，测算了 2014—2019 年我国城市层面的财政压力情况，为财政压力的测度提供了新的角度。此外，在财政压力值的测算基础上，本文进一步结合城市地理空间因素对财政压力的空间集聚性进行了评估，识别了财政风险的持续性集聚城市，为城市层面的系统性风险识别提供了一定的参考。

三、财政压力值测算与风险评估

（一）财政压力值测算

本文财政压力值的计算主要分为 PPP 财政支出责任的测算与地方政府债务余额的确定两方面。为考量 PPP 所带来的财政压力，需要深入至具体的项目层面进行考量。根据《政府和社会资本合作项目财政承受能力论证指引》（财金〔2015〕21 号，以下简称“21 号文”），地方政府在 PPP 项目周期内的财政支出责任包括股权投资、运营补贴、风险承担、配套投入。由于财政部的 PPP 综合信息平台（以下简称“CPPPC”）要求所有入库项目在识别阶段公布未来年度财政支出责任数额的测算数据①，因此 CPPPC 的公开信息为本文中城市层面 PPP 财政支出责任的测算提供了依据，并使得该测算具备了操作上的可实现性。基于明树数据的数据支持②，本文收集了 CPPPC 公布的 2014 至 2019 年 12 535 个 PPP 项目未来年度的财政支出责任信息。然而由于 PPP 的落地存在较多不确定性因素，PPP 项目的入库不必然形成地方政府的财政支出责任，因此本文进一步对已完成中标的项目进行筛选，并去除省本级与中央本级的项目，最终得到 7 697 个已中标 PPP 项目的财政支出责任信息，其中财政支出责任的数据跨度为 2014 至 2067 年。描述性统计详见表 1。

表 1　　PPP 项目信息描述性统计

指标		观测值	合计值	最大值	最小值	平均值
按回报机制划分（亿元）	合作年限（年）	7 697	—	50.00	5.00	19.27
	投资金额（亿元）	7 697	111 150.62	886.39	0.07	14.44
	政府付费	3 031	31 997.95	319.00	0.07	10.56
	可行性缺口补助	4 192	71 969.96	647.35	0.16	17.17
	使用者付费	474	7 182.71	886.39	0.15	15.13
按行业类型划分（亿元）	交通运输	1 049	28 111.53	886.39	0.30	26.80
	农林水利与环境保护	1 155	12 745.48	268.99	0.13	11.04
	社会事业与其他	1 370	12 360.53	350.00	0.07	9.02
	市政公用事业	3 491	37 532.34	647.35	0.09	10.75
	综合开发	632	20 400.74	490.62	0.79	32.28

根据上表可知，筛选出的 PPP 项目平均合作期限为 20 年，平均投资金额为 14.44 亿元，其中约 93.8%的项目属于政府付费或可行性缺口补助类型，这意味着绝大多数 PPP 项目均涉及地方政府的财政支出责任，只有少数使用者付费项目可以通过自身经营收入实现运转。从行业类型来看，交通运输、市政

① 财政部对项目信息的公开要求见《政府和社会资本合作（PPP）综合信息平台信息公开管理暂行办法》（财金〔2017〕1 号），CPPPC 平台网址：www.cpppc.org。

② 本文得到了明树数据“明日之树”项目的数据支持，在此表示感谢。项目网址：http://bridata.com/about/article/detail?id=5935&code=0103&title=。

公用事业是 PPP 资金的主要流向，合计占比达 59.1%。对相关数据进行提取，并将各年度 PPP 项目的财政支出责任数据进行加总（如图 1），可知支出责任在未来将呈现先升后降趋势，在 2024 年前后达到峰值。根据项目实施机构定位至区域来看，东部地区面临的未来支出的绝对规模相对最高，中西部较为接近，东北部地区相对最低。①

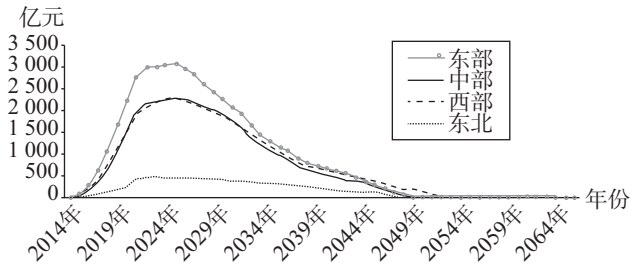

图 1　2014 至 2067 年 PPP 财政支出责任累计值

为了测算地方政府当年面临的 PPP 财政压力，需要将未来财政支出责任的流量数据转变为当前的存量数据，即财政支出责任的现值。在年度折现率的确定方面，21 号文提出年度折现率应“参照同期地方政府债券收益率合理确定”。因此，本文测算了各省级政府发行的 10 年期以上地方政府债券的平均发行利率，结合实践经验②，将各城市的 PPP 年度折现率确定为所在省份 10 年期以上债券平均发行利率上浮 50%，其中最高值为贵州的 5.64%，最低值为西藏的 5.17%。基于这一设定，单个 PPP 项目的财政支出责任现值计算公式为：

$$P_{it}=E_{it}+\frac{E_{i,t+1}}{(1+\theta_j)}+\frac{E_{i,t+2}}{(1+\theta_j)^2}+\cdots+\frac{E_{i,t+n}}{(1+\theta_j)^n} \tag{1}$$

其中，i 代表项目，j 代表城市，t 代表年份（t 的取值范围为自中标年份起，至 2019 年），n 代表当年至支付周期最后一年的间隔年数，E_{it} 代表项目 i 在年份 t 面临的财政支出责任，θ_j 代表项目 i 所在城市 j 的 PPP 年度折现率。在 P_{it} 基础上，可进一步汇总得到城市 j 在年份 t 的 PPP 财政支出责任现值 PPP_{jt}，以衡量各城市在不同年份面临的财政支出压力，计算公式如下：

$$PPP_{jt}=\sum P_{it} \tag{2}$$

对于地方政府债务水平的测算，学者有多种不同的测算方式，从而导致结果差异较大。地方政府债务余额是当前可直接衡量地方债务负担的官方指标，虽然这一口径忽视了隐性债务负担，但一般说来，地方政府暴露出来的债务水平在某种程度上可反映地方政府的真实债务水平。为获取地方债务数据，本文通过人工搜集方式，在各城市的预决算报告或直接登录财政局网站搜索“债务”关键词查找相关债务余额数据。通过大量的人工搜集工作后发现，2016—2019 年地方政府债务数据的披露情况相对较好，80%以上地区均能查找到债务余额的相关数据。而由于财政部于 2014 年至 2016 年才相继发布《地方政府存量债务纳入预算管理清理甄别办法》《关于对地方政府债务实行限额管理的实施意见》《地方政府一般债务预算管理办法》《地方政府专项债务预算管理办法》等政策文件并建立预算管理框架，因此 2014—2015 年债务数据的披露程度相对较差，本文结合新增余额数据进行缺失值补充。根据财政部公布的债务发行信息③，可知各年度的债务余额计算公式为：

本年度债务余额 = 上一年度地方债务余额 + 本年地方政府债务发行额 - 本年度地方政府债务还本额

根据财政部公布的数据，可计算 2015、2016 年新增的债务余额分别为 -6 505.93 亿元、5 989.22 亿元，与相邻年份的债务余额相比，变动比率不到 5%。因此可在 2016—2017 年各地区债务余额数据基础上，合理填充 2014 至 2015 年的缺失数据。本文借鉴毛捷和黄春元（2018）[24]的做法，将未来一年新增的债务余额按照各地级市 GDP 占省级 GDP 的比例分配至各地市，从而得到相应年份的债务余额数据，最终整理出 338 座城市 2014—2019 年间的债务规模（包括一般债务余额与专项债务余额）全样本数据。由于财政压力不仅取决于存量规模，还取决于经济规模（刘尚希和赵全厚，2002[25]），因此本文采用比值指标衡量地方政府的财政压力。财政压力的测算公式定义为：

$$Press_{jt}=\frac{Debt_{jt}+PPP_{jt}}{General_{jt}+Revenue_{jt}} \tag{3}$$

① 本文中对东中西部和东北地区的划分方法依据国家统计局划分方法确定。

② 基于对相关行业专家的访谈，实践中 PPP 年度折现率的取值一般在 5%~6%之间。

③ 财政部地方债务管理信息，网址：http://yss.mof.gov.cn/zhuantilanmu/dfzgl/sjtj/index.html。

其中，*Press* 代表财政压力值，*Debt* 代表政府债务余额，*General* 代表一般公共预算收入，*Revenue* 代表政府性基金收入。在式（3）中，分子整体衡量的是各城市在当期面临的未来财政支出责任，分母则代表地方政府的综合财力。与传统的测算方式相比，式（3）的优势主要在于：综合考量了债务与 PPP 对地方政府带来的未来支出责任，更为全面地考量了地方政府的支付压力。此外，由于政府性基金收入的数据较难获取，传统的地方政府债务率的测算容易忽视这一因素。然而，由于政府债务余额中的专项债务余额是由政府性基金收入进行偿还，因此综合财力的测算不应忽视政府性基金收入。① 式（3）在将我国地方政府综合财力定义为当年一般公共预算收入与政府性基金收入之和，更为全面地考量了地方政府的支付能力。

根据式（3）计算所得的 338 座城市财政压力值绘制核密度图，可清晰地观察其分布特征。图 2 左图展示了各城市 2014—2019 年财政压力值的分布变动情况，从中可以看出，财政压力值的分布呈现明显的右偏特征，有相当多城市远大于平均财政压力水平，暴露了一定的财政风险。从时间趋势来看，曲线逐年向右移动反映财政压力值呈现明显的上升趋势，显然这与 PPP 近年来的迅速推进有着时间关系。图 2 右图展示了本文的财政压力值与传统的债务率计算公式的差别。债务率的计算通常只关注债务对综合财力的影响，而加入 PPP 财政支出责任的现值以后，曲线出现了较大幅度的右移，说明在不考虑 PPP 财政支出责任的情况下，地方政府面临的财政压力水平被明显地低估。

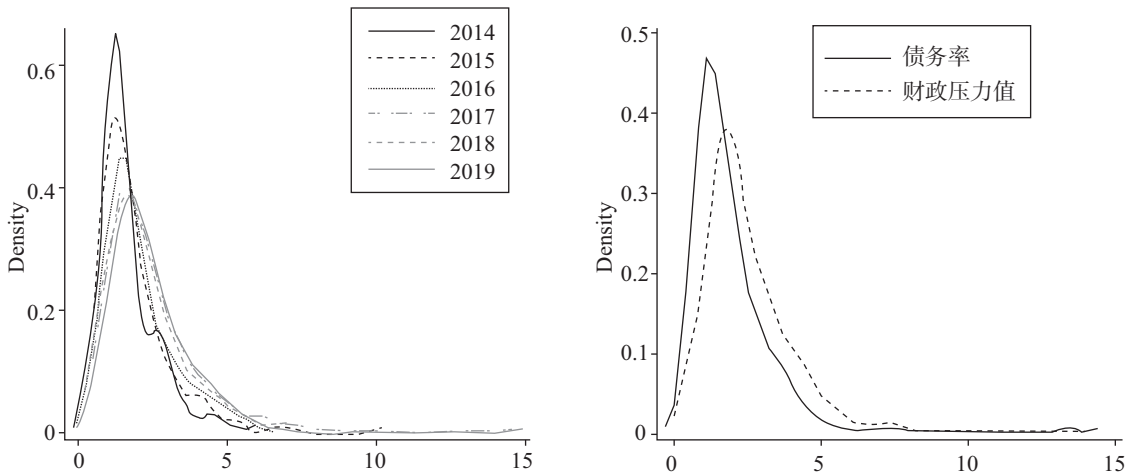

图 2　城市财政压力值分布情况

基于上文的测算，可对地方政府债券与 PPP 两种形式的筹资模式进行粗略比较。以 2019 年为例，城市层面的债务余额合计约为 18.4 万亿元，② 而 PPP 财政支出责任现值合计约为 6.5 万亿元，说明尽管近年来 PPP 推广迅速，城市地方政府面临的最大偿还压力仍然来自于地方政府债券的兑付。但不可忽视的是，PPP 在整体的财政压力中已占有了较大的比重，而且由于 PPP 的年度折现率（5%~6%）相较地方政府债券利率（3%~4%）更高，PPP 模式的筹资成本显著高于地方政府债券。因此，未来 PPP 支出责任的履行以及在项目运营过程中面临的不确定性将是财政风险的重要来源，值得加以关注。

（二）KMV 风险评估

财政压力值衡量的是地方政府未来所面临的财政支出责任。对于地方政府而言，即使当期财政压力值较大，只要未来财政支付的期限结构分布合理，仍然可以使风险在时间维度上得以分散，从而有效缓解财政收支矛盾。同样地，若财政支付的期限结构分布较为集中，即使当期财政压力值较小，仍有可能在未来某一年形成收支矛盾，从而触发财政危机。因此，在上节财政压力的测算理念基础上，本节从当期财政支出责任及其对应的可偿债收入角度，结合

① 鉴于国有资本经营预算收入规模相对较小，且数据较难获取，本文在综合财力的测算中未纳入这部分收入。

② 由于本文侧重于城市层面的分析，未包含省级政府的债务余额，因此本文测算的城市层面债务余额水平（18.4 万亿）低于 2019 年地方政府债务的总余额（21.3 万）。

KMV 模型，进一步衡量城市地方政府的财政违约风险。

KMV 模型是由美国 KMV 公司于 1997 年开发的一种信用评估方法，在地方政府财政风险的测度中得到了广泛应用（王学凯和黄瑞玲，2015[26]；王周伟等，2019[27]；魏蓉蓉等，2020[22]）。其主要原理是通过引入财政收入的可支配部分与波动率、增长率，测算当期违约距离（DD），其中 DD 越小代表风险越高。相关测算公式如下：

$$g=\frac{1}{n-1}\sum_{t=1}^{n-1}\ln\frac{A_{t+1}}{A_t}+\frac{1}{2}\sigma^2 \tag{4}$$

$$\sigma=\sqrt{\frac{1}{n-2}\sum_{t=1}^{n-1}\left(\ln\frac{A_{t+1}}{A_t}-\frac{1}{n-1}\sum_{t=1}^{n-1}\ln\frac{A_{t+1}}{A_t}\right)^2} \tag{5}$$

在上式中，g 代表地方政府财力的增长率，σ 代表波动率，A_t 代表在 t 年份可用于担保的财力水平。参考蒋忠元（2011）[28]、李腊生等（2013）[29]的测算方法，本文选取综合财力的 30% 作为可担保财力的估计值。此外，与上节财政压力的测算理念类似，本文使用债务偿还本息与 PPP 支出责任的合计值测算各年的偿还规模，其计算公式如下：

$$B_t=DE_t+PE_t \tag{6}$$

其中，DE_t 代表 t 年的债务偿还本息，PE_t 代表 t 年的 PPP 财政支出责任。其中 DE_t 的城市数据使用 WIND 数据库上地方政府债券的兑付金额与省内债务余额的分配比例进行测算，PE_t 数据使用 CPPPC 上公开的 PPP 项目财政支付责任数据进行汇总测算。这一测算方式不仅纳入了 PPP 的支出责任，而且对债务的偿还规模更为直观。基于上述数据，可测算城市地方政府面临的违约距离：

$$DD=\frac{\ln\frac{A_t}{B_t}+\left(g-\frac{1}{2}\sigma^2\right)\Delta t}{\sigma\sqrt{T}} \tag{7}$$

其中，$\Delta t=T-t$，$t=0$。根据上述计算结果，可得 2014—2019 年各城市的违约距离。将 2019 年各省份的平均违约距离展示如图 3。

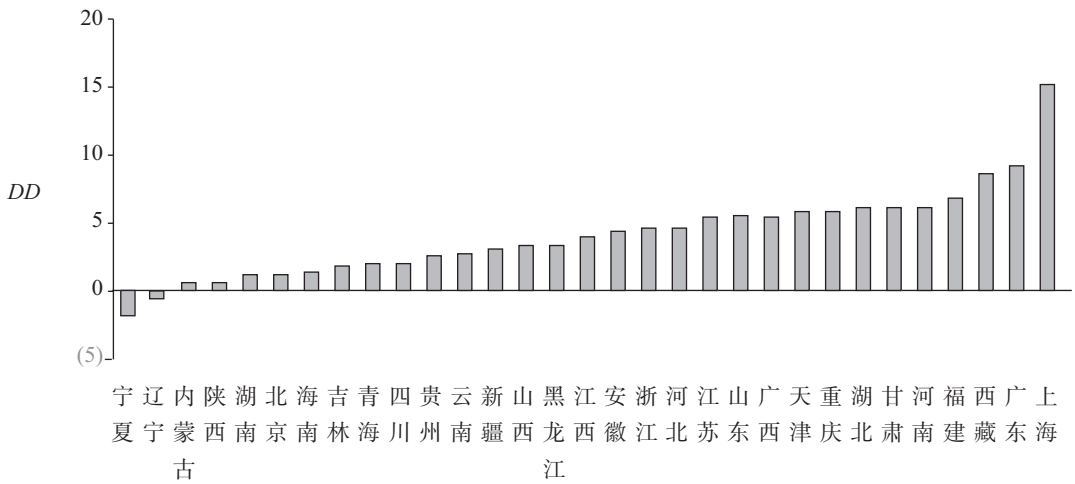

图 3　2019 年各省份平均违约距离

根据结果可知，在 30% 的可担保财力且考虑 PPP 财政支出责任的估计下，上海、广东、西藏的财政风险相对较小，而宁夏、辽宁、内蒙古的财政风险相对较大，甚至达到负值。而在城市层面，2019 年违约距离最近的 10 座城市分别为：陕西商洛、宁夏石嘴山、宁夏固原、宁夏中卫、湖南湘西州、湖南张家界、辽宁铁岭、辽宁鞍山、贵州六盘水、辽宁抚顺；违约距离最远的 10 座城市分别为广东茂名、湖北荆门、福建厦门、广东深圳、甘肃酒泉、广东江门、上海、河南郑州、湖北武汉、广东惠州。

通过预测各城市财力变化情况，可进一步测算未来的城市违约距离。本文使用往年的平均增长率对财力的波动进行预测，鉴于 2020 年初爆发的新冠肺炎疫情对地方财政收入产生了深远影响，本文引入各省份 2020 年第一季度的财政收入增长率对未来增长率的预测进行修正，从而预测得到各城市未来 5 年的财力水平。基于前文测算的未来支出责任，计算得到各城市未来五年的违约距离（按由近至远排序）。根据结果（见表 2）可知，未来五年财政风险较大的城市有着明显的持续性与空间集聚特征。

表 2 未来五年城市违约距离最近的十座城市

2020 年	2021 年	2022 年	2023 年	2024 年
陕西商洛	宁夏中卫	宁夏中卫	宁夏中卫	宁夏中卫
宁夏石嘴山	陕西商洛	陕西商洛	陕西商洛	江西新余
宁夏中卫	宁夏石嘴山	江西新余	江西新余	宁夏石嘴山
江西新余	宁夏固原	宁夏石嘴山	黑龙江佳木斯	陕西商洛
黑龙江佳木斯	黑龙江佳木斯	黑龙江佳木斯	宁夏石嘴山	宁夏固原
贵州六盘水	贵州六盘水	吉林松原	甘肃定西	吉林延边州
吉林延边州	江西新余	吉林延边州	安徽铜陵	黑龙江佳木斯
吉林松原	内蒙古呼伦贝尔	贵州六盘水	宁夏固原	吉林松原
宁夏固原	吉林延边州	宁夏固原	吉林延边州	贵州六盘水
内蒙古呼伦贝尔	吉林松原	内蒙古呼伦贝尔	吉林松原	吉林辽源

四、财政压力的空间特征分析

（一）空间相关性检验

财政压力在城市间的分布并非孤立的，由于地方政府的行为具有策略互动性，邻近地区间地方政府的举债行为具有显著的空间关联（刁伟涛，2016[30]；王周伟等，2019[27]）。而这种空间关联性一旦固化，可能会逐步形成风险聚集区域，值得加以关注。为了深入探究我国城市层级财政压力的空间分布特征，本节使用莫兰指数（Moran's I）对上文测算的财政压力的相关指标（包括债务率、PPP 负担率、财政压力值、违约距离）进行空间相关性检验。① 根据地理学第一定律，距离较近的地区关联更为密切，因此首先构建空间权重矩阵 W，其上的元素为：

$$w_{ij}=\frac{1}{d_{ij}^2} \tag{8}$$

其中，d_{ij}代表城市 i 与城市 j 间的地理距离。在此设定下，相距较远的城市其所对应的空间矩阵元素相对较小，由此可识别邻近地区间的相关关系。对空间权重矩阵进行标准化后，莫兰指数的计算公式可表示为：

$$I=\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{9}$$

其中，x 表示城市的财政压力指标。若 I 为正值，代表相邻地区间财政压力呈现空间正相关，值越大代表相关性越强；若 I 为负值，代表相邻地区间财政压力呈现空间负相关，值越小则差异越大；若 I 显著为 0，则表示无空间相关性。使用 STATA16.0 计算的莫兰指数值见表 3。

表 3 2015—2019 年莫兰指数测算值

年份	债务率	PPP 负担率	财政压力值	违约距离
2014 年	0.224***	−0.013	0.224***	0.036**
2015 年	0.231***	0.008	0.230***	0.029**
2016 年	0.250***	0.035**	0.232***	0.009
2017 年	0.236***	0.082***	0.214***	0.051***
2018 年	0.157***	0.093***	0.164***	0.165***
2019 年	0.184***	0.099***	0.185***	0.178***

注：***、**、*分别表示 1%、5%、10%的显著性水平。

① 基于上节的测算，债务率为债务余额与综合财力的比值，PPP 负担率为 PPP 财政支出责任现值与综合财力的比值，财政压力值为债务率与 PPP 负担率之和。

由表 3 可知，城市层面的四类指标均呈现出显著的空间正相关，说明无论是债务投资还是 PPP 投资，城市地方政府在进行投资决策时，不仅考虑到自身的财政承受水平，也考虑到周边相邻城市的影响，呈现近邻模仿效应。其中，由于 PPP 在进入运营阶段之前需要约两年的谈判、建设期，前期财政支出责任较小，自 2016 年才开始展现出空间相关性特征。此外，城市违约距离展现出的空间相关性，反映了我国城市间的财政风险具有集聚性特征，值得加以关注。为了清晰展示城市间财政压力空间相关性的特征，绘制局部莫兰散点图（图 4）。如图 4 所示，四类指标的散点均大多集中于一三象限，说明呈现显著的正向空间相关性，财政风险较大的城市往往也与财政风险较大的城市相邻。

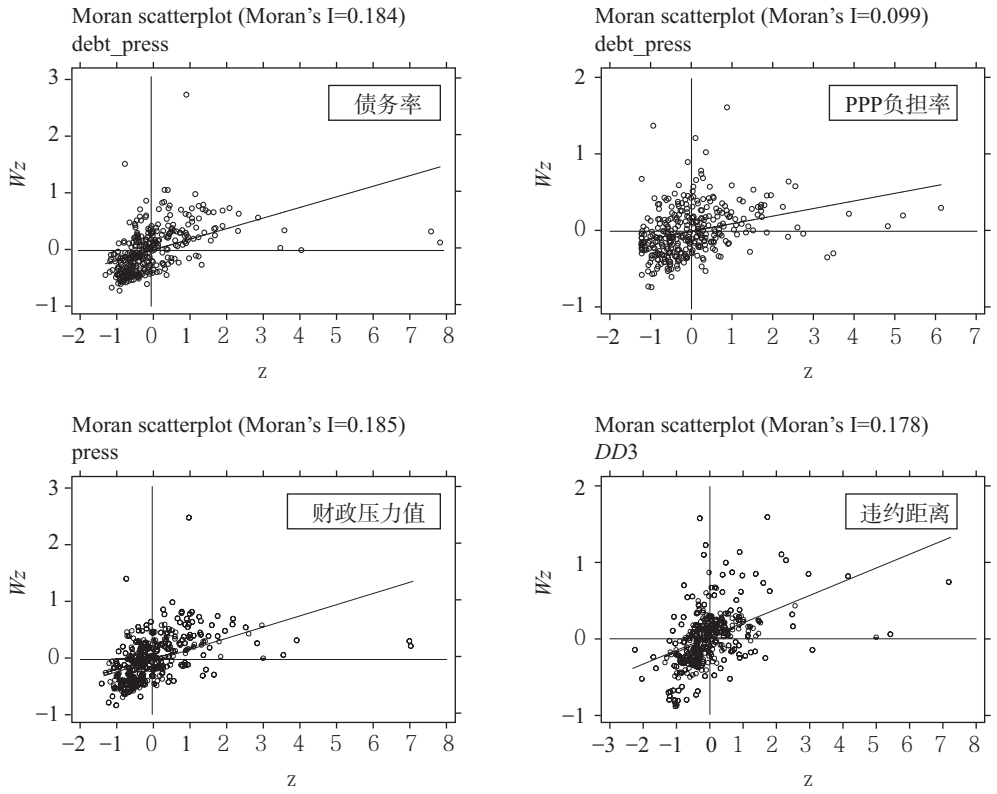

图 4　莫兰散点图（2019 年）

（二）空间集聚性分析

由于我国地域特征明显，不同城市财政压力的空间特征必然存在异质性。尽管使用莫兰指数可以衡量整体上空间相关性的存在，但却难以识别空间的局部特征，无法从城市层面做出深入分析。因此本节进一步引入 Ord 等 Getis（1995）[31] 构建的 Getis-Ord 统计量 G_i^*，用于衡量局部空间集聚。这一统计量的优势在于能够识别城市中的热点（hot spots）与冷点（cold spots），从而描述空间异质性。其计算公式为：

$$G_i^* = \frac{\sum_{j=1}^{N} w_{ij}(d)\, y_j}{\sum_{j=1}^{N} y_j} \tag{10}$$

其中，i 与 j 代表城市，d 表示地理距离，$w_{ij}(d)$ 为根据距离规则定义的空间权重矩阵元素，y_j 为 j 城市的观测值。根据城市的经纬度坐标计算地理距离，发现我国城市间的平均距离约为 1 423 公里，考虑到城市集聚的辐射范围，将 d 设定为 200 公里。因此 $w_{ij}(d)$ 可表示为：

$$w_{ij}(d) = \begin{cases} 1 & \text{若 } d_{ij} \leqslant 200 \text{ 公里} \\ 0 & \text{若 } d_{ij} > 200 \text{ 公里} \end{cases} \tag{11}$$

在式（10）中，分子表示距离 i 市中心点 200 公里半径范围内的数值之和，分母表示所有城市的数值之和。因此，这一统计量测算了与所有地区相比，i 市及其周边城市（即位于 d 公里半径范围内的城市）是否具有相对较大或较小的比率。若该统计量在 5% 的显著性水平上偏高（偏低），则该地区被确定为热（冷）点。由此可对 337 座城市进行财政风险的空间集聚识别，识别结果见表 4。

表 4 热点与冷点城市识别① （单位：座）

指标	2014 年	2015 年	2016 年	2017 年	2018 年	2019 年	持续五年	最近两年
热点城市数量								
债务率	41	51	52	51	35	44	8	34
PPP 负担率	13	29	27	36	29	31	0	27
财政压力值	40	52	58	52	34	39	11	33
违约距离	33	26	12	21	54	48	0	40
冷点城市数量								
债务率	50	49	62	69	38	53	27	38
PPP 负担率	0	0	0	6	13	13	0	11
财政压力值	51	47	50	58	45	53	22	45
违约距离	0	0	5	9	19	28	0	19

如表 4 所示，从持续五年的时间维度来看，有 11 座城市在财政压力值指标下连续成为热点城市，同时有 22 座城市持续成为冷点城市，说明已形成了持续性的财政高/低风险区域（见表 5）。② 观察表 5 可发现，热点城市主要来自内蒙古、湖南、贵州、云南、宁夏所在的西部、中部地区，而冷点城市主要来自上海、江苏、浙江、安徽、河南、广东所在的长三角、珠三角地区。可见某些地区已累积了与经济发展水平不相匹配的财政压力，并呈现持续性的空间集聚特征，蕴含较大的财政风险。

表 5 财政压力值热/冷点城市（持续五年）

类型	城市名称
持续五年财政压力值热点城市	内蒙古包头、内蒙古乌兰察布、湖南张家界、湖南怀化、湖南湘西州、贵州铜仁、贵州黔东南州、云南临沧、云南德宏州、宁夏吴忠、宁夏中卫
持续五年财政压力值冷点城市	上海、江苏南京、江苏无锡、江苏常州、江苏苏州、江苏南通、江苏扬州、江苏镇江、江苏泰州、浙江嘉兴、浙江湖州、安徽马鞍山、安徽滁州、河南开封、河南鹤壁、河南新乡、河南许昌、河南漯河、河南周口、广东深圳、广东珠海、广东惠州

此外，由于 2017 年以前城市 PPP 的投资较为分散，因此在表 3 持续五年的时间维度中，PPP 负担率指标下并未出现热/冷点城市。但随着 PPP 模式的推广，近两年已有 27 座城市持续成为热点城市，表现出财政风险的空间集聚趋势（见表 6）。观察表 6 可发现，近两年 PPP 的风险空间集聚城市分布于吉林、湖南、四川、贵州、云南等，主要集中于西部地区，其中数量最多的四川占比达 33.3%。而在债务方面，近两年热点城市主要分布于内蒙古、辽宁、吉林、黑龙江、湖北、湖南、宁夏等，较多集中于东北地区，其中数量最多的辽宁占比达 32.4%。可见城市间在进行 PPP 投资与债务投资的选择中有着一定差异。这可能是因为，PPP 是一种新型的投融资模式，各地市的接受程度与本省的相关配套政策有关，因此忽视 PPP 对地方政府带来的财政压力容易使某些低债务地区的财政风险被低估。

表 6 PPP 负担率与债务率热点城市对比（最近两年）

类型	城市名称
持续两年 PPP 热点城市	吉林长春、吉林吉林、吉林四平、吉林辽源、吉林松原、湖北恩施州、湖南张家界、湖南怀化、湖南湘西州、四川自贡、四川泸州、四川绵阳、四川广元、四川乐山、四川南充、四川宜宾、四川广安、四川巴中、贵州遵义、贵州毕节、贵州铜仁、云南保山、云南临沧、云南大理州、云南德宏州、云南怒江州、陕西汉中

① 对于债务率、PPP 负担率、财政压力值三组数据，数值越大说明财政支出压力越大，因此通过热点城市识别高风险区域。而对于违约距离数据，数值越大说明违约风险越小，因此通过冷点城市识别高风险区域。

② 由于债务率指标下持续五年的热/冷点城市与财政压力值结果类似，因此不在此展示。

续前表

类型	城市名称
持续两年债务热点城市	内蒙古呼和浩特、内蒙古包头、内蒙古乌海、内蒙古乌兰察布、辽宁沈阳、辽宁大连、辽宁鞍山、辽宁抚顺、辽宁本溪、辽宁丹东、辽宁锦州、辽宁营口、辽宁辽阳、辽宁盘锦、辽宁铁岭、吉林辽源、吉林通化、吉林白山、吉林延边州、黑龙江鹤岗、黑龙江牡丹江、湖北宜昌、湖北恩施州、湖南张家界、湖南怀化、湖南湘西州、重庆、贵州铜仁、陕西安康、甘肃白银、宁夏银川、宁夏石嘴山、宁夏吴忠、宁夏固原

在违约距离方面，由于地方政府债券在 2017 年以前还未大规模进入兑付阶段，因此并没有持续五年的集聚城市出现。然而统计最近两年的持续冷点城市可发现，违约距离较近的城市几乎全部来自于东北地区。其中，辽宁城市占比达 73.7%，可见东北地区当年的财政收入用于支付当年的 PPP 财政支出责任与地方政府债券本息已有较大压力，且在空间集聚方面已显示出一定的持续性，面临较大的财政风险。

表 7　　违约距离冷点城市（最近两年）

类型	城市名称
持续两年违约距离冷点城市	辽宁沈阳、辽宁大连、辽宁鞍山、辽宁抚顺、辽宁本溪、辽宁丹东、辽宁锦州、辽宁营口、辽宁阜新、辽宁辽阳、辽宁盘锦、辽宁铁岭、辽宁朝阳、辽宁葫芦岛、吉林四平、吉林辽源、吉林通化、宁夏银川、宁夏吴忠

基于上述分析可初步推断，财政风险在区域内的空间集聚可能与地方政府行为的策略互动机制有关。策略互动的具体表现形式一般包括地方政府之间的学习、模仿、竞争机制（Marsh 和 Sharman，2009[32]），其中学习指参考其他地区成功经验的行为，模仿往往有着更多的非理性元素，竞争则重点在于经济层面的互动。根据相关实证研究，我国地方政府之间的竞争往往在同一省份的城市之间更为敏感（Yu 等，2016[33]），官员对政绩的追求可能是影响财政风险集聚的重要因素（祝志勇和高扬志，2010[34]）。因此，同一省内城市财政风险的大规模集聚在某种程度可解释为省内地方政府之间的非良性竞争，调整对地方政府的激励手段可能是缓解财政压力的根本性举措。

五、结论与政策含义

伴随着我国外部环境不确定性的加剧与内部经济增速的放缓，进一步防范与化解重大风险成为当务之急。根据对财政压力的新口径测度以及财政压力空间特征的分析，本文发现：第一，我国城市层面的财政压力正在逐年加大，其中地方政府债券的兑付与 PPP 的财政支出责任支出是财政压力的主要来源。第二，城市间的财政压力水平分布呈现显著差异，对于部分东北部、中西部城市，其违约风险已经远远超出平均水平，伴随新冠肺炎疫情的持续冲击，其当期的偿还能力将受到严峻考验。第三，由于地方政府行为具有策略互动性，城市财政压力呈现显著的空间集聚性特征，这将导致持续性风险高发地区的形成，严重制约当地的经济发展与民生保障。

基于研究结果，本文提出如下建议：

第一，加强对 PPP 财政支出责任的风险监测。尽管 PPP 与债务有着本质不同，但其带来的财政偿付压力是类似的，值得重点关注。当前对 PPP 的风险监测仍主要集中于流量层面，且存在一定的主观因素（如部分地区高估了未来财政支出增长率），难以完整反映 PPP 带来的财政压力。PPP 存量层面的统计需要综合全部 PPP 项目的信息，对数据的要求更高。随着政府信息开放程度的增加，建议对地方政府 PPP 财政支出责任的风险监测可进一步延展至存量层面，如引入 PPP 财政支出责任与债务的综合统计指标，将 PPP 项目未来支出责任的累计现值在政府预决算报告中予以公布，从而完整反映地方政府的财政压力。这种统计方式可能存在的难点在于：未来 PPP 项目的支出责任难以预测，根据现有政策，PPP 支出责任应当与当年社会资本的运营绩效直接挂钩，因此与债务偿付不同，真实的 PPP 未来支出责任是一个未知数。然而，根据前期估算的政府支出责任以及项目当前的运营状况，仍然能够近似测度 PPP 支出责任的区间范围，从而给出具有参考价值的存量信息。唯有尽可能充分地掌握 PPP 存量信息，才能全面地监测地区财政风险，并为地方政府的中长期财政规划提供参考依据。

第二，建议充分关注财政压力较大城市所面临的财政风险。在一系列重大公共危机的冲击下，部分城

市面临较大的违约风险。由于城市间存在空间上的相互影响，财政风险的累积不仅可能造成城市自身收不抵支，还可能使得财政风险在区域内蔓延，形成持续的财政风险集聚地区。因此，防范财政风险应当尽可能关注财政压力较大城市，并及时施以相应的措施。建议通过转移支付等手段加强对财政风险高发地区的应急性救援，兜实“三保”底线。对地方政府自身而言，可在短期内通过存量资产处置、削减非必要支出等缓解收支矛盾，并在长期通过土地规划、混合所有制改革、平台转型等方式提升城市经营能力，实现从规划到建设的良性循环。

第三，从体制机制的角度调整激励方式，打破财政风险的持续性聚集。由于地方政府面临上级政府的激励，因此城市之间（特别是同一省份内城市）存在策略互动，导致地方政府可能忽视自身条件进行非理性投资，形成区域内财政风险的聚集。建议进一步完善官员考核制度，如通过引入财政压力的综合测算作为考核指标，将财政风险作为评价地区政绩的参考依据，从而遏制地方政府的投资冲动。此外，绩效管理作为当前财政体制改革的重点内容，可在制度激励层面发挥重要作用。例如，通过在绩效评价中对财政资金使用效益进行科学评估，并纳入当地居民的满意度，实现多方协同治理，保障财政资金的合理使用。实践中，建议重点关注绩效评价指标的科学有效性，避免绩效评价流于形式，真正使绩效管理起到对地方政府的激励与约束作用。

本文的不足在于，财政压力值的测度未能考量地方政府的隐性债务负担，如融资平台债务等。当前唯一较为透明的数据为城投债数据，但城投债如何进行界定至今仍未明确（曹婧等，2019[35]），且由于融资平台融资的形式较多（如银行贷款、资产证券化等），与地方政府的财务关系较难厘清，暂时无法获得较为完整与准确的统计。财政压力是识别区域财政风险的重要依据，对财政压力的度量方式需要不断与时俱进以适应新的客观现实。随着政府会计权责发生制改革的推进、预算透明度的不断提高，对财政压力的度量也应不断加以完善，从而为长远规划提供决策依据，有效防范与化解地方政府财政风险。

附录

2019 年城市财政压力值排名（由高到低排列）

排名	区域	省份	城市	排名	区域	省份	城市	排名	区域	省份	城市
1	中部	湖北	恩施州	16	东北	吉林	延边州	31	西部	内蒙古	包头
2	西部	宁夏	石嘴山	17	东部	广东	肇庆	32	西部	贵州	毕节
3	西部	宁夏	固原	18	东北	吉林	松原	33	西部	内蒙古	呼伦贝尔
4	西部	陕西	商洛	19	西部	内蒙古	阿拉善盟	34	中部	湖南	张家界
5	西部	内蒙古	乌兰察布	20	西部	内蒙古	通辽	35	东北	吉林	四平
6	东北	吉林	白山	21	西部	新疆	克孜勒苏州	36	西部	贵州	黔南州
7	西部	云南	保山	22	西部	四川	巴中	37	西部	贵州	贵阳
8	西部	贵州	铜仁	23	西部	青海	海东	38	东北	辽宁	铁岭
9	中部	湖南	湘西州	24	东北	黑龙江	伊春	39	西部	内蒙古	乌海
10	东北	吉林	辽源	25	西部	新疆	博尔塔拉州	40	东北	辽宁	盘锦
11	东北	黑龙江	牡丹江	26	西部	四川	广元	41	东北	辽宁	抚顺
12	东北	黑龙江	佳木斯	27	西部	甘肃	白银	42	西部	广西	河池
13	西部	四川	资阳	28	西部	陕西	安康	43	中部	湖北	咸宁
14	西部	贵州	六盘水	29	东北	辽宁	鞍山	44	东北	辽宁	营口
15	西部	内蒙古	兴安盟	30	西部	宁夏	中卫	45	东北	吉林	白城

续前表

排名	区域	省份	城市	排名	区域	省份	城市	排名	区域	省份	城市
46	西部	云南	玉溪	81	西部	内蒙古	鄂尔多斯	116	东部	海南	海口
47	西部	广西	来宾	82	西部	甘肃	平凉	117	西部	广西	钦州
48	西部	内蒙古	锡林郭勒盟	83	西部	陕西	铜川	118	西部	四川	宜宾
49	东北	黑龙江	哈尔滨	84	西部	新疆	阿勒泰地区	119	中部	安徽	滁州
50	东部	海南	儋州	85	西部	贵州	黔东南州	120	东北	辽宁	大连
51	西部	甘肃	陇南	86	东北	辽宁	阜新	121	东部	河北	秦皇岛
52	西部	陕西	汉中	87	中部	湖南	怀化	122	西部	四川	遂宁
53	中部	湖南	郴州	88	东北	辽宁	锦州	123	西部	广西	崇左
54	西部	甘肃	定西	89	东北	黑龙江	七台河	124	西部	甘肃	庆阳
55	西部	云南	德宏州	90	中部	湖南	益阳	125	中部	湖北	随州
56	东北	黑龙江	双鸭山	91	中部	湖南	常德	126	东部	江苏	宿迁
57	中部	湖北	黄冈	92	东北	黑龙江	鹤岗	127	中部	山西	阳泉
58	中部	湖北	十堰	93	中部	江西	萍乡	128	西部	云南	怒江州
59	西部	云南	临沧	94	东北	吉林	吉林	129	东部	福建	龙岩
60	西部	云南	昭通	95	西部	四川	内江	130	中部	湖南	永州
61	东北	吉林	长春	96	东部	福建	三明	131	中部	湖南	邵阳
62	东北	辽宁	丹东	97	东部	广东	潮州	132	西部	贵州	黔西南州
63	中部	湖南	娄底	98	中部	安徽	铜陵	133	西部	内蒙古	巴彦淖尔
64	西部	贵州	遵义	99	东部	广东	阳江	134	东部	广东	河源
65	西部	云南	普洱	100	东部	广东	汕头	135	中部	安徽	安庆
66	中部	江西	新余	101	西部	四川	南充	136	中部	山西	运城
67	东北	辽宁	葫芦岛	102	东北	辽宁	辽阳	137	西部	甘肃	张掖
68	西部	云南	丽江	103	中部	湖北	鄂州	138	西部	新疆	吐鲁番
69	东北	吉林	通化	104	西部	云南	曲靖	139	中部	安徽	六安
70	西部	宁夏	吴忠	105	中部	湖南	湘潭	140	西部	四川	攀枝花
71	中部	湖北	黄石	106	东部	广东	梅州	141	西部	四川	达州
72	西部	贵州	安顺	107	西部	甘肃	临夏州	142	东北	黑龙江	鸡西
73	西部	广西	百色	108	东部	福建	南平	143	西部	甘肃	酒泉
74	西部	新疆	乌鲁木齐	109	中部	山西	临汾	144	中部	山西	大同
75	西部	甘肃	武威	110	西部	新疆	阿克苏地区	145	西部	广西	贺州
76	中部	安徽	池州	111	西部	内蒙古	呼和浩特	146	西部	广西	玉林
77	西部	新疆	哈密	112	中部	山西	忻州	147	西部	云南	文山州
78	东北	辽宁	本溪	113	西部	云南	迪庆州	148	西部	四川	自贡
79	西部	陕西	渭南	114	东部	福建	宁德	149	西部	广西	防城港
80	东部	福建	莆田	115	西部	云南	昆明	150	西部	四川	凉山州

续前表

排名	区域	省份	城市	排名	区域	省份	城市	排名	区域	省份	城市
151	西部	四川	雅安	186	西部	甘肃	兰州	221	东部	江苏	盐城
152	西部	陕西	咸阳	187	中部	安徽	黄山	222	西部	新疆	昌吉州
153	东北	辽宁	朝阳	188	东部	广东	汕尾	223	中部	河南	南阳
154	东部	福建	泉州	189	中部	河南	驻马店	224	东部	广东	茂名
155	东部	广东	云浮	190	西部	四川	绵阳	225	东部	江苏	扬州
156	西部	云南	大理州	191	中部	河南	安阳	226	西部	云南	红河州
157	西部	陕西	延安	192	西部	内蒙古	赤峰	227	中部	河南	开封
158	西部	陕西	西安	193	西部	甘肃	甘南州	228	西部	四川	泸州
159	中部	山西	吕梁	194	西部	四川	阿坝州	229	东部	山东	枣庄
160	中部	湖北	孝感	195	西部	四川	甘孜州	230	西部	四川	德阳
161	东北	黑龙江	齐齐哈尔	196	西部	广西	贵港	231	东部	山东	日照
162	西部	新疆	巴音郭楞州	197	中部	安徽	亳州	232	中部	江西	九江
163	西部	云南	西双版纳州	198	西部	青海	果洛州	233	东部	浙江	衢州
164	东部	广东	韶关	199	西部	甘肃	嘉峪关	234	中部	河南	信阳
165	西部	四川	乐山	200	东部	江苏	镇江	235	东部	江苏	淮安
166	东北	黑龙江	大兴安岭	201	中部	河南	平顶山	236	西部	青海	海西州
167	东部	山东	临沂	202	西部	青海	海北州	237	中部	河南	新乡
168	中部	江西	鹰潭	203	西部	广西	梧州	238	东部	河北	保定
169	西部	广西	桂林	204	西部	青海	西宁	239	西部	青海	玉树州
170	西部	云南	楚雄州	205	西部	广西	柳州	240	东部	山东	泰安
171	中部	湖北	荆州	206	中部	安徽	宣城	241	中部	河南	周口
172	东部	海南	三亚	207	中部	湖南	株洲	242	东部	山东	济宁
173	西部	青海	海南州	208	西部	四川	广安	243	东部	山东	德州
174	中部	河南	濮阳	209	东部	广东	湛江	244	西部	陕西	宝鸡
175	西部	新疆	克拉玛依	210	中部	河南	鹤壁	245	中部	河南	漯河
176	西部	甘肃	金昌	211	中部	山西	晋中	246	中部	山西	晋城
177	东部	福建	漳州	212	东部	山东	滨州	247	东部	河北	唐山
178	中部	安徽	宿州	213	中部	河南	洛阳	248	东部	山东	东营
179	中部	安徽	阜阳	214	中部	湖北	襄阳	249	西部	广西	南宁
180	中部	湖南	衡阳	215	东部	广东	清远	250	东部	山东	聊城
181	中部	山西	朔州	216	中部	湖北	宜昌	251	中部	湖南	长沙
182	中部	山西	长治	217	西部	甘肃	天水	252	西部	青海	黄南州
183	西部	新疆	和田地区	218	西部	广西	北海	253	东部	天津	天津
184	中部	安徽	芜湖	219	东部	广东	揭阳	254	东部	河北	邢台
185	中部	安徽	淮南	220	东北	辽宁	沈阳	255	东部	浙江	丽水

续前表

排名	区域	省份	城市	排名	区域	省份	城市	排名	区域	省份	城市
256	西部	重庆	重庆	284	东北	黑龙江	绥化	312	中部	江西	南昌
257	中部	河南	焦作	285	西部	新疆	伊犁州	313	西部	西藏	那曲
258	东部	浙江	台州	286	中部	山西	太原	314	西部	陕西	榆林
259	中部	湖北	荆门	287	西部	宁夏	银川	315	东部	广东	佛山
260	中部	河南	三门峡	288	东部	浙江	绍兴	316	中部	江西	景德镇
261	东部	广东	惠州	289	东部	江苏	徐州	317	东部	浙江	金华
262	中部	湖南	岳阳	290	西部	四川	眉山	318	中部	安徽	合肥
263	东部	河北	邯郸	291	东部	浙江	温州	319	西部	西藏	阿里地区
264	中部	湖北	武汉	292	中部	江西	吉安	320	东部	福建	福州
265	东部	河北	衡水	293	东部	广东	东莞	321	东部	北京	北京
266	中部	江西	抚州	294	东部	山东	莱芜	322	东部	江苏	无锡
267	东部	江苏	连云港	295	东部	山东	青岛	323	中部	江西	宜春
268	西部	四川	成都	296	东部	河北	承德	324	中部	江西	上饶
269	东部	广东	中山	297	东部	江苏	南通	325	东部	浙江	杭州
270	东部	河北	石家庄	298	东部	浙江	湖州	326	东部	江苏	常州
271	中部	河南	商丘	299	东部	江苏	南京	327	西部	西藏	日喀则
272	东部	山东	烟台	300	西部	新疆	喀什地区	328	东部	福建	厦门
273	东北	黑龙江	黑河	301	中部	河南	郑州	329	东部	山东	济南
274	东部	河北	张家口	302	东部	浙江	嘉兴	330	东部	上海	上海
275	中部	安徽	蚌埠	303	中部	安徽	淮北	331	西部	西藏	昌都
276	东北	黑龙江	大庆	304	中部	安徽	马鞍山	332	东部	河北	廊坊
277	东部	江苏	泰州	305	东部	浙江	宁波	333	西部	西藏	山南
278	东部	山东	威海	306	东部	广东	广州	334	西部	西藏	林芝
279	东部	浙江	舟山	307	中部	河南	许昌	335	东部	江苏	苏州
280	东部	山东	淄博	308	东部	河北	沧州	336	西部	西藏	拉萨
281	东部	山东	潍坊	309	东部	广东	珠海	337	东部	广东	深圳
282	东部	广东	江门	310	东部	山东	菏泽				
283	西部	新疆	塔城地区	311	中部	江西	赣州				

参考文献

[1] 罗必良. 分税制、财政压力与政府“土地财政”偏好 [J]. 学术研究，2010（10）：27-35.

[2] 王雅龄，王力结. 地方债形成中的信号博弈：房地产价格——兼论新预算法的影响 [J]. 经济学动态，2015（4）：59-68.

[3] Han L，Kung J K. Fiscal Incentives and Policy Choices of Local Governments，Evidence from China [J]. Journal of Development Economics，2015，9：89-104.

[4] Chen S X. The Effect of a Fiscal Squeeze on Tax Enforcement：Evidence from a Natural Experiment in China [J]. Journal of Public Economics，2017：62-76.

[5] 曹春方，马连福，沈小秀. 财政压力、晋升压力、官员任期与地方国企过度投资 [J]. 经济学（季刊），2014（4）：1415-1436.

[6] 刘尚希. 中国财政风险的制度特征："风险大锅饭" [J]. 管理世界，2004（5）：39-44.
[7] 陈晓光. 财政压力、税收征管与地区不平等 [J]. 中国社会科学，2016（4）：53-70.
[8] 孙开，张磊. 分权程度省际差异、财政压力与基本公共服务支出偏向——以地方政府间权责安排为视角 [J]. 财贸经济，2019：18-32.
[9] 毛捷，韩瑞雪，徐军伟. 财政压力与地方政府债务扩张——基于北京市全口径政府债务数据的准自然实验分析 [J]. 经济社会体制比较，2020（1）：22-33.
[10] 亓寿伟，毛晖，张吉东. 财政压力、经济刺激与以地引资——基于工业用地微观数据的经验证据 [J]. 财贸经济，2020（4）：20-34.
[11] 朱军，寇方超，宋成校. 中国城市财政压力的实证评估与空间分布特征 [J]. 财贸经济，2019（12）：20-34.
[12] 席鹏辉，梁若冰，谢贞发，苏国灿. 财政压力、产能过剩与供给侧改革 [J]. 经济研究，2017（9）：86-102.
[13] Osei-Kyei R，Chan A P C. Comparative Study of Governments' Reasons/Motivations for Adopting Public-Private Partnership Policy in Developing and Developed Economies/Countries [J]. International Journal of Strategic Property Management，2018，5：403-414.
[14] Buso M，Marty F，Phuong T T. Public-Private Partnerships from Budget Constraints：Looking for Debt Hiding? [J]. International Journal of Industrial Organization，2017：56-84.
[15] Oliver H. Incomplete Contracts and Public Ownership：Remarks，and an Application to Public-Private Partnerships [J]. The Economic Journal，2003：69-76.
[16] Hodge G A，Greve C. Public-Private Partnerships：An International Performancereview [J]. Public Administration Review，2007，3：545-558.
[17] 马恩涛，李鑫. PPP 政府或有债务风险管理：国际经验与借鉴 [J]. 财政研究，2018（5）：35-45.
[18] 谭艳艳，邹梦琪，张悦悦. PPP 项目中的政府债务风险识别研究 [J]. 财政研究，2019（10）：47-57.
[19] Wang L，Zhou L，Xiong Y，Yan D. Effect of Promotion Pressure and Financial Burden on Investment in Public-private Partnership Infrastructure Projects in China [J]. Asian-Pacific Economic Literature，2019，2：128-142.
[20] 陈少强. 完善 PPP 财政风险管理研究——基于流量管理和存量管理相结合的视角 [J]. 中央财经大学学报，2018（12）：3-13.
[21] 张牧扬，卢小琴，汪峰. 地方财政能够承受起 PPP 项目财政支出责任吗？——基于 2010—2018 年 PPP 项目的分析 [J]. 财政研究，2019（8）：49-59.
[22] 魏蓉蓉，李天德，邹晓勇. 我国地方政府 PPP 隐性债务估算及风险评估——基于空间计量和 KMV 模型的实证分析 [J]. 社会科学研究，2020（2）：66-74.
[23] 刁伟涛，傅巾益. 我国县级政府债务风险的分类度量、区域分布和变化特征：2015—2017 [J]. 财政研究，2019（5）：58-76.
[24] 毛捷，黄春元. 地方债务、区域差异与经济增长——基于中国地级市数据的验证 [J]. 金融研究，2018（5）：1-19.
[25] 刘尚希，赵全厚. 政府债务：风险状况的初步分析 [J]. 管理世界，2002（5）：22-32.
[26] 王学凯，黄瑞玲. 基于 KMV 模型的地方政府性债务违约风险分析——以长三角地区为例 [J]. 上海经济研究，2015（4）：62-69.
[27] 王周伟，赵启程，李方方. 地方政府债务风险价值估算及其空间效应分解应用 [J]. 中国软科学，2019（12）：81-95.
[28] 蒋忠元. 地方政府债券发行过程中的信用风险度量和发债规模研究——基于 KMV 模型分析江苏省地方政府债券 [J]. 经济研究导刊，2011（19）：61-62.
[29] 李腊生，耿晓媛，郑杰. 我国地方政府债务风险评价 [J]. 统计研究，2013（10）：30-39.
[30] 刁伟涛. 我国省级地方政府间举债竞争的空间关联性研究 [J]. 当代财经，2016（7）：36-45.
[31] Ord J K，Getis A. Local Spatial Autocorrelation Statistics：Distributional Issues and anApplication [J]. Geographical Analysis，1995，4：286-306.
[32] Marsh D，Sharman J C. Policy Diffusion and Policytransfer [J]. Policy Studies，2009，3：269-288.
[33] Yu J，Zhou L，Zhu G. Strategic Interaction in Political Competition：Evidence from Spatial Effects Across Chinese Cities [J]. Regional Science and Urban Economics，2016：23-37.
[34] 祝志勇，高扬志. 财政压力与官员政绩的牵扯：细究地方政府投融资平台 [J]. 改革，2010（12）：30-35.
[35] 曹婧，毛捷，薛熠. 城投债为何持续增长：基于新口径的实证分析 [J]. 财贸经济，2019（5）：5-22.

（责任编辑：孙亦军　张安平）

信息透明度、银行挤兑与风险传染
——基于实验经济学的实证研究

Information Transparency, Bank Run and Risk Contagion: Empirical Study Based on Experimental Economics

王丽珍　张简荻　陈　华
WANG Li-zhen　ZHANG Jian-di　CHEN Hua

［摘　要］针对银行挤兑风险的研究主要聚焦于单一银行内部客户的取款行为，或者两个相同规模银行的挤兑风险传染问题。本文基于 DD 模型设计不同信息透明度下存在两个不同规模银行主体的经济学实验，研究银行保险制度背景下银行挤兑风险从大银行传染至小银行以及从小银行传染至大银行的潜在原因和传染路径。研究表明：当信息透明度较高时，大银行客户倾向于关注银行基本面信息，以银行流动性水平作为取款动机的决策依据，而小银行客户的取款行为则主要取决于与银行其他存款人的内部协调；当信息透明度不足时，银行存款人易受外界信息干扰而出现恐慌挤兑行为，无论大小银行是否存在关联，挤兑风险由小银行传染至大银行的可能性较低，但由大银行传染至小银行的可能性很高；当信息透明度极低时，银行存款人的利益很可能受到损害。本研究为银行“羊群效应”可能引发的系统性风险防范与化解提供了理论借鉴。

［关键词］信息透明度　银行规模　银行挤兑　风险传染　实验经济学
［中图分类号］F832.5　［文献标识码］A　［文章编号］1000-1549（2020）10-0026-10

Abstract: The research on bank run mainly focuses on the withdrawal behavior of customers in the same bank, or the contagion risk between banks with the same size. Using experimental economics method based on DD model of banks with different size and different information transparency, this paper studies the potential causes and contagious mechanisms. The results show that, when the information transparency is high, depositors in big bank intend to care about the bank fundamental and take the liquidity level as a decision basis, while depositors in small bank pay more attention to the coordination of the other depositors. When the information transparency is inadequate, depositors are easy to panicly withdraw owning to the noisy information. No matter the two banks are linked or not, the probability of bank-run contagion from big bank to small bank is high, but from small bank to big bank is low. When the information transparency is quite low, the interests of depositors are likely to be damaged. The conclusion offers important theoretical guidance for the treatment of bank-run risks.

Key words: Information transparency　Bank size　Bank run　Risk contagion　Experimental economics

［收稿日期］　2020-03-19
［作者简介］　王丽珍（通讯作者），女，1984 年 7 月生，中央财经大学保险学院副教授，研究方向为金融系统性风险；张简荻，女，1998 年 1 月生，中央财经大学保险学院硕士研究生，研究方向为金融系统性风险；陈华，1973 年 11 月生，中央财经大学保险学院副教授，研究方向为风险管理与保险。
［基金项目］　国家社会科学基金青年项目“基于区块链技术的小额保险深度精准扶贫机制构建与成效评估研究”（项目编号：19CJL033）。感谢匿名评审人提出的修改建议，笔者已做了相应修改，本文文责自负。

一、引言

一直以来银行挤兑频繁发生，尤其是在金融管制放松、经济全球化进程加快的背景下，银行挤兑现象更是屡见不鲜。例如，2001 年阿根廷发生了严重的银行挤兑，导致银行系统的短暂关闭；2004 年俄罗斯经历了局部银行挤兑；2007 年英国北岩银行遭遇挤兑风潮，最后被政府救助并且国有化；2011 年釜山储蓄银行出现了挤兑高潮。银行挤兑很可能造成银行危机，进而出现银行破产和风险传染，甚至形成系统性风险，波及整个金融市场乃至实体经济。虽然包括我国在内的许多国家尚未发生大规模银行挤兑，但是潜在的挤兑风险仍然困扰着各个国家。存款保险制度的实施能够在很大程度上保证存款的安全，然而信息不对称条件下的趋同性、恐慌预期和负反馈等因素可能影响存款人对金融市场的信心，单一事件通过传染效应会产生群体性的挤兑风潮，这种“羊群效应”可能导致系统性风险。与此同时，随着互联网金融的高速发展，互联网金融产品不断涌现，如果由于某些原因，众多用户同时大规模赎回资金，与银行挤兑类似，这势必也会造成风险传染。如何采取有效措施预防和应对这些风险以维护金融体系的稳定性，已经成为各国共同面临的一个棘手问题。

关于银行挤兑的开创性研究成果是 Diamond 和 Dybvig（1983）[1]的相关研究，后期研究大多基于该理论研究框架展开，主要采用的研究方法包括理论研究（Goldstein 和 Pauzner，2005[2]）、实证分析（Iyer 和 Puri，2012[3]）和实验经济学方法（Davis 和 Reilly，2016[4]；Shakina 和 Angerer，2018[5]）。由于理论假设的局限性和实证数据的不可得性，实验经济学方法已经成为当前研究的主流方法。通过开展经济学实验能够较好地模拟并反映银行挤兑这一经济现象，从而借助实验数据分析银行挤兑的潜在原因与传染路径。

采用实验经济学方法对银行挤兑问题的研究最早可以追溯到 21 世纪初。Madies（2006）[6]在存在非挤兑均衡的背景下研究了银行挤兑的广泛性，发现当银行流动性水平较高时，银行挤兑发生概率较低。Garratt 和 Keister（2009）[7]研究了银行挤兑中导致存款人协调失败的因素，研究表明，当银行取款需求已知并且随机、频繁发生时，银行挤兑出现的可能性极低；当允许多次取款时，存款人更倾向于多次取款而非一次。Kiss 等（2012）[8]、王海巍（2014）[9]验证了存款保险和其他存款人行为的可观察性对银行挤兑的影响。Arifovic 等（2013）[10]则识别得到协调参数的一个不确定区域，研究了银行挤兑中协调需求对协调失败的影响。后期关于银行挤兑的研究则更加具体化。例如：Kiss 等（2014）[11]研究了取款人性别对银行挤兑行为的影响；Dijk（2017）[12]研究了存款人害怕、悲伤、高兴等心理状态对于银行挤兑的影响；Kiss 等（2018）[13]证实了纯粹恐慌导致银行挤兑的可能性；Shakina 和 Angerer（2018）[5]则研究了存款人交流和文化对于银行挤兑的影响。不难发现，以上研究主要基于单一银行展开，主要针对银行特征以及银行内部存款人的行为特点等方面进行分析。

目前基于两个或者更多银行来分析银行挤兑风险跨行传染的研究极少。Chakravarty 等（2014）[14]对两个同等规模银行的挤兑风险传染问题做了研究，研究发现，即使在两个银行相互独立的情况下，银行存款人的取款行为也会受到可观察银行存款人取款行为的显著影响。Brown 等（2017）[15]则研究了银行挤兑风险跨行传染的条件和渠道，研究表明：一个银行的挤兑风险之所以会触发另外一个银行存款人的挤兑行为，主要是因为增加了另一个银行存款人对本银行其他存款人取款可能性的预期；观察到取款行为只会影响存款人的预期，当基于这些预期形成对银行基本面的认识时，挤兑风险才会出现传染。基于银行挤兑跨行传染的研究都是基于相同规模的两个银行展开，但是现实金融市场中银行规模存在差异，研究不同规模银行的挤兑风险传染具有重要的现实意义。

关于银行挤兑导致系统性风险发生的原因，目前主要存在两种对立观点。一种观点认为，银行挤兑是存款人协调失败的结果（Chang 和 Velasco，2001[16]；Martin，2006[17]），存款人是否提前取款主要受其他存款人行为的影响。另一种观点则认为，银行挤兑是银行基本面的负向信息导致的（Jacklin 和 Bhattacharya，1988[18]；Allen 和 Gale，1998[19]；Chen 和 Hasan，2006[20]；Gu，2011[21]），银行资产质量、流动性水平等是存款人决定是否提前取款的主要原因。当然，也有一些学者认为银行挤兑是存款人协调失败和银行基本面信息共同作用的结果（Balkenborg 等，2011[22]）。由于信息不对称，在银行基本面状况既定的情况下，信息披露程度或者信息透明度会对银行挤兑产生直接而重要的影

响。因此，在采用实验经济学方法进行研究时，需要事先给予被试者关于银行的私人或者公共信息，这些信息对于取款决策具有重要的影响。Schotter 和 Yorulmazer（2009）[23]研究了银行挤兑的动态性和严重性，研究发现存在银行危机时，存款人了解的信息越多，他们约束自己不取款的意愿越高。He 和 Manela（2012）[24]研究了当谣言导致具备偿付能力的银行出现挤兑时的信息获取与动态取款决策问题，研究表明，关于流动性水平的私人信息获取能够约束偿付能力充足但流动性不足的银行挤兑，关于偿付能力的公共信息则能够减弱银行挤兑。Chakravarty 等（2014）[14]、Brown 等（2017）[15]则分析了其他银行取款人的行为信息对于本银行存款人的影响。存款人所掌握的信息或者信息透明度显著影响银行存款人的取款行为，是产生银行挤兑的重要原因之一。那么在不同的信息透明度下，尤其是当银行规模存在差异时，银行挤兑产生原因以及挤兑风险传染路径是否存在差异？目前尚没有学者针对性地研究不同信息透明度对银行挤兑造成的具体影响以及传染路径。

本文基于 Diamond 和 Dybvig（1983）[1]针对银行挤兑的奠基性研究成果，借鉴 Brown 等（2017）[15]、Chakravarty 等（2014）[14]的研究范式，采用实验经济学研究方法，以银行挤兑和风险传染为研究对象，以信息透明度和银行规模为研究切入点，将不同信息透明度场景和不同规模银行纳入研究框架，在银行保险制度背景下研究银行挤兑风险从大银行传染至小银行以及从小银行传染至大银行的可能原因和传染路径。主要创新之处体现在三个方面：其一，在不同信息透明度下对取款人的行为做比较分析，验证信息透明度对于银行挤兑以及可能的系统性风险传染的重要影响；其二，分析银行挤兑从大银行传染至小银行以及从小银行传染至大银行的路径差异，揭示银行规模在银行挤兑风险传染中的重要性，为“大而不倒”还是“关联紧密而不倒”提供进一步的证据；其三，已有的针对挤兑风险跨银行传染的研究中，均没有考虑存款保险制度，本文基于存款保险调整支付矩阵中当期取款或者等待期满的最低给付额度，研究结论更贴近现实。本文的研究结论将进一步丰富当前关于银行挤兑的研究成果，完善系统性金融风险的研究框架，尤其对于银行挤兑风险的防范化解具有重要的借鉴价值。

二、实验设计

在 Brown 等（2017）[15]的实验中，构建了只存在两个规模相同的银行——“左银行”和“右银行”的银行系统，通过重复博弈研究了银行存款人之间的风险传染问题。Chakravarty 等（2014）[14]则采用单次博弈对两家规模相同银行的挤兑风险传染做了分析。本文的实验设计借鉴 Brown 等（2016）[15]、Chakravarty 等（2014）[14]的设计思路，并对其进行拓展。在不同的信息透明度下，将规模不同的两家银行纳入实验研究框架，同时考虑存款保险制度，调整支付矩阵中当期取款或者等待期满的最低给付额度，研究存款保险背景下，银行挤兑风险的跨行传染问题。

本实验通过招募、报名、筛选等流程选择了北京某高校 72 名大学生作为被试，在 2017 年年底采用 Z_Tree计算机软件平台分两轮在实验室开展实验。为了避免实验过程的需求效应，并且保证实验结果的无偏性，所选择的被试均来自于不同的专业学院，并且对金融风险存在基本常识性的了解。在实验开始前，实验助理宣读并解释实验规则，被试可以对不明之处进行咨询。为了保证被试能够充分理解实验规则，他们需要准确无误地回答关于实验规则的若干相关问题，针对回答错误的被试，实验助理将对其做进一步的解释，直至所有人充分理解实验规则为止。

为了保证实验的内部有效性，并且实现实验的引致价值（induced-value），在实验过程中，被试每人操作一台电脑，彼此通过隔板隔开，严禁任何形式的交流。为了保证实验结果的广泛性和实证结果的稳健性，将被试分为 8 个不同的组，每组包含 9 个人，其中 3 个属于小银行的客户，6 个属于大银行的客户。在实验开始时，所有的被试被随机指定为大银行或者小银行的客户，并且这一角色在后续的实验中保持不变。为了表示银行的偿付能力，假定银行存在两种流动性水平①，当 L 为 0.7 时，银行资产的流动性水平高、偿付能力强；当 L 为 0.3 时，银行资产流动性水平低、偿付能力弱。银行的流动性水平服从马尔科夫过程，以 1/3 的概率保持不变，以 2/3 的概率转移为

① 此处借鉴 Brown 等（2016）[15]、Chakravarty 等（2014）[14]的术语，可以将其解读为存款准备金比率。

另一种流动性水平①。在实验中具体表现为，后一轮实验银行的流动性水平以 1/3 概率与上一轮实验相同，以 2/3 概率与上一轮实验不同。

被试需要进行不同信息透明度下的三类实验，分别是“领导组”实验、“追随组”实验、和“基准组”实验，每类试验包含 10 轮，总共每人进行了 30 轮试验。不妨假定被试 i 属于大银行 I，与之相对应的小银行记为 J，则在第一个 10 轮实验中，I 是领导组，J 是追随组，而在第二个 10 轮实验中，I 和 J 身份互换，即 J 变为领导组，I 为追随组。在第三个 10 轮实验中，I 和 J 均变为基准组。对于被试 i 而言，其前 10 轮实验属于领导组实验，中间 10 轮为追随组实验，后 10 轮为基准组实验。在领导组实验中，大银行和小银行的存款人被告知该银行的实际流动性水平。在追随组实验中，大银行和小银行的存款人并不知道自己所在银行的流动性水平，但是他们被告知相应领导组银行的总取款人数。在基准组实验中，大银行和小银行的存款人不能获得关于本银行或者其他银行的任何信息。为了进一步得到挤兑风险的具体传染路径，在追随组实验中，设定领导组小银行和追随组大银行之间以及追随组小银行和领导组大银行之间存在关联和独立两种状态。当处于关联状态时，追随组银行的流动性水平不再按照前述的马尔可夫过程转移，而是主要受领导组银行流动性水平的影响。若领导组大银行属于问题银行，它对追随组小银行在 90%水平上产生影响；但是若领导组小银行属于问题银行，它对追随组大银行在 60%水平上产生影响②。当两个银行处于独立状态时，大银行和小银行的流动性水平不受彼此影响。

每轮实验开始时，被试都会被赋予初始效用值 100 作为 0 时刻的存款。在每一类实验中，他们都要根据实验类别和得到的信息做出当期取款或者等待期满的决策。而银行是否有充足的钱来满足取款需求，取决于银行自身的现金流动性和总的当期取款人数，银行通过清算贴现优先满足当期取款人的需求。在计算支付矩阵时，当期取款最多支付效用值 100，如果当期取款人数较多，则银行可能无法同时满足客户的取款请求，此时需要基于银行当时的流动性水平进行计算。借鉴 Chakravarty 等（2014）[14] 的思路，假定存款利率 i 为 50%③，当期取款人数为 n，银行总人数为 N，初始效用值为 100，当期取款时支付额假定为 W，等待期满时支付额假定为 D。目前存款保险制度已经相当普遍，由于存款保险的存在，即使银行出现了破产，存款人仍然能够得到部分或者全部存款。基于此，在支付矩阵中假定最低赔偿的效用值为 30。记 $n/N=r$ 为取款比例，则当期取款或者等待期满支付额的计算公式如式（1）和式（2）所示。

$$W=\begin{cases}100, & r\leqslant L\\ 100\dfrac{L}{r}, & r>L \text{ 且 } 100\dfrac{L}{r}\geqslant 30\\ 30, & r>L \text{ 且 } 100\dfrac{L}{r}<30\end{cases} \tag{1}$$

$$D=\begin{cases}\left(100N-100\dfrac{n}{L}\right)\dfrac{1+i}{N-n}, & r\leqslant L \text{ 且 } \left(100N-100\dfrac{n}{L}\right)\dfrac{1+i}{N-n}\geqslant 30\\ 30, & r>L \text{ 或者 } r\leqslant L \text{ 且 } \left(100N-100\dfrac{n}{L}\right)\dfrac{1+i}{N-n}<30\end{cases} \tag{2}$$

不同银行存款人决定当期取款或者等待期满时，通过式（1）和式（2）计算得到的具体支付矩阵如表 1 所示。在支付矩阵中列出了不同当期取款人数下，

表 1　　支付矩阵

支付矩阵		大银行取款人数							小银行取款人数			
		0	1	2	3	4	5	6	0	1	2	3
低流动性	当期取款		100	90	60	45	36	30		90	45	30
	等待期满	150	80	30	30	30	30		150	30	30	
高流动性	当期取款		100	100	100	100	84	70		100	100	70
	等待期满	150	137	117	85	30	30		150	117	30	

① 在现实情况下，银行状态应该是以更高（低）的概率保持不变（发生变化）。但是在实验中，为了得到银行不同流动性水平对客户行为产生的影响，需要增加银行流动性水平的不确定程度，以体现它所导致的客户取款行为的差异。

② 之所以这样设计主要是因为当大小银行存在关联时，若大银行出现问题，它对小银行的影响会比较大，而当小银行出现问题时，它对大银行的影响会相对较小。

③ 借鉴 Chakravarty 等（2014）[14] 的研究，假定利率较高，主要为了调整支付矩阵。

取款人可以获得的效用值。例如，对于处于高流动性水平下大银行的存款人，当有 3 个人选择当期取款时，选择当期取款的 3 个人可以获得效用值 100，而选择等待期满的另外 3 个存款人则获得效用值 85。存款人依据自己所在银行的情况、所进行的实验类别以及支付矩阵做出最终决策。根据实验结果，一单位效用值对应 0.5 元，每个被试的平均报酬大约为 40 元。

三、均值比较检验

通过实验获取的主要变量的描述性统计结果如表 2 所示。其中：*With* 代表存款人的取款行为，当提前取款时其值为 1，否则为 0；*Num* 代表总取款人数；*Liq* 表示银行流动性水平；*Lin* 代表追随组与领导组银行的关联状态；*Judg*1 是追随组银行对于相应领导组银行流动性水平的判断；*Judg*2 是追随组银行对于相应领导组银行总取款人数的判断。不难发现，不同组别的流动性水平相当，*Liq* 在领导组、追随组和基准组的均值分别是 0.47、0.42 和 0.49，对它们进行信息差异检验发现不存在异质性特征。然而总取款人数和取款行为在不同组别差别较大，这表明不同信息透明度下，存款人的取款行为可能存在差异，后文将通过均值比较检验和多元回归分析进行验证。

表 2　　**主要变量的描述性统计结果**

	变量	*Lin*	*With*	*Judg*1	*Judg*2	*Liq*	*Num*
领导组	均值	0.50	0.50	—	—	0.47	2.66
	标准差	0.50	0.50	—	—	0.50	2.02
	最小值	0.00	0.00	—	—	0.00	0.00
	中位数	0.50	0.00	—	—	0.00	3.00
	最大值	1.00	1.00	—	—	1.00	6.00
追随组	均值	0.50	0.66	3.16	3.58	0.42	3.37
	标准差	0.50	0.48	1.22	1.34	0.49	1.94
	最小值	0.00	0.00	1.00	1.00	0.00	0.00
	中位数	0.50	1.00	3.00	4.00	0.00	3.00
	最大值	1.00	1.00	5.00	5.00	1.00	6.00
基准组	均值	—	0.67	—	—	0.49	3.66
	标准差	—	0.47	—	—	0.50	2.25
	最小值	—	0.00	—	—	0.00	0.00
	中位数	—	1.00	—	—	0.00	4.00
	最大值	—	1.00	—	—	1.00	6.00

为了得到不同信息透明度下存款人取款行为的差异，本部分首先对领导组、追随组和基准组进行均值比较检验，具体结果如表 3 所示。

通过观察表 3 中的取款均值发现，无论是大银行还是小银行的领导组，当流动性水平高时，其取款人数明显少于流动性水平低时的人数；但是基准组和追随组的行为则不能较好地反映银行的流动性水平。尤其是基准组，当银行的流动性水平较低时，存款人反而选择等待期满。可见，信息透明度对存款人行为存在直接的影响，信息透明度低甚至信息误导会损害存款人的利益。同时，对不同组别进行均值比较发现，领导组、基准组和追随组三者之间的取款行为基本上在 5%的水平上存在显著差异。这与直接观察取款均值的结果相吻合，存款人对信息的掌握程度对于银行挤兑存在直接的影响。

但是也存在无显著差异的一些情况。当银行流动性水平较低时，小银行作为领导者或者关联状态以及独立状态的追随者，它们的取款行为不存在显著差异。这一方面表明，大银行作为领导者，其取款行为能够较为真实地将自身的流动性水平反馈给小银行追

表 3 不同信息透明度下取款行为的均值比较检验

组别	领导组				基准组			
银行	大银行		小银行		大银行		小银行	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
流动性	高	低	高	低	高	低	高	低
均值	0.36 (0.03)	0.77 (0.03)	0.23 (0.03)	0.56 (0.06)	0.88 (0.02)	0.77 (0.03)	0.47 (0.05)	0.21 (0.04)
样本数	216	264	168	72	258	222	111	129
组别	追随组							
	关联状态				独立状态			
银行	大银行		小银行		大银行		小银行	
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
流动性	高	低	高	低	高	低	高	低
均值	0.65 (0.04)	0.59 (0.06)	0.46 (0.06)	0.61 (0.07)	0.86 (0.03)	0.67 (0.04)	0.6 (0.06)	0.63 (0.06)
样本数	174	66	63	57	120	120	60	60
T 检验	均值比较							
	(1)~(5)	(2)~(6)	(3)~(7)	(4)~(8)	(5)~(13)	(6)~(14)	(7)~(15)	(8)~(16)
	0	0.88	0	0	0.49	0.03	0.10	0
	(5)~(9)	(6)~(10)	(7)~(11)	(8)~(12)	(1)~(13)	(2)~(14)	(3)~(15)	(4)~(16)
	0	0	0.92	0	0	0.03	0	0.37
	(1)~(9)	(2)~(10)	(3)~(11)	(4)~(12)	—	—	—	—
	0	0	0	0.51	—	—	—	—

注释：均值栏显示的是平均值和标准差，均值比较中显示的是 *T* 检验的 *P* 值。下同。

随者；另一方面，因为独立状态和关联状态下的追随组均与领导组无显著差异，这也意味着小银行存在较强的恐慌挤兑行为，即使在没有任何关联的情况下，受大银行的影响也很大。流动性水平较高时，小银行作为独立状态或者关联状态下的追随者，其行为与基准组小银行的行为无显著差异。在这种情况下，大银行作为领导者没有出现显著的银行挤兑行为，小银行作为追随者，不能通过观察大银行存款人的行为得到明确信息，所以无论是与大银行处于关联状态还是独立状态，均表现出与基准组存款人类似的取款行为。

相对于小银行，大银行存款人的取款行为则更为理性。只有在流动性水平较高时，大银行作为基准组与自身作为独立状态下的追随者才不存在显著差异。而在其他情况下，基于不同的信息透明度，大银行存款人均表现出不同的取款行为，不同信息透明度对大银行存款人的行为影响更大。

领导组存款人能够直接观察到所在银行的流动性水平，追随组则通过观察领导组存款人的行为对其所在银行的流动性水平和其他客户取款的可能性做出预测。通过分析领导组存款人的取款行为我们发现，当银行的流动性水平不同时，领导组银行的取款人数存在明显的阈值。当大银行作为领导组时，观察到两人以下取款条件下流动性水平为 0.8 的概率为 80%，五人以上取款条件下流动性水平为 0.2 的概率为 95%；当小银行作为领导组时，观察到一人以下取款条件下流动性水平为 0.8 的概率为 87%。基于此，下面对领导组大小银行取款人数分组，通过均值比较分析取款人数对追随组产生的影响，结果如表 4 所示。

表 4　　**基于不同领导组行为的追随组取款均值比较检验结果**

状态	关联				不关联			
小银行	(1)	(2)	(5)	(6)	(3)	(4)	(7)	(8)
	$w\geqslant 2$	$W<2$	$W\geqslant 5$	$w<5$	$w\geqslant 2$	$W<2$	$W\geqslant 5$	$w<5$
With	0.666 (0.053)	0.285 (0.07)	0.648 (0.065)	0.439 (0.061)	0.678 (0.051)	0.472 (0.084)	0.75 (0.073)	0.559 (0.054)
*Judg*1	0.519 (0.039)	0.714 (0.054)	0.476 (0.050)	0.678 (0.041)	0.544 (0.026)	0.541 (0.035)	0.458 (0.041)	0.58 (0.023)
*Judg*2	0.673 (0.035)	0.321 (0.047)	0.675 (0.044)	0.446 (0.041)	0.738 (0.028)	0.604 (0.063)	0.756 (0.042)	0.672 (0.034)
N	78	42	54	66	84	36	36	84
小银行	*T* 检验							
	(1)~(2)	(3)~(4)	(5)~(6)	(7)~(8)	(1)~(3)	(2)~(4)	(5)~(7)	(6)~(8)
With	0.000	0.033	0.022	0.05	0.872	0.091	0.311	0.146
*Judg*1	0.004	0.949	0.002	0.008	0.590	0.012	0.792	0.032
*Judg*2	0.000	0.026	0.000	0.164	0.147	0.001	0.215	0.000 0
N	120	120	120	120	162	78	90	150
大银行	关联		不关联		*T* 检验			
	(9)	(10)	(11)	(12)				
	$W>1$	$w\leqslant 1$	$W>1$	$w\leqslant 1$	(9)~(10)	(11)~(12)	(9)~(11)	(10)~(12)
With	0.666 (0.058)	0.620 (0.036)	0.736 (0.041)	0.785 (0.036)	0.51	0.376	0.319	0.002
*Judg*1	0.518 (0.034)	0.642 (0.021)	0.460 (0.026)	0.430 (0.027)	0.002	0.441	0.185	0.000
*Judg*2	0.655 (0.038)	0.587 (0.026)	0.622 (0.028)	0.783 (0.029)	0.166	0.000	0.495	0.000
N	66	174	114	126	240	240	180	300

结果（1）~（8）反映了追随组小银行受领导组大银行的影响情况。通过观察均值不难发现，小银行提前取款人数与大银行取款人数存在正向关系。进一步观察小银行存款人基于大银行取款人数做出的心理预期发现，无论大小银行是否关联，大银行取款人数越多，小银行存款人预期自己所在银行的流动性水平越低，其他人取款的可能性越高，而且该结果在关联条件下更加显著。这表明，小银行受大银行的影响容易产生基于实质信息的或者纯粹的心理恐慌，进而出现银行挤兑，风险由大银行传染至小银行。结果（9）~（12）反映了大银行追随组受小银行领导组的影响情况。结果表明，无论二者是否关联，大银行存款人是否提前取款、对自己所在银行流动性水平的预期以及银行其他人取款可能性的预期基本不受小银行存款人取款行为的影响。

表 4 同时列出了均值比较 *T* 检验的结果。与前述观察取款均值的结果相一致，（1）~（2）、（3）~（4）、（5）~（6）、（7）~（8）的比较结果证实了大银行存款人是否提前取款显著影响小银行存款人的取款行为和心理预期；结果（1）~（3）、（5）~（7）表明当大银行取款人数较多时，大小银行是否处于关联状态对小银行存款人的取款行为不存在显著影响，而（2）~（4）、（6）~（8）则意味着大银行取款人数较少时，小银行存款人会相对较为理性，它们在判断自己银行的流动性水平或者本银行其他客户是否取款时，会考虑是否与大银行之间存在关联；结果（9）~（10）、（11）~（12）、（9）~（11）、（10）~（12）则表明无论是否与小银行存在关联，大银行存款人基本不受小银

行取款行为的影响。可见，规模因素对于风险传染具有非常重要的影响，规模指标无疑在系统重要性判定中具有举足轻重的地位。

四、多元回归分析

为了进一步分析领导组、追随组和基准组取款行为的影响因素及其差异，针对这三个组别分别构建多元回归模型，如式（3）、（4）、（5）所示：

$$With_{it}^{leader}=\gamma_i+\alpha_1 Num_{it-1}^{leader}+\alpha_2 Liq_{it}^{leader}+\alpha_3 Liq_{it-1}^{leader}+\alpha_4 Los_i^{leader}+AX_{it}^{leader}+\varepsilon_{it} \quad (3)$$

$$With_{it}^{baseline}=\lambda_i+\mu_1 Num_{it-1}^{baseline}+\mu_2 Liq_{it}^{baseline}+\mu_3 Liq_{it-1}^{baseline}+\mu_4 Los_i^{baseline}+ZX_{it}^{baseline}+\sigma_{it} \quad (4)$$

$$With_{it}^{follower}=\eta_i+\varphi_1 Num_{it-1}^{follower}+\varphi_2 Liq_{it}^{follower}+\varphi_3 Liq_{it-1}^{follower}+\varphi_4 Los_i^{follower}+\varphi_5 Othnum_{it}+\varphi_6 Lin_{it}^{follower}+BX_{it}^{follower}+\tau_{it} \quad (5)$$

其中：*leader*、*follower* 和 *baseline* 分别代表领导组、追随组和基准组；$With_{it}$代表存款人 i 在第 t 轮的取款行为，当提前取款时其值为 1，否则为零。Num_{it-1}代表存款人 i 所在银行在第 $t-1$ 轮的总取款人数；$Othnum_{it}$表示追随组存款人 i 所对应的领导组银行的总取款人数；Liq_{it}和 Liq_{it-1}分别表示存款人 i 所在银行在第 t 和 $t-1$ 轮的流动性水平；Los_i 代表存款人 i 对损失的态度；Num_{it}代表存款人 i 所在银行在第 t 轮的总取款人数；Lin_{it}代表存款人 i 所在银行在第 t 轮与领导组银行的关联状态；X_{it}代表其他控制变量，包括存款人的性别、收入、家庭背景等；ε_{it}、τ_{it}和 σ_{it}均为扰动项。对模型（3）~模型（5）采用 Probit 回归分析，结果如表 5 所示。

表 5　　不同组别取款行为的影响因素分析

	领导组		追随组		基准组	
	大银行	小银行	大银行	小银行	大银行	小银行
Liq_{it}	1.582*** (8.89)	1.223*** (4.49)	−0.751*** (−3.83)	−0.002 (−0.01)	−0.157 (−0.85)	−0.352 (−1.12)
Liq_{it-1}	−0.730*** (−3.08)	−0.334 (−1.18)	0.704*** (3.58)	0.614*** (3.01)	0.051 (0.28)	−0.242 (−0.78)
Num_{it-1}	0.278*** (4.33)	0.994*** (8.13)	0.264*** (4.08)	0.207* (1.73)	0.235*** (2.99)	1.252*** (7.74)
$Othnum_{it}$			−0.003 (−0.03)	0.183*** (2.94)		
Lin_{it}			0.636* (1.76)	0.074 (0.24)		
Los_i	−0.136* (−1.77)	−0.170 (−1.43)	−0.033 (−0.27)	−0.084 (−0.67)	0.088 (0.98)	−0.294* (−1.80)
_cons	−1.742*** (−2.94)	−1.396* (−1.69)	−0.306 (−0.32)	−1.775** (−1.97)	−1.064 (−1.42)	−1.680 (−1.53)
N	432	216	432	216	432	216

首先分析领导组大银行和小银行的回归结果。在领导组实验中，银行存款人被明确告知自己银行的流动性水平，存款人基于此决定当期取款还是等待期满。当银行流动性水平较低时 Liq_{it}取值为 1，否则为零。所以无论是大银行还是小银行，流动性水平越低，银行的取款人数越多，而且相对于小银行，大银行存款人对银行流动性水平更加敏感。上期取款总人数对于当前存款人的取款行为也存在显著影响，大银行和小银行的回归系数分别为 0.278 和 0.994，小银行上一期的取款人数对本期存款人的影响更大。由此可见，领导组银行在选择是否取款时，均会考虑自身的流动性水平和银行其他存款人的取款行为。但是相对而言，大银行存款人更为理性，比较重视自身银行的流动性水平，而小银行存款人则比较关注与其他存款人的内部协调。

其次分析追随组大银行和小银行的回归结果。在追随组实验中，银行存款人被告知自身所在银行上一期的流动性水平和取款人数、与领导组银行之间的关联或者独立状态以及领导组银行当期的取款人数。与预期结果相一致，根据表 5 中针对追随组的回归结

果，无论大银行还是小银行，上一期的流动性水平和取款人数对取款行为在 10% 的显著水平上产生影响。追随组银行存款人在决定是否提前取款时会关注自身银行的流动性水平和其他存款人的行为，但是因为信息不完备，他们只能基于上一期的流动性水平和取款人数做出判断。而追随组 Liq_{it} 的回归系数表明，他们的判断并不能真实反映当期银行的实际流动性水平。进一步观察 $Othnum_{it}$ 和 Lin_{it} 的回归结果，当小银行作为追随者时，$Othnum_{it}$ 在 1% 的水平上显著影响取款行为；当大银行作为追随者时，Lin_{it} 在 10% 的水平上显著影响取款行为。这一方面表明，无论与领导组大银行是否存在关联，追随组小银行存款人受大银行的影响容易出现恐慌挤兑行为，但是追随组大银行的存款人并不受领导组小银行的影响；另一方面，追随组大银行存款人相对比较理性，在决定是否提前取款时会考虑与领导组小银行之间的关系，该结果与前述均值比较检验的结论相吻合。

最后分析基准组大银行和小银行的回归结果。在基准组实验中，银行存款人不能获得任何有效信息，唯一可以参考的是，在决定是否提前取款后他们能够看到自己的报酬，基于支付矩阵他们能够判断自身所在银行上一期的总取款人数。表 5 中基准组实证检验结果表明，只有上一期的取款人数在 1% 的水平上显著影响存款人的取款决策，这与预期结果一致。基准组 Liq_{it} 的回归系数并不显著，因为信息透明度极低，银行存款人的取款行为完全不能反映银行的实际流动性水平，如果银行濒临破产危机，无疑会损害银行存款人的利益。

综上所述，信息透明度对银行存款人的行为存在显著影响。当信息透明度较高时，银行存款人会做出理性取款决策，对银行流动性水平做出积极反馈；当信息透明度不足时，银行存款人的取款决策容易受到外部干扰，尤其是规模大的银行因为破产危机出现银行挤兑时，规模小的银行存款人很容易因为心理恐慌而产生挤兑行为；当信息透明度极低时，银行存款人不能对银行经营状况做出有效响应，很可能因为不能及时获得有效银行信息而做出错误决策。

五、结论与政策建议

伴随着金融危机的不断出现，银行挤兑风险一直是学术界的研究热点之一。已有研究主要聚焦于单一银行内部客户的取款行为，或者两个相同规模银行的挤兑风险传染问题，而且尚没有关于不同信息透明度的针对性研究。本文基于 DD 模型设计不同信息透明度下存在两个不同规模银行主体的经济学实验，通过开展实验研究不同情境下银行客户的取款行为，实证分析出现银行挤兑的可能性和风险传染的路径，研究结论对于减少银行挤兑风险传染，防范化解系统性风险具有重要意义。

均值比较检验和多元回归分析结果都表明，信息透明度对银行存款人的取款行为存在显著影响。当信息透明度较高时，银行存款人的取款行为相对理性，银行所披露的信息和其他存款人的取款行为是主要决策依据。当信息透明度不足时，银行存款人进行取款决策的依据缺失，他们容易受外界干扰信息的影响，进而出现恐慌预期和恐慌挤兑行为。当信息透明度极低时，银行存款人不能获得充足有效的信息及时做出决策，如果银行濒临破产危机，他们的利益很可能受到损害。由此可见，提高银行信息透明度有助于银行客户做出理性决策，维护银行客户的利益，促进对银行等金融机构的市场约束。正如王宗润等（2015）[26] 所指出的，监管部门应努力提高银行强制信息披露的水平，以约束并降低银行更大风险承担行为发生的可能。

针对领导组银行的分析表明，当信息透明度较高时，大银行存款人倾向于关注银行基本面信息，以银行流动性水平作为取款动机的决策依据，而小银行存款人的取款行为则主要取决于与银行其他存款人的内部协调。关于银行挤兑的潜在原因，一些学者认为是存款人之间协调博弈的结果，而另外一些学者则认为是银行基本面信息不对称导致的。本文发现导致银行挤兑的原因可能因风险传染中银行规模的不同而存在差异，大银行存款人相对理性，主要基于银行信息做出决策，而小银行存款人则易于恐慌，容易受银行其他存款人的影响。

针对追随组银行的分析表明，当小银行出现挤兑风险时，无论是否与其处于关联状态，大银行基本不会受小银行的影响，挤兑风险由小银行传染至大银行的可能性较低；但是当大银行出现挤兑风险时，即使小银行与其不存在关联，也容易出现心理恐慌而导致银行挤兑，挤兑风险由大银行传染至小银行的可能性很高。Chakravarty 等（2014）[14]、Brown 等（2016）[15] 针对两个同等规模银行进行研究表明，即使两个银行处于独立状态，一个银行出现挤兑风险时，另外一个银行的存款人很容易产生心理恐

慌而导致风险传染，最终产生银行挤兑。本文的研究结果是已有研究的有效补充，在基于存款人心理预期的挤兑风险传染中，银行规模无疑是一个决定性的重要因素，这也与监管部门将规模作为评价金融机构系统重要性核心指标的实践相一致。

本文采用实验经济学方法展开研究，主要考虑了不同信息透明度和不同银行规模下的银行挤兑与风险传染问题。但是，现实中银行的主营业务、经营地域、银行性质，以及公众认可度、资产负债管理水平、风险控制能力等都会影响银行挤兑和风险传染。因此，如何将更多的影响因素纳入实验范畴，进行更贴合实际的实验设计是本文进一步的研究方向。

参考文献

[1] Diamond D，Dybvig P. Bank Runs，Deposit Insurance，and Liquidity [J]. Journal of Political Economy，1983，91 (3)：401-419.

[2] Goldstein I，Pauzner A. Demand-deposit Contracts and the Probability of Bank Runs [J]. The Journal of Finance，2005，60 (3)：1293-1327.

[3] Iyer R，Puri M. Understanding Bank Runs：The Importance of Depositor-Bank Relationships and Networks [J]. The American Economic Review，2012，102 (4)：1414-1445

[4] Davis D D，Reilly R J. On Freezing Depositor Funds at Financially Distressed Banks：An Experimental Analysis [J]. Journal of Money，Credit and Banking，2016，48 (5)：989-1017.

[5] Shakina E，Angerer M. Coordination and Communication during Bank Runs [J]. Journal of Behavioral and Experimental Finance，2018：115-130.

[6] Madies P. An Experimental Exploration of Self-fulfilling Banking Panics：Their Occurrence，Persistence，and Prevention [J]. The Journal of Business，2006，79 (4)：1831-1866.

[7] Garratt R，Keister T. Bank Runs As Coordination Failures：An Experimental Study [J]. Journal of Economic Behavior and Organization，2009，71 (2)：300-317.

[8] Kiss H J，Rodriguezlara I，Rosagarcia A，et al. On the Effects of Deposit Insurance and Observability on Bank Runs：An Experimental Study [J]. Journal of Money，Credit and Banking，2012，44 (8)：1651-1665.

[9] 王海巍. 存款保险及可观测度对银行挤兑的分析——基于 z-Tree 软件平台的实验方法 [J]. 中国管理科学，2014，22：389-395.

[10] Arifovic J，Jiang J H，Xu Y，et al. Experimental Evidence of Bank Runs As Pure Coordination Failures [J]. Journal of Economic Dynamics and Control，2013，37 (12)：2446-2465.

[11] Kiss H J，Rodriguez-Lara I，Rosa-Garcia A. Do Women Panic More than Men? An Experimental Study of Financial Decisions [J]. Journal of Behavioral and Experimental Economics，2014，52：40-51.

[12] Dijk O. Bank Run Psychology [J]. Journal of Economic Behavior and Organization，2017，144：87-96.

[13] Kiss H J，Rodriguezlara I，Rosagarcia A，et al. Panic Bank Runs [J]. Economics Letters，2018：146-149.

[14] Chakravarty S，Fonseca M A，Kaplan T R，et al. An Experiment on the Causes of Bank Run Contagions [J]. European Economic Review，2014：39-51.

[15] Brown M，Trautmann S T，Vlahu R，et al. Understanding Bank-run Contagion [J]. Management Science，2017，63 (7)：2272-2282.

[16] Chang R，Velasco A. A Model of Financial Crises in Emerging Markets [J]. Quarterly Journal of Economics，2001，116 (2)：489-517.

[17] Martin A. Liquidity Provision Vs. Deposit Insurance：Preventing Bank Panics without Moral Hazard [J]. Economic Theory，2006，28 (1)：197-211.

[18] Jacklin C J，Bhattacharya S. Distinguishing Panics and Information-based Bank Runs：Welfare and Policy Implications [J]. Journal of Political Economy，1988，96 (3)：568-592.

[19] Allen F，Gale D. Optimal Financial Crises [J]. Journal of Finance，1998，53 (4)：1245-1284.

[20] Chen Y，Hasan I. The Transparency of the Banking System and the Efficiency of Information-based Bank Runs [J]. Journal of Financial Intermediation，2006，15 (3)：307-331.

[21] Gu C. Herding and Bank Runs [J]. Journal of Economic Theory，2011，146 (1)：163-188.

[22] Balkenborg D，Kaplan T R，Miller T J，et al. Teaching Bank Runs with Classroom Experiments. [J]. Journal of Economic Education，2011，42 (3)：224-242.

[23] Schotter A，Yorulmazer T. On the Dynamics and Severity of Bank Runs：An Experimental Study [J]. Journal of Financial Intermediation，2009，18 (2)：217-241.

[24] He Z，Manela A. Information Acquisition in Rumor-based Bank Runs [J]. Journal of Finance，2012，71 (3)：1113-1158.

[25] 王宗润，万源沅，周艳菊. 隐性存款保险下银行信息披露与风险承担 [J]. 管理科学学报，2015 (4)：84-97.

(责任编辑：韩　嫄　张安平)

我国商业银行系统性风险的识别与预警研究

Identification and Early Warning of Systemic Risk of Chinese Commercial Banks

李辰颖

LI Chen-ying

［摘　要］防范金融系统性风险是十九大提出的重要任务，而银行是金融系统非常重要的子系统，识别银行系统性风险因素、构建风险预警模型是防范风险的重要手段。首先，从宏观经济和银行系统两方面建立评估银行系统性风险的指标体系，并采用方差阈值法对指标进行初步筛选，接着提出了基于孤立森林的指标筛选法进一步筛选出与银行系统性风险高度相关的 4 个指标。其次，为了识别银行系统性风险，提出了包含孤立森林法、基于角度的离群点检测法和局部异常因子法的多种异常检测算法相结合的银行系统性风险识别模型，并选取 2012 年 1 月到 2018 年 12 月的数据进行风险识别。从识别结果来看，所构建的风险识别模型基本准确识别出了历史风险时点，与历史事件切合度高，结果具有可解释性。最后，采用 DeepAR 方法分别预测入选的 4 个指标，并将预测结果输入到银行系统性风险识别模型，以此构建银行系统性风险预警模型，预警结果表明该模型在一定程度上能够对银行系统性风险进行预警。

［关键词］商业银行　系统性风险　识别　预警　DeepAR

［中图分类号］F830.593　［文献标识码］A　［文章编号］1000-1549（2020）10-0036-18

Abstract: Preventing financial systemic risks is an important task proposed by the 19th National Congress of the Communist Party of China. As banks are very important subsystems of the financial system, identifying bank systemic risk factors and building risk early warning models are critical tools for preventing risks. Firstly, an index system that including macroeconomic and banking systems aspects was established to evaluate the systemic risk of banks. Then, 11 indicators were extracted using a variance threshold method, in order to further extract the indicators. A method based on Isolation Forest was proposed and only four indicators left, which are highly relevant to the systemic risk of banks. Secondly, a bank systemic risk identification model compounded from Isolation Forest, Angle-based Outlier Detector and Local Outlier Factor was proposed. Moreover, the data from 2012 to 2018 were selected for risk identification. The results showed that the model could identify historical risk time points which were highly consistent with the historical events. Finally, a bank systemic risk early warning model was proposed based on DeepAR. First, four indicators were predicted respectively by DeepAR, then the prediction results were input into the bank systemic risk identification model, and the identification (warnning) results were calculated. The results showed that the model could give a technical support for the early warning of the bank's systemic risk to a certain extent.

Key words: Commercial bank　Systemic risk　Identification　Early warning　DeepAR

［收稿日期］　2020-05-21

［作者简介］　李辰颖，女，1981 年 12 月生，上海财经大学会计学院博士后，北京林业大学经济管理学院副教授，研究方向为财务会计。

［基金项目］　国家社会科学基金一般项目“基于影子银行渠道的房地产泡沫与系统性金融风险关系研究”（项目编号：14BJY172）。

感谢匿名评审人提出的修改建议，笔者已做了相应修改，本文文责自负。

一、引言

金融危机让人们意识到风险的全球性和系统性，风险的传播是牵一发而动全身。近年来，金融系统仍然处于风险易发、多发期，多种不健康现象和不确定因素（例如国际国内经济政策的动态调整、金融机构内控机制的不健全、违法违规金融行为时有发生等）都可能对金融体系造成冲击，导致金融发展环境不稳定。金融风险具有传染性和突发性，且其在经济中具有核心地位，一旦产生系统性金融风险，就会严重影响经济社会发展，危害性极强。因此，防范威胁金融系统稳定的“黑天鹅”和“灰犀牛”事件具有重要意义。习近平主席强调，金融安全是国家安全的重要组成部分，准确判断风险隐患是保障金融安全的前提。党的十九大明确提出将防范化解系统性金融风险作为三大攻坚战的首要任务。

银行业作为我国金融系统中至关重要的子系统，其资产规模占金融资产总量的 80%以上，在经济改革与发展中具有举足轻重的地位。在经济环境不稳定、银行不良资产反弹压力较大的情况下，银行业运作与发展面临的隐患也随之扩大。如何使银行在动态变化的环境中及时准确地识别风险因素，并尽早做出应对之策，从而降低风险带来的损失与影响，是保证银行系统稳定运行的关键所在。因此，从我国宏观环境与银行系统的实际出发，选取一套行之有效的银行系统性风险预警指标体系，并构建预警模型，是辅助银行识别风险，提高事前防范能力，加强事中控制水平，强化事后反馈机制的重要手段。只有做到了对系统性风险的防范与控制，才能保障银行系统稳健运行，进而为金融系统的安全与稳定奠定基础。

本文的结构安排为：首先回顾了银行系统性风险评价指标和评价方法，并进行总结与述评，得出进一步研究的方向。其次，在总结文献并结合本文研究内容的基础上，构建银行系统性风险度量指标体系，采用方差阈值法对指标进行初步筛选，并提出了基于孤立森林法的指标选取方法，进一步筛选指标，其目的是对指标进行降维，筛选出与银行系统性风险相关度高的指标，为银行系统性风险识别奠定基础。再次，通过多种异常检测算法相结合的方法，构建银行系统性风险识别模型，对银行系统性风险进行识别。接着，采用 DeepAR 模型分别预测入选指标，构建银行系统性风险预警模型。最后，对本文研究进行总结，并得出结论及给出建议。

本文研究内容与其他研究的主要区别有：（1）在银行系统性风险度量指标的选取上完全以数据为驱动，避免了人为主观因素的干扰，在指标筛选中提出了一种基于随机森林法的适用于无监督学习的指标筛选方法；（2）采用 Soft-DTW 来度量银行股票回报率相似度，间接衡量银行系统传染性，该方法可以在不关注时间序列先后关系及是否存在压缩的情况下，获取两条时间序列的最佳匹配路径，非常适合银行回报率的相关性度量；（3）在银行系统性风险的度量上，提出了多种异常检测算法相结合的风险识别方法，避免了指标结合过程中给指标权重赋值这一难点；（4）在指标的预测中，将指标看做时间序列，采用了 DeepAR 对指标进行预测，不但避免了预测指标时对其他数据的收集，而且相较其他方法提高了预测精度。

二、文献综述

（一）银行系统性风险评价指标

由于金融系统安全与稳定的重要性，以及其影响因素的复杂性，国内外政府机构和一些学者均针对这一问题进行了研究，下面分别对政府机构与学者构建的银行风险评价或预警指标体系进行回顾。

政府机构的预警指标体系具有代表性的是美国金融管理当局的 CAMELs 评级体系，该体系是对商业银行的业务经营、信用状况等制定的综合等级评定制度，从资本充足率（Capital adequacy）、资产质量（Asset quality）、管理水平（Management ability）、盈利状况（Earning performance）和流动性（Liquidity）5 个方面来评价商业银行风险。CAMELs 评价方法因其有效性被世界上大多数国家所采用。目前国际上对商业银行评价基本与 CAMELs 评级体系一致，也主要包括这 5 个方面。

我国银监会结合我国经济、金融环境的特点，制定了更具体的“CARPALs”监管体系，该体系包括 7 方面，加上有限自由裁量，共计 8 个方面：资本充足性、拨备覆盖、贷款质量、大额风险集中度、附属机构、流动性和案件防控。与 CAMELs 体系相比，CARPALs 同样重视资本充足率、资产质量和流动性，在此基础上监管指标对集中度风险、操作风险均有所体现，强化了对大型银行集团跨境、跨业风险的并表监管。

在 CAMELs 评级体系的基础上，学者们也开展了具体的研究。Kaminsky 和 Reinhart（2000）[1] 研究发现银行危机爆发的影响因素包括汇率、利率、债券价格、货币乘数以及实体经济发展情况。Demirgüç-Kunt 和 Detragiache（2002）[2] 从宏观经济、制度环境、金融体系三个方面构建了影响银行安全的指标体系。宏观经济层面考虑了 GDP 增长、利率、通货膨胀率、货币政策等；金融体系因素有 M2（广义货币）、外汇储备率、公共部门信贷规模等；制度环境因素是经济发展水平。

我国学者孙强和崔光华（2017）[3]、刘松林等（2018）[4]、赵丹丹和丁建臣（2019）[5] 同样是从这三个方面选择了预警指标，如反映宏观经济环境的国民生产总值增长率、居民消费价格指数增长率和广义货币增长率等；反映微观银行经营状况的资本充足率、不良贷款率、拨备覆盖率、存贷比、净资产收益率和资产负债率等；在此基础上均考虑了行业冲击，用固定资产投资额增长率、房地产投资增长率、国房景气指数来衡量。孙强和崔光华（2017）[3] 还增加了银行间同业拆借利率，赵丹丹和丁建臣（2019）[5] 增加了国际经济环境指标，包括经常账户差额/GDP、外汇储备增长率、进出口额增长率等。

除了上述研究普遍采用的宏微观指标外，还有学者将指标选取重点放在上市银行的股市收益率，通过上市银行在股市上的表现来预测银行系统性风险，例如肖璞等（2012）[6]、陈忠阳和刘志洋（2013）[7]、田海山等（2017）[8]、蒋海和张锦意（2018）[9]、吴敏灵（2018）[10] 则主要利用上市银行微观数据来度量银行系统性风险，选取了银行系统收益率、股市收益率、股市波动率等指标，在此基础上加入了一些银行特征数据，如银行的杠杆率、规模，同时也考虑了宏观经济 GDP 增长率、国房景气指数等。Patro 等（2013）[11] 利用金融机构的股票收益率相关性和 CDS 的违约率相关性来衡量系统性风险。高国华和潘英丽（2013）[12]、王琳和沈沛龙（2017）[13]、毛昌梅等（2020）[14] 利用上市银行收益率两两间的时变相关系数构建了银行体系风险联动的预警指标。

此外，还有研究从风险角度选取指标。陆静和王捷（2012）[15] 依据巴塞尔协议从银行面临的信用风险、市场风险、操作风险、流动性风险四个方面逐层去寻找关键风险的诱因指标，指标中包含了宏观经济指标，也包括了具体的银行操作指标。汪冬华等（2013）[16] 基于银行面临的信用风险、市场风险、操作风险，选取了上市商业银行的财务数据和金融市场公开数据来研究银行系统性风险。方意（2016）[17] 通过构建银行资产负债表直接关联网络模型，计算了银行脆弱性和传染性指标，研究了银行系统性风险的四种传染渠道以及各渠道对风险影响程度的差异，并得出“常态”系统性风险由杠杆率驱动，而“危机”系统性风险由银行关联性驱动。吴念鲁等（2017）[18] 在利用银行同业数据建立银行同业网络的基础上构建银行机构网络模拟模型，研究不同冲击下银行同业间流动性风险的传染机制和后果。赵越和韩永辉（2016）[19] 在研究流动性风险和传染性风险时采用银行两两之间的同业交易头寸，构建风险头寸矩阵。吴卫星等（2016）[20] 采用银行流动资产总量与同业存款之比来作为衡量流动性风险的指标，用CoVaR 度量银行间风险溢出效应。

通过上述银行系统性风险指标的总结可以得出，评价指标可以大致分为三类：第一类是从宏观经济环境、微观银行财务数据来选取指标；第二类是侧重通过反映银行业在股市上表现的指标来预测风险；第三类是从银行面临的风险出发逐层寻找诱因指标。

（二）银行系统性风险评价方法

针对这三类评价指标，评价方法也可以归为三大类。

第一类从宏观经济环境、微观银行财务数据来选取指标预测风险，最关键的问题是如何确定权重以及如何找到核心指标。刘松林等（2018）[4] 采用变异系数法、熵值法和拉开档次法进行赋权，然后采用相对熵法对三种方法求出的权重进行综合赋权，确定商业银行预警指标的权重。杨霞和吴林（2015）[21]、谢贤芬和王斌会（2017）[22] 运用主成分分析法进行赋权，构建了银行宏观安全状态指数，然后建立 Logit 模型识别预警银行宏观安全状态的关键性指标。孙强和崔光华（2017）[3] 采用 FR 概率模型构建了我国银行业系统性风险压力指数。Li 等（2013）[23]、赵丹丹和丁建臣（2019）[5] 采用支持向量机（SVM）和核主成分分析法构建了中国银行业系统性风险预警模型。

第二类主要是基于股票收益指标来建立预警模型，主要的方法有条件在险价值法（CoVaR）、系统性期望损失（SES）、DIP（Distress Insurance Premium）和 Shapley Value。Adrian 和 Brunnermeier（2016）[24] 提出的 ΔCoVaR 是具有代表性的系统性风险测度方

法，该方法以银行处于危机或正常状态下的 CoVaR 之差作为其所承担的系统性风险度量。肖璞等（2012）[6]、陈忠阳和刘志洋（2013）[7]采用 CoVaR 方法，运用我国上市商业银行的股票收益率数据，对我国上市银行的系统性风险进行了评价。田海山等（2017）[8]结合广义双曲线分布族与 Copula 函数扩展系统性风险测度 CoVaR 的方法，并用该方法对我国商业银行系统性风险进行了测度。

Acharya 等（2017）[25]提出的 SES（Systemic Expected Shortfall）将整个金融系统资本不足时，某一单个银行的预期资本短缺定义为该银行的 SES，并将其作为对系统性风险的贡献，用每日边际期望损失（Marginal Expected Shortfall，MES）表示某一单个银行对金融系统 ES（Expected Shortfall）的边际贡献。吴敏灵（2018）[10]首先估算了上市银行 MES，然后利用 MES 的可加性，以每家银行的月度总资产账面价值的份额为权重，对各银行月度平均 MES 进行加权平均，计算出总边际期望损失，对银行整体系统性风险预警。

高国华和潘英丽（2013）[12]采用多元 GARCH-BEKK 测算银行收益率的动态相关系数，应用 Markov 区制转换模型来识别和判断银行系统性风险状态。王琳和沈沛龙（2017）[13]利用 DCC-GARCH 测算上市银行收益率两两间的时变相关系数，研究了银行间的风险联动关系。毛昌梅等（2020）[14]采用多元 GARCH-BEKK 构建波动溢出网络，并通过分析网络拓扑指标研究了银行网络的溢出效应。

第三类是基于银行面临的风险选取指标进行评价。相应的风险度量方法包括：陆静和王捷（2012）[15]通过构建商业银行风险的拓扑结构，度量了各类指标对全面风险的贡献。汪冬华等（2013）[16]采用风险因子模型来构建市场和信用风险敞口回报的边际分布，用损失分布法来构建操作风险敞口回报边际分布，然后基于 Copula 理论的整体风险度量法，度量我国商业银行整体风险。李楠等（2014）[26]在度量商业银行的市场风险时提出了基于组合权重伪似然函数的半参数方法。戴丽娜（2017）[27]基于非参数方法估计了商业银行操作风险。方意（2016）[17]、吴念鲁等（2017）[18]通过构建基于银行资产负债表的关联网络模型研究了系统性风险的传染渠道和后果。

（三）文献评述

通过对现有文献中银行系统性风险评价指标及评价方法的回顾与总结，发现评价指标主要有三种类型。第一类是从宏观经济环境、微观银行财务数据去寻找指标；第二类是从银行的市场表现上寻找指标；第三类是从银行面临的风险诱因中寻找指标。这三类指标各有优劣。第一类指标涉及的方面众多，优点是可以从多方面、多角度去识别引发风险的因素，但缺点是指标众多时很难及时识别风险。第二类指标重点依赖于银行的市场表现，优点是市场指标更新较快，可以及时发现风险，但缺点是市场表现会受到众多因素的干扰，可能并不源于银行自身的风险。第三类指标直接从银行面临的各类风险着手，层层分解地寻找指标，优点是识别出的风险因素更具体、更直接，但缺点是这样的风险识别方式更适合于事后的原因分析，而对于风险预测则较不适用。不适用的原因有两点：一是某一因素出现问题可能并不会引发系统性风险；二是要随时去识别与关注这些细小因素的变动不符合成本效益原则。

关于方法的选择，则是根据评价指标的不同，适用的评价方法也不同。针对第一类指标常用的方法有信号分析模型、支持向量机、人工神经网络、熵值法和主成分分析法等；第二类指标常用方法有条件在险价值法、系统性期望损失、DIP、Shapley Value 和 GARCH-BEKK 等；第三类指标常用方法有贝叶斯网络、风险因子模型、资产负债表的关联网络模型和损失分布法等。

通过文献回顾与总结，结合本文研究目的，银行系统性风险预警主要需要解决三个方面的问题：一是采用哪些指标来度量银行系统性风险；二是如何识别银行业是否有风险；三是采用什么方法对银行系统性风险进行预警。对于第一个问题，风险预警指标不一定是越多越好，指标越多越不具有预警性，只有把握关键性指标，才能使预警具有可行性、符合成本效益原则、及时有效，因此在综合考虑三类指标优缺点的基础上，首先选择将第一类和第二类指标融合，然后利用适用于指标降维的方差阈值法初步筛选指标，接着提出了基于孤立森林法的指标筛选方法，进一步筛选出与系统性风险具有密切关系的指标。对于第二个问题，首先采用局部线性嵌入法（Locally Linear Embedding，LLE）[28]对指标进行降维，然后提出了基于角度的离群点检测法（Angle-based Outlier Detector，ABOD）[29]、孤立森林（Isolation Forest，IF）[30][31]和局部异常因子法（Local Outlier Factor，LOF）[32]相结

合的银行系统性风险识别模型。对于第三个问题，把度量指标看作时间序列，采用 DeepAR[33] 方法基于历史数据对未来数据进行预测，并将预测结果作为银行系统性风险识别模型的输入，对银行系统性风险进行预警。下面就这三方面问题分别进行阐述。

三、构建银行系统性风险度量指标

（一）银行系统性风险度量指标的提出

在银行系统层面，通过分析具有权威性的银行系统性风险评价体系“CAMELs”，并结合我国银监会的“CARPALs”监管体系，可以得出，对银行系统性风险的度量应涵盖银行资本充足性、资产质量、盈利情况、流动性和对风险的抵抗能力等方面。因此，本文在选择评价指标的时候也尽量涵盖这些方面。在参照前人研究的基础上从宏观经济和银行系统两个层面进行了指标选取。

1. 宏观经济指标的选取。

在宏观经济层面，总结已有文献并结合本文的研究内容进行指标选取。银行业的收入主要来源于利息收入、中间业务收费和投资收益等等，国家经济状况、各行各业的经营状况、国民收入及消费水平等都会影响银行业的经营状况，因此选取反映国家整体经济状况的 GDP 增长率、反映工业企业和银行业经营状态和发展趋势的景气指数及企业家和银行家信心指数、反映市场变动情况的中采制造业采购经理指数、反映房地产业发展变化趋势和变化程度的国房景气指数、反映民众消费情况的居民消费价格同比涨跌幅及消费者预期指数。

2. 银行系统指标的选取。

银行系统方面，主要从商业银行的财务数据和市场数据两个方面选取指标。指标大部分都有公认的计算方法，因此只介绍指标的选取，只有银行股票回报相关性属于复杂合成指标，单独介绍其计算方法。

（1）银行财务数据指标。

参照杨霞和吴林（2015）[21]、孙强和崔光华（2017）[3]、刘松林等（2018）[4]、赵丹丹和丁建臣（2019）[5]、胡历芳和刘纪鹏（2019）[34] 的研究，从银行的财务表现，选取了反映资本充足性的资本充足率、反映银行业市场资金价格和状态的银行间同业拆借利率、反映风险抵抗能力的拨备覆盖率、反映资产质量的不良贷款率、反映收益能力的资产利润率与资本利润率、反映流动性的流动性比例。

考虑到由于银行同业业务而形成的银行关联性，进而引发银行之间的传染性，参照方意（2016）[17]、赵越和韩永辉（2016）[19]、吴念鲁等（2017）[18] 的研究，选择反映银行间关联性的同业负债关联度指标，该指标是银行负债中的同业负债与总资产的比值。首先选取上市银行季度报告数据，计算单家银行的同业负债与总资产的比值，然后求取当季度所有银行该比值的平均值，最后采用线性插值法转化为月度数据。该指标的值越大表明银行之间的关联度越高，同时从系统整体来看，该值越大说明流动性缺口也越大。

（2）银行市场数据指标。

从银行的市场表现选取评价风险的指标，通过收益率及波动率等来度量银行的市场表现，但由于单个企业的日数据不适合与宏观经济指标以及银行系统的其他指标进行结合研究，因此选用可以反映商业银行整体市场表现的月度数据——中证银行指数的月度收益率来与其他指标结合研究。

由于传染性是系统性风险产生的重要原因，因此单独选取衡量银行传染性的指标。当一家银行遭遇危机，由于银行间的关联性，危机会传染给系统中的其他银行，通过市场股价数据可以推导投资者对于银行的风险预期，这种冲击的速度和程度最终会体现为各家银行收益回报的同向变动。参照 Patro 等（2013）[11]、高国华和潘英丽（2013）[12]、王琳和沈沛龙（2017）[13]、毛昌梅等（2020）[14] 的研究，采用上市银行股票收益率的相关性衡量银行系统传染性（指标命名为银行股票回报相关性）。

在对银行传染性风险指标选取及度量方法文献总结与分析的基础上，提出了选用银行股票回报率数据，采用 Soft-DTW[35] 来度量银行股票回报率相似度，间接衡量银行系统传染性的方法。由于该指标属于复杂合成指标，下面单独对其生成过程进行介绍：

第一步是指标特征分析。银行回报率存在如下特征：（1）某一事件的冲击，可能影响某些银行的回报率，而如果事件持续作用或由于银行间的传导作用，则可能影响其他银行，也即其时间存在先后关系；（2）事件的冲击对某些银行回报率影响大，而有些影响弱，即幅值存在大小问题；（3）事件的冲击对某些银行的影响时间长，而对有些影响时间短，即持续时间存在长短问题（也即是否存在压缩）。

第二步是研究方法选择。采用银行股票回报相关性来度量传染性，关键是选取相关性分析方法。相关

性分析常见的主要方法有皮尔逊相关系数、时间滞后互相关法（TLCC）、动态时间规整法（DTW）和瞬时相位法等。其中皮尔逊相关系数的优点是简单明了，但一般来说无法体现时间序列的先后顺序；TLCC 法和瞬时相位法可以处理时间序列的前后关系，但对于压缩扭曲的时间序列分析效果相对较差；DTW 是通过计算两条时间序列的相似度来判断其相关度，可以在不关注时间序列先后关系，是否存在压缩的情况下，获取两条时间序列的最佳匹配路径。综合考虑银行回报率的特征及 DTW 法的优势，选择 DTW 法计算银行回报率相似度。DTW 虽然能够较好地计算两条时间序列相似度，但也存在一些问题，其中之一便是损失函数不可微，Soft-DTW 正是为了解决这一问题提出的，故本文采用 Soft-DTW 计算银行系统传染性。

第三步是指标的计算。在计算中选取了 2012 年 1 月到 2018 年 12 月所有上市银行的回报率日数据。假设共有 m 家银行，以计算第 n 月的银行系统传染性为例，步骤如下：首先，采用 soft-DTW 方法，分别计算第 i 家银行月回报率和其他银行相似度，得到 $m-1$ 个值，求取平均值，得到该银行对其他银行平均相似度；其次，计算完所有银行相对于其他银行的平均相似度后得到 m 个平均值，对这 m 个平均值再求平均值，得到该月相似度平均值，也即该月银行股票回报相关性；最后，计算完所有月份的银行股票回报相关性，得到该阶段银行股票回报相关性。计算得到的银行股票回报相关性部分统计特征如表 1 所示。

表 1　　银行股票回报相关性部分统计特征

均值	方差	四分位距	四分位发散系数	变异系数	分桶熵
-1.572 8	0.035 5	0.253 5	0.079 4	-0.119 8	2.079 0

从以上统计特征可以得出，银行之间回报率曲线是比较相似的，相似度的分布是比较集中的，这反映出银行之间联系比较紧密，传染性较高。

3. 指标汇总与数据来源。

所有指标按照类型分为宏观经济指标和银行系统指标两类，共计 24 个指标。样本区间选择 2012 年 1 月到 2018 年 12 月。指标名称如表 2 所示。数据来源于国家统计局网站、中经网统计数据库、中国人民银行网站、国泰安数据库和银监会网站等。

表 2　　银行系统性风险度量指标

	指标名称		指标名称
宏观经济指标	国民经济同比实际增速	银行系统指标	商业银行存贷比
	货币和准货币期末同比增速		商业银行资本充足率
			商业银行拨备覆盖率
	工业企业家信心指数		商业银行流动性比例
	银行家信心指数		商业银行不良贷款率
			商业银行人民币超额备付金率
	工业企业景气指数		银行间同业拆借加权平均利率
	银行业景气指数		商业银行成本收入比
			商业银行资产利润率
	中采制造业采购经理指数		商业银行资本利润率
	国房景气指数		银行指数月收益率
			银行指数月波动率
	居民消费价格同比涨跌幅		银行同业负债关联度
	消费者预期指数		银行股票回报相关性

注：以上数据均为月度数据，对于季度数据进行线性差值转换为月度数据。

（二）银行系统性风险度量指标的筛选

由于初步提出的银行系统性风险的度量指标有 24 个，而数据样本数为 84 个，也即样本数较少，而维度较大，如果不对样本进行筛选，主要可能会造成以下后果：一是指标之间可能有相关性，导致后续用于银行系统性风险识别的异常检测算法结果不稳定；二是容易导致异常检测算法和各指标时间序列预测算法过拟合；三是不利于结果的分析和结论的得出。因此，必须对指标进行合理的筛选。

在指标的筛选过程中首先采用方差阈值法对指标进行初步筛选，然后提出了基于孤立森林法的指标选取方法，对剩余指标进行进一步筛选，得到最终的度量指标。

1. 基于方差阈值法的指标筛选。

方差阈值法的原理是首先估计各指标的方差，然后根据阈值对指标进行筛选。采取该方法的主要原因是：文中采用的是基于多种异常检测相结合的方法，对银行系统性风险进行识别和判断，如果一个指标的方差较小，则说明该指标的整体波动较小，也即对异常的贡献较少，可以较为安全地删除该指标。计算步骤如下：（1）对样本数据标准化；（2）对标准化后的数据计算各指标方差；（3）根据阈值筛选指标。得到的计算结果如表 3 所示。

表 3　　方差阈值法计算结果

指标	方差	指标	方差
银行指数月波动率	1.549	中采制造业采购经理指数	0.477
银行间同业拆借加权平均利率	1.276	商业银行成本收入比	0.457
银行股票回报相关性	1.123	商业银行资产利润率	0.447
商业银行资本充足率	1.034	商业银行资本利润率	0.412
商业银行流动性比例	0.861	银行家信心指数	0.368
银行指数月收益率	0.819	国民经济同比实际增速	0.367
工业企业家信心指数	0.741	银行业景气指数	0.320
居民消费价格同比涨跌幅	0.707	货币和准货币期末同比增速	0.295
消费者预期指数	0.670	银行同业负债关联度	0.275
工业企业景气指数	0.635	国房景气指数	0.249
商业银行存贷比	0.615	商业银行拨备覆盖率	0.236
商业银行人民币超额备付金率	0.549	商业银行不良贷款率	0.227

若将阈值取为各指标方差均值 0.6，那么筛选出 11 个指标，分别为银行指数月波动率、银行间同业拆借加权平均利率、银行股票回报相关性、资本充足率、工业企业家信心指数、银行指数月收益率、居民消费价格同比涨跌幅、商业银行存贷比、消费者预期指数、工业企业景气指数和商业银行流动性比例。

2. 基于孤立森林法的指标筛选。

经过阈值筛选法得到了 11 个备选指标，备选指标个数仍相对较多，需要进一步筛选。指标的选择即特征选择，常用的特征选择方法有过滤法、包裹法和嵌入法，除过滤法的一些算法外，其他方法的多种算法一般用于监督学习，而本文所研究的银行系统性风险识别问题，不易事先得到样本所属类别，属于无监督学习，故大部分特征选择方法均不适用。本文对银行系统性风险识别采用的是异常检测算法，因此可以利用异常检测算法的结果构造一个度量值来筛选指标，这样既可以提高异常检测算法的准确度，又可以对指标进行筛选，降低数据维度。

基于以上分析提出了基于孤立森林法的指标筛选方法。下面先给出该方法的步骤，然后对应用这些步骤的原因及结果进行分析。

基于孤立森林的指标筛选方法主要步骤为：（1）随机选取备选指标的子集；（2）对子集指标数据标准化；（3）采用局部线性嵌入法对标准化后的子集指标数据降维；（4）对降维后的数据采用孤立森林法进行异常检测，得到每个样本点的得分；（5）重复步骤（4）10 次，求取每个样本点的平均得分；（6）分别求取子集中每个异常样本点到正常样本中心的欧氏距离，并求取所有欧式距离的中位数；（7）重复 n 次步骤（1）~（6）得到 n 个欧式距离中位数，欧式距离中位数最大者所对应的子集即为最佳子集。

步骤（2）中对子集数据进行标准化，是因为第（6）步中要求取异常样本到正常样本中心的欧式距离，如果不进行标准化，那么欧式距离受子集指标数值大者影响较大，不能正确反映两者之间的关系，而且标准化后也有利于局部线性嵌入法的降维。

步骤（3）中采用了局部线性嵌入法对数据进行降维。局部线性嵌入法是一种流形学习，它能够在局部保持样本之间的几何关系，因而在降维的过程中相较于其他降维方法更有可能保持邻近样本之间的关系，能够尽量避免将异常点塌缩到正常样本点中，或将正常点映射到异常样本中，而且降维之后也有助于孤立森林法检测异常值。

步骤（4）中采用孤立森林法的原因是：孤立森林法一般用于结构化数据的异常检测，其核心思想是随机选择超平面对数据空间进行切割，处于低密度区域的点和较高密度区域的点就容易被孤立起来，把孤立起来的点作为异常值，就可以完成异常检测任务。孤立森林法已被证明是一种有效的异常检测算法，而且可以对每个样本点进行评分，故而采用孤立算法对指标进行筛选。

步骤（5）中重复 10 次求取平均值是为了减少孤立森林算法随机因素的干扰，使算法更稳健。

步骤（6）中采用中位数可以更好地避免极端异常值给算法带来的影响。

3. 指标筛选结果及分析。

鉴于备选指标只有 11 个，综合考虑计算量和时间成本，可以采用穷举法选取所有的子集，并计算所有子集每个异常样本点到正常样本中心欧氏距离的中位数，欧式距离中位数最大者所对应的指标子集为：银行指数月波动率、银行间同业拆借加权平均利率、商业银行流动性比例以及银行指数月收益率。欧式距离中位数最大的前 5 个子集如表 4 所示。

表 4 孤立森林法筛选出的欧式距离最大的 5 个子集

指标子集	距离
商业银行流动性比例、银行间同业拆借加权平均利率、银行指数月收益率、银行指数月波动率	0. 352
银行间同业拆借加权平均利率、银行指数月波动率、银行股票回报相关性	0. 331
银行间同业拆借加权平均利率、商业银行存贷比、商业银行资本充足率、银行指数月波动率	0. 328
居民消费价格同比涨跌幅、银行间同业拆借加权平均利率、银行指数月收益率、银行指数月波动率	0. 327
商业银行流动性比例、银行指数月收益率、银行指数月波动率、银行股票回报相关性	0. 298

从筛选结果可得：入选的 4 个指标商业银行流动性比例、银行间同业拆借加权平均利率、银行指数月收益率、银行指数月波动率，均属于银行系统指标，而没有宏观经济指标。这其中可能的原因有：（1）银行是否有系统风险，与宏观经济指标相比，银行业自身的指标更为重要；（2）宏观经济指标一般都属于长周期指标，而本文所选取的数据量可能不足以观察到明显的宏观经济指标对银行系统性风险的影响；（3）相对于银行业自身的指标，宏观经济指标波动较小，异常值也较少，对异常的贡献相对较少；（4）我国经济长期处于快速发展中，整个经济和金融业总体处于上升趋势，并且银行业也没有发生过大规模系统性风险，这可能也是宏观经济指标影响较小的一个原因。

入选指标中商业银行流动性比例和银行间同业拆借加权平均利率均是反映商业银行流动性风险的重要指标，而流动性风险是商业银行所面临的重要风险之一。商业银行的经营特点（以存放贷）使其天生存在发生流动性风险的可能。随着金融改革的深入和宏观经济的调整，银行业流动性风险长期化、结构化趋势日益明显，并成为重要的风险来源，该风险一旦发生会给金融系统造成灾难性后果。2008 年爆发的国际金融危机，使国际银行业都意识到了流动性风险管理的重要性，强化对金融机构流动性风险的监管成为国际监管改革的重要内容。因此，这两个指标的入选也较为符合现实情况。

银行指数月收益率和月波动率是银行的市场表现，可以反映上市银行总体股价变动和走势，通过该指标可以反映出银行经营的盈利状况和收益的稳定情况，而银行稳定的盈利是其持续经营与发展的重要基础。因此，从理论上分析可以得出这两个入选指标也对银行系统性风险的预警至关重要。

除理论分析可以说明入选指标与银行系统性风险具有重要关系外，已有文献的研究结论也能证明这一点。吴卫星等（2016）[20]认为由于银行同业拆借市场化程度高，使其更能真实反映银行的流动性状况；孙强和崔光华（2017）[3]选取的银行系统压力指数指标包含流动性比例、银行同业拆借利率；赵丹丹和丁建臣（2019）[5]选取“中证银行”股票指数的波动率度量中国银行业系统性风险等级。从计算结果和以上分析可得：入选的指标在一定程度上可以反映银行的系统性风险，指标具有可解释性。

4. 入选指标与系统性风险的逻辑关系（参见图 1）。

银行之间由于同业业务的普遍存在，使流动性风险在银行系统内传染很快且后果严重（吴念鲁等，2017[18]）。银行间同业拆借关系是银行间传染风险形成的主要渠道之一，使银行系统中的任何一家银行都难以独善其身。流动性短缺银行会选择流动性过剩银行作为潜在债权人，然后从中选择同业拆借利率低的流动性过剩银行进行同业拆借，如果一家银行无法满足其拆借需求，则会按照利率从低到高的排序进行逐家拆借，直到流动性短缺银行的流动性需求获得满足为止；而流动性过剩银行如果无法满足所有向其申请贷款银行的需求，则会按照一定的原则进行筛选，例如按照贷款银行的净资产进行筛选，直至流动性分配完成（李智和牛晓健，2017[36]）。拆借完成后相关银行的资产负债结构发生了相应变化，银行的资产中出现了同业资产，而负债中也出现了同业负债，这样的过程循环往复，形成银行间同业拆借关系，辅助整个银行系统维持持续的流动性，使其得以正常运转。

银行间相互存款具有流动性重新分配的作用，当银行系统流动性充足时，银行关联对流动性转移的作用占主导，使流动性分布更均衡，银行间关联有助于保持系统稳定；若系统中流动性短缺，则风险传染作用会占主导，且银行关联可能加速流动性风险的传染（邓晶等，2013[37]）。如果出现超额流动性冲击，银行普遍流动性紧张，无法满足所有银行的流动性需求，基于货币资金的供求关系，当需求大于供给时，价格会上涨，表现为资金成本的增加，也即同业拆借利率的上升。

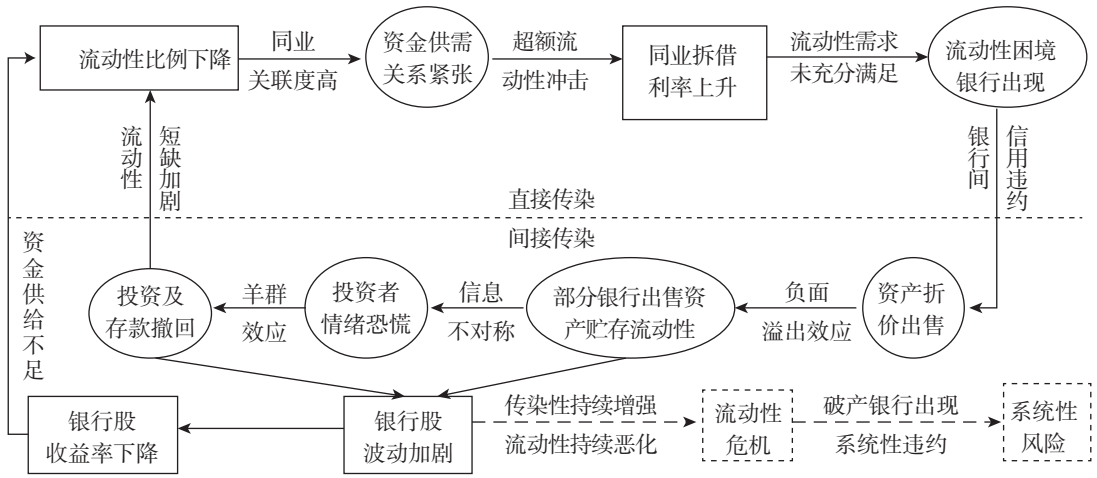

图 1　入选指标与银行系统性风险的逻辑关系

银行系统整体流动性下降，同业拆借利率上升，都意味着银行系统资金的紧张，而资金是银行的血液，当部分流动性短缺银行的流动性需求未能得到满足，因缺血而陷入流动性困境，可能出现银行间信用违约，而被迫将资产进行折价抛售（吴念鲁等，2017[18]）。银行系统又由于同业拆借、同业存单和同业理财等同业业务形成直接关联而交织成的网络，使彼此之间联系紧密，当某家或某些银行陷入困境，势必会对与之关联的银行（例如债权、债务方）造成负面的溢出效应。

银行系统除了因关联业务而形成的直接传染外，还有因经营“同质化”而面临相似风险敞口产生的间接传染，例如当外部冲击或内部影响较为严重，引发极端风险，陷入流动性困境的银行会通过资产折价抛售来满足资金需求，使银行系统面临资产价格下跌的风险。与风险生成银行高度关联的银行也成为风险承受银行（方意和郑子文，2016[38]），这些银行债权受损，坏账增加，资产规模和权益缩减，这时受传染的银行也可能出售资产以贮存流动性。银行间关联以及当流动性冲击出现时银行间采取的一致性行动，加速了风险的传染（李晓伟和宗计川，2018[39]）。

由于资本市场信息的不对称，投资者难以得知所有银行的实际经营情况，从众心理和恐慌情绪使投资者容易出现“羊群效应”。投资者恐慌情绪的蔓延，使银行流动性风险的传染范围、概率和速度显著提升（沈沛龙和李志楠，2020[40]）。某家或某些银行的负面信息，可能使投资者争相收回银行投资和存款，反映到银行股收益率的大幅波动，进而引起银行股收益率的下跌。在我国优质上市企业较为稀缺，因此银行股一直受到投资者的追捧，可以看到上市银行股的收益率普遍较高也较为稳定。但一旦银行股收益率下降、波动率提高，说明投资者对银行业失去信心，如果这种情绪迅速蔓延，大规模的撤资会引发银行业“挤兑”现象，使银行资金更加紧缺，加剧银行系统的流动性风险。同业拆借市场连接了各个商业银行，使银行间交织成的巨大网络成为流动性风险传染的重要渠道（吴卫星等，2016[20]）。

当银行间传染性持续增强、流动性持续恶化，陷入流动性危机的银行会逐渐增加，银行系统的稳定受到威胁，破产银行出现，债务违约产生连锁反应，引发“多米诺骨牌”效应，更多的银行和金融机构受到影响，风险的传染最终引发系统性风险。通过分析可知，流动性冲击极易演变成系统性风险，而银行间的直接和间接关联产生的一致性行动又会加速与放大流动性风险的传染速度。流动性比例、银行同业拆借利率、银行股收益率和波动率的变化在整个过程中的重要地位，使得综合监控这四个指标的表现和变化趋势有助于银行系统性风险的预警。

四、银行系统性风险识别

对银行系统性风险进行识别是银行系统性风险预警的前提，对银行风险进行识别不仅可以对历史数据是否有风险进行标注，而且还可以对银行系统性风险预警结果进行评价。

（一）银行系统性风险识别模型

为了对银行系统性风险进行识别，本文提出了多种异常检测算法相结合的方法。将多种异常检测算法相结合的优点是：（1）一般来说，不同的异常检测

算法对不同的异常值点关注程度是不一样的，例如有的异常检测算法较为关注全局异常点，有的检测算法更关注数据的密度等，把不同的检测方法结合起来，就更可能利用每种检测算法的优势；（2）本文所采用的样本点数较少，采用多种检测算法相结合的方式，在一定程度上有利于降低过拟合。

在构建银行系统性风险识别方法中主要有以下三个问题需要解决：（1）采取哪几种异常检测算法；（2）如何将多种算法进行结合；（3）异常点的比例如何选取。

对于问题（1），常用的异常检测算法有单类支持向量机（One-class SVM）[41]、孤立森林（IF）、局部异常因子法（LOF）、基于角度的离群点检测法（ABOD）和 Feature Bagging[42]等。单类支持向量机严格来说属于新奇检测（Novelty Detection），即训练数据不能包含异常值，本文所用数据训练时无法区分正常值和异常值，因而不适用于该算法。Feature Bagging 是一种集成方法，其基学习器可以相同也可以不同，是一种优秀的异常检测算法，但对本数据集来说，单独采用该算法的结果与单独采用其他算法的结果相比，差异较大，而且结果难于解释，故不采用该算法。经计算孤立森林法、基于角度的离群点检测法和局部异常因子法较为适用本数据集，故采用以上 3 种方法。

对于问题（2），常用的结合方法[43]有平均值法（Average）、中位数法（Median）、最大值法（Maximization）、Average of Maximum（AOM）和 Maximum of Average（MOA）[44]法等。本文所需要结合的异常检测算法只有 3 种，不适宜采用 AOM 和 MOA 方法。综合考虑最终结果的稳定性，排除最大值法，而平均值法和中位数法差异不大，最后采用平均值法。

对于问题（3），Du-Mouchel（1983）[45]经过试验统计，选择了 10%左右的数据作为金融风险样本，林宇等（2016）[46]通过回顾总结前人研究成果，选择总体样本的 9%作为金融风险样本。因而本文选择总样本数的 10%作为银行系统性风险较大的样本，其个数为 9 个，也即异常样本点个数为 9 个。

基于角度的离群点检测法、孤立森林法和局部异常因子法相结合的银行系统性风险识别模型的步骤为：首先对选取出的 4 个指标数据进行 Z-score 标准化；然后采用局部线性嵌入法（LLE）对指标进行降维；接着将降维后的数据分别输入到 3 个异常值检测算法中：基于角度的离群点检测法、孤立森林法和局部异常因子法，得到 3 种算法对每一个样本点的评分，采用简单的平均值法，得到每个样本点的最终评分，评分越高越有可能是异常点；最后根据异常点占比 10%得到异常点，即银行系统性风险较大的点。为了降低随机因素的影响，增强鲁棒性，重复 10 次，求取平均值。

（二）识别结果及分析

采用上述方法得到 9 个异常点如表 5 所示。4 个指标银行指数月波动率，银行间同业拆借加权平均利率，商业银行流动性比例以及银行指数月收益率曲线如图 2 所示，其中银行间同业拆借加权平均利率和商业银行流动性比例纵轴为百分比，同时图 2 中标注了识别出的风险较大点的位置。

表 5　　银行系统性风险识别结果

年月	得分
2012 年 1 月	1. 865
2012 年 12 月	1. 379
2013 年 6 月	3. 165
2014 年 12 月	3. 322
2015 年 1 月	1. 289
2015 年 7 月	2. 153
2015 年 8 月	1. 389
2016 年 1 月	1. 777
2018 年 10 月	1. 033

从银行系统性风险识别结果来看，整体上异常点基本分布在曲线的波峰或波谷，这个结果是符合实际情况的，如一般认为指数月波动率、月收益率大幅波动时，系统性风险较大；而银行间同业拆借加权平均利率处于高位或大幅上升时银行业也面临较大风险；商业银行流动比例较低或处于一段时期内相对低位时，银行系统性风险较其他时期高。

具体到异常点来看，2012 年 1 月的银行指数月波动率较大，且银行间同业拆借加权平均利率处于高位；2012 年 12 月银行指数月收益率急剧升高；而 2013 年 6 月银行指数月收益率大幅下跌，银行间同业拆借加权平均利率处于最高点，同时商业银行流动比例处于低位；2014 年 12 月和 2015 年 1 月银行指数月收益率大幅变化，银行指数月波动率增大，商业银行流动比例也有明显滑坡；2015 年 7 月、2015 年 8 月和 2016 年 1 月银行指数月收益率大幅下降，银行

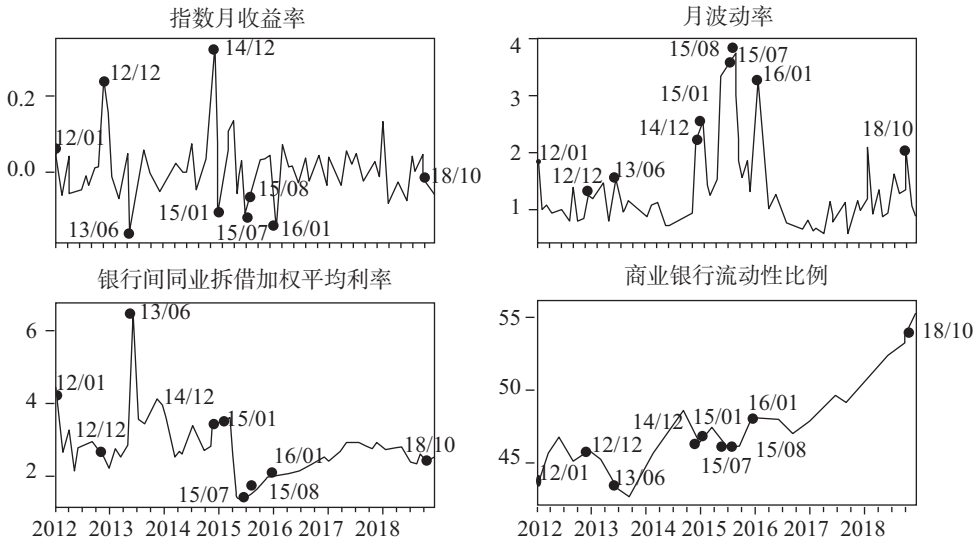

图 2　指标曲线及异常点

指数月波动率处于最高值附近，且商业银行流动比例处于一段时期内的低位；而 2018 年 10 月银行指数月波动率是近两年来的最大值，其他指标较为正常，这一点也能从其异常程度的得分体现出来。

从银行系统性风险识别结果的得分来看，风险最高的 9 个时点分别是 2012 年 1 月、12 月，2013 年 6 月，2014 年 12 月，2015 年 1 月、7 月、8 月，2016 年 1 月和 2018 年 10 月。通过历史事件回顾，可发现在这 9 个时点，确实发生过对银行系统造成影响的典型事件。

2011 年在通货膨胀逐步走高的背景下，央行采取了持续紧缩的货币政策。2011 年第一、二季度，央行数次上调准备金率并加息。该政策的影响是银行体系的超储率下跌，可用资金减少，2011 年 6 月超储率下滑至 0.8%，创历史新低。2011 年的“钱荒”问题贯穿了整个 2011 年，所以 2012 年 1 月银行系统性风险比较高。

2013 年 6 月份由于美联储退出 QE 引发的恐慌以及我国多家金融机构的期限错配造成了短期流动性不足，“钱荒”问题再次出现，银行间 7 天回购利率在 6 月 20 日创历史新高。资金价格的飙升使流动性风险升高，从而引发系统性风险的上升，这次的黑天鹅事件使银行系统受到极大冲击。

2014 年 11 月美国正式结束量化宽松政策，中国人民银行 11 月 22 日起逐步推进新的改革，包括下调贷款和存款基准利率、推进利率市场化、调高存款利率浮动区间的上限等，这些行为都可能放大银行的风险。因此，2014 年 12 月也是银行系统的高风险点。

2015 年，中国经济继续深化“新常态”，国内经济下行、GDP 增速持续下滑、人民币快速贬值、房地产投资增速回落，同时国际经济政治环境的变化也对中国经济提出了巨大的挑战。同年银行业也发生了很多变革，显性存款保险体制使存款利率高低成为储户选择银行的唯一标准，银行为图高回报，而放大经营风险；商业银行法修订，将存贷比不得超过 75% 的指标删除，将存贷比由法定监管指标转为流动性监测指标，商业银行获得了更大信贷投放空间，同时也放大了风险。2015 年 5 月 11 日起再次降息，存款利率浮动区间扩至 1.5 倍。2015 年 6 月份，中国股市出现了一次“股灾”，2015 年 7 月中国股市继续大跌，2015 年 8 月后，资本外流问题更加严重。2015 年年初至 2015 年，第三季度末，我国银行业不良贷款率持续攀升，拨备覆盖率持续下降，因此在 2015 年识别出三个银行系统性风险比较高的点。

2016 年 1 月 1 日证监会正式实施指数熔断相关规定，政策出台的目的是稳定股市，然而熔断机制出台的短期内，引起大规模恐慌，造成证券市场大跌。2016 年 1 月 4 日，A 股遇到史上首次“熔断”；2016 年 1 月 7 日，沪深 300 指数再度触发熔断线。为维护市场稳定，指数熔断机制自 1 月 8 日起暂停实施。2016 年 1 月股市的持续动荡也影响了银行系统的稳定。

2018 年下半年，国际国内潜在风险的集中释放，使金融市场产生高波动。上证综指一度下探至 2449 点，为 2015 年中以来最低值；美元兑人民币汇率突破 6.97，创 10 年来新低；中美贸易摩擦，美国在 9

月正式对原产中国的商品加征 10%的进口关税。金融行业强监管和去杠杆利好长期经济发展，但短期内可能使得潜在风险集中释放，例如商业银行不良贷款率同比上升有所上升，银行风险也升高，但相对来讲，2018 年 10 月的银行系统性风险较小，从表 5 中的得分也可以看出来。

总体来说，银行系统性风险识别方法的检测结果可以合理解释，基本正确地检测出了异常点。

五、银行系统性风险预警模型的构建

银行系统性风险预警模型的构建主要有以下三种思路：

第一种思路是将银行系统性风险识别结果看作已知标签，那么预警模型构建问题可以转换成为监督学习问题，可以采用多种监督学习方法建模。但这一思路面临三个主要难点：一是银行系统性风险识别结果是由另外的方法推断得来的，本身可能并不准确。二是有风险的样本点明显少于无风险的样本点，属于类别不平衡学习。虽然有多种解决类别不平衡的方法可供选择，但样本异常有不同的原因，对于本研究来说，从数据空间来看，异常样本点可能稀散地分布于数据空间，表示不充分，解决类别不平衡的方法难于奏效。三是样本数据偏少，不利于监督学习。

第二种思路是构建银行系统性风险指数，用指数来衡量银行系统的风险程度，再通过预测指数或和指数相关的其它量，对银行系统性风险进行预警。这种方法面临的难点有：一是怎样选取指标构建指数；二是怎样将指标合成指数；三是如何预测指数或其相关变量；四是如何判断风险。

第三种思路是将银行系统性风险分解为信用、市场、操作风险和流动性风险，然后从这四类风险出发，运用逻辑推理去推导出引发风险的原因，构建银行系统性风险的拓扑结构，然后去检验各节点的风险概率，建立预警信号灯。这种预警思路面临的难点：一是难以穷尽所有因素；二是如何给各风险点设置风险临界值；三是难以判断某一或某些节点出现问题能否引发系统风险。

本文的思路是用历史数据对指标进行预测，然后再采用银行系统性风险识别模型对银行系统性风险进行预警，这种思路的关键点在于指标的预测和异常检测算法。因为本文所构建的银行系统性风险识别指标只有 4 个，同时还构建了基于 3 种异常检测算法相结合的银行系统性风险识别模型，而这一模型也可用于银行系统性风险预警，所以采用这种思路构建银行预警模型具有可行性。由于银行系统性风险预警中的风险识别模型已构建完成，下面主要介绍 4 个度量指标的预测方法。

（一）度量指标预测模型的构建

由于 4 个度量指标的值都存在时间先后关系，可以将度量指标看作时间序列。建模中可以采用的技术路径主要有两种：一种是收集其他银行业指标、宏观经济指标和其他相关指标，将这些指标作为输入，采用某种预测方法，可以同时预测所有度量指标或其中的某几个度量指标；另一种路径是只通过度量指标的历史数据预测未来数据。这两种路径各有优劣，在都能完成任务的前提下，本文选择第二种路径。

1. 误差度量方法的选择。

常用的预测模型误差度量方法有：均方误差（Mean Square Error，MSE）、均方根误差（Root Mean Square Error，RMSE）、平均绝对误差（Mean Absolute Error，MAE）、中位数绝对误差（Median Absolute Error，MAE）、平均绝对百分比误差（Mean Absolute Percentage Error，MAPE）、Mean Relative Absolute Error（MRAE）等。Hyndman 和 Koehler（2006）[47]分析了多种误差度量方法的优劣，提出并推荐采用平均绝对标准化误差（Mean Absolute Scaled Error，MASE）来度量预测误差，同时该方法也是 M4-competition 采用的误差度量方法之一。因而本文也采用 MASE 来度量模型误差。

2. 预测模型的回顾。

时间序列建模方法主要有如下几种：

（1）传统时序建模方法，包括差分整合移动平均自回归模型（Autoregressive Integrated Moving Average model，ARIMA）、自回归条件异方差模型（Autoregressive Conditional Heteroskedasticity model，ARCH）和动态线性回归模型（Dynamic Linear Regression models）等，这类方法主要适用于小规模、单变量的预测。

（2）基于时间序列分解的方法。这类方法先将时间序列按照一定的方法分解为几部分，然后分别对分解后的序列进行建模，最后重构组合并预测，按照分解方法的不同，主要又分为：

STH 分解[48]，即将时间序列分解为周期项（Season）、趋势项（Trend）、节假日项（Holiday）。

这种方法有许多成功的应用，但主要的问题在于周期的确定，趋势改变点的检测，需要预先知道节假日等。

将时间序列看作信号按信号分析的方法来处理，主要分解方法有傅里叶分解（Fourier Transformation，FT）、小波分解（Wavelet Decomposition，WD）、经验模态分解（Empirical Mode Decomposition，EMD）[49]和变分模态分解（Variational Mode Decomposition，VMD）[50]等。在实际应用中 EMD 可能存在模态混叠现象，分离后的信号包含多个频率，影响精度；VMD 的主要局限性是边界效应和突发信号，长期模态会随着时间的推移而急剧变化并且会在全局范围内重叠，以及需要预先定义模态数等。

LN 分解，即将时间序列数据分为线性项（L）和非线性项（N）。

（3）机器学习模型。一般将问题转化为监督学习，通常采用 Extreme Gradient Boosting（XGBoost）[51]、Support Vector Regression（SVR）、随机森林（Random Forest）等方法：这类方法一般需要进行特征选择及提取，有着较广泛的应用。

（4）深度学习神经网络模型。用于时间序列预测的典型模型有长短期记忆网络（Long Short-Term Memory，LSTM）[52]、循环神经网络（Recurrent Neural Network，RNN）、门控循环神经网络（Gated Recurrent Unit，GRU）和时间卷积网络（Temporal Convolutional Networks，TCN）[53]等。近期热门的深度学习神经网络有 LSTNet[54]、TPA-LSTM[55]和 Memory Time-series Network（MTNet）[56]等。基于神经网络和深度学习的时间序列预测模型越来越受到学者的重视，是近几年的研究热点。

（5）基于图模型的时间序列预测。应用较多的是隐马尔科夫模型（Hidden Markov Model，HMM），一般来说对于没有复杂记忆的时间序列，HMM 可以较好地进行预测，如果时间序列是复杂历史依赖的，HMM 预测结果可能变差。

3. 预测模型的构建。

基于预测模型回顾总结，结合本文研究的内容，采用 DeepAR 方法对 4 个度量指标（银行指数月收益率，银行指数月波动率，银行间同业拆借加权平均利率以及商业银行流动性比例）进行预测。

DeepAR 是亚马逊于 2017 年提出的基于自回归循环神经网络的时间序列预测方法，该方法相较于其他方法的优点有：（1）仅需要少量的特征工程；（2）采用蒙特卡洛进行概率预测；（3）通过相似数据的学习，可以对仅有一点历史数据或者根本没有历史数据的项目进行预测；（4）不一定是通过高斯噪声实现，允许用户选择一个适合数据统计特性的似然函数。正是由于以上优点，特别是优点（1）和（3）非常适合本文研究的数据，故选择该方法建模。

本文分别对 4 个度量指标采用 DeepAR 方法进行建模，在建模中将 2012 年 1 月到 2018 年 6 月的数据作为训练数据，将 2018 年 7 月到 12 月的数据作为测试数据，进行单步预测，也即每次仅预测一个月的数值。考虑到数据为月度数据，对时效性要求相对不高，为提高预测结果的准确度，在预测测试集第二个月的数据时，将第一个月的真实值放入训练集，重新训练模型，再预测第二个月的数值，直到最后一个测试集数据。DeepAR 中的循环神经网络采用 LSTM 网络，主要超参数的确定采用网格搜索，概率分布采用学生氏分布。

4. 预测结果及分析。

采用 DeepAR 方法对 4 个指标进行建模，预测结果如图 3 所示，其中银行间同业拆借加权平均利率和商业银行流动性比例纵轴为百分比。图 3 给出了 4 个指标的真实值（只给出了 2018 年 1 月到 2018 年 12 月部分）和预测值曲线图，竖线是训练集与测试集的分割线，深色阴影部分代表 50% 的预测区间，浅色阴影部分代表 90% 的预测区间，预测值采用中位数。从图中可以得出预测值基本反映了真实值的走势。

为了对比分析还采用了 MQ-CNN（Multi-Horizon forecasting with a Convolutional Neural Network（CNN）as an encoder）、MQ-RNN[57]、Non-Parametric Time Series Forecaster（NPTS）[58]、ARIMA、指数平滑法（ETS）和多层感知器（Multi-Layer Perceptron，MLP）6 种方法分别对 4 个指标进行预测，得到测试集的 MASE 如表 6 到表 9 所示。其中：NPTS 采用指数核（Exponential Kernel）函数；ARIMA 自回归过程阶数 p，差分阶数 d，移动平均阶数 q 采用 Hyndman（2008）[59]的方法确定；ETS 误差项（Error）、趋势项（Trend）、周期项（Seasonal）为无、乘性或加性，由 Hyndman 等（2008）[60]给出的算法自动选择；MLP 有 2 个隐藏层，每层 40 个单元。每种方法均计算 10 次，求取平均值，减少随机因素影响。

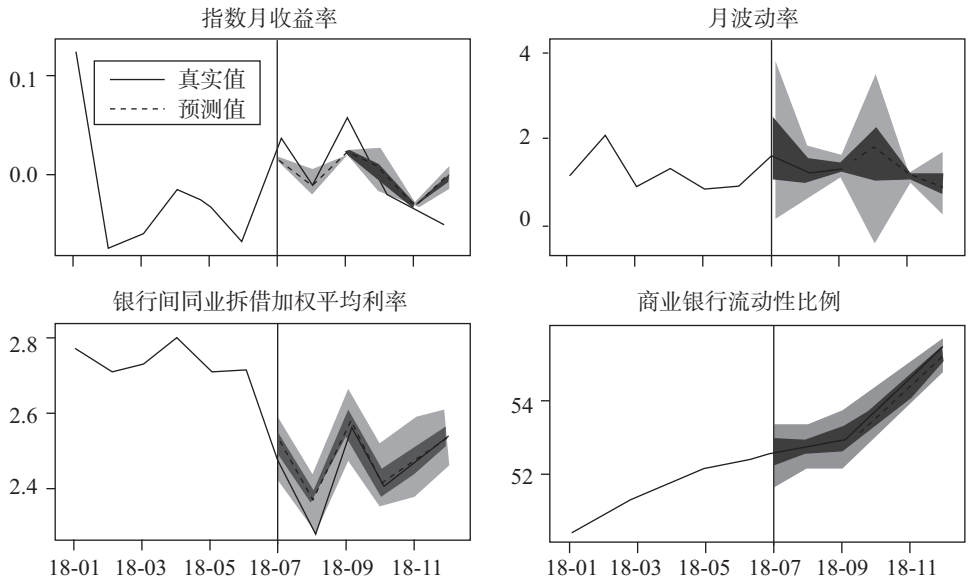

图 3　采用 DeepAR 方法预测结果

表 6 至表 9 给出了上述 7 种方法在 4 个指标测试集（2018 年 7 月至 2018 年 12 月）的 MASE 及其平均值，通过比较所有预测时间 MASE 的平均值可得，DeepAR 方法在银行指数月收益率、银行指数月波动率和银行间同业拆借加权平均利率 3 个指标的预测中，优于其他 6 种方法，而在商业银行流动性比例指标的预测中 DeepAR 与 MQ-CNN 和 ARIMA 方法性能相当。

表 6　　银行指数月收益率测试集的 MASE

年月	MQCNN	MQRNN	NPTS	ARIMA	ETS	MLP	DeepAR
2018 年 7 月	1. 522	1. 497	0. 576	0. 521	0. 401	0. 555	0. 632
2018 年 8 月	0. 699	0. 617	0. 254	0. 197	0. 270	0. 153	0. 150
2018 年 9 月	0. 932	0. 878	1. 065	0. 772	0. 581	0. 569	0. 768
2018 年 10 月	0. 851	0. 806	0. 262	0. 181	0. 300	0. 439	0. 163
2018 年 11 月	0. 389	0. 343	0. 887	0. 509	0. 603	0. 252	0. 490
2018 年 12 月	0. 269	0. 277	0. 783	0. 815	0. 887	1. 293	0. 746
平均	0. 777	0. 736	0. 638	0. 499	0. 507	0. 543	0. 492

表 7　　银行指数月波动率测试集的 MASE

年月	MQCNN	MQRNN	NPTS	ARIMA	ETS	MLP	DeepAR
2018 年 7 月	0. 869	0. 858	0. 940	0. 482	0. 882	0. 927	0. 776
2018 年 8 月	0. 977	0. 864	0. 460	0. 426	0. 259	0. 468	0. 341
2018 年 9 月	0. 208	0. 246	0. 657	0. 434	0. 079	0. 166	0. 115
2018 年 10 月	0. 896	0. 899	1. 556	0. 972	0. 951	0. 934	0. 931
2018 年 11 月	1. 739	1. 620	0. 202	0. 439	0. 877	1. 148	0. 830
2018 年 12 月	0. 168	0. 236	0. 050	1. 086	0. 690	0. 365	0. 426
平均	0. 809	0. 787	0. 644	0. 639	0. 623	0. 668	0. 570

表 8　　**银行间同业拆借加权平均利率测试集的 MASE**

年月	MQCNN	MQRNN	NPTS	ARIMA	ETS	MLP	DeepAR
2018 年 7 月	0. 325	0. 324	0. 411	0. 289	0. 263	0. 347	0. 303
2018 年 8 月	0. 062	0. 076	0. 727	0. 395	0. 364	0. 258	0. 376
2018 年 9 月	0. 490	0. 470	0. 410	0. 159	0. 141	0. 360	0. 145
2018 年 10 月	0. 464	0. 417	0. 371	0. 266	0. 139	0. 226	0. 098
2018 年 11 月	0. 183	0. 167	0. 196	0. 051	0. 092	0. 065	0. 091
2018 年 12 月	0. 047	0. 060	0. 280	0. 047	0. 089	0. 085	0. 055
平均	0. 262	0. 252	0. 399	0. 201	0. 181	0. 224	0. 178

表 9　　**商业银行流动性比例测试集的 MASE**

年月	MQCNN	MQRNN	NPTS	ARIMA	ETS	MLP	DeepAR
2018 年 7 月	0. 089	0. 083	2. 733	0. 034	0. 086	0. 082	0. 095
2018 年 8 月	0. 038	0. 094	3. 001	0. 039	0. 061	0. 077	0. 022
2018 年 9 月	0. 029	0. 084	3. 209	0. 030	0. 083	0. 060	0. 014
2018 年 10 月	0. 341	0. 412	3. 451	0. 339	0. 234	0. 307	0. 341
2018 年 11 月	0. 021	0. 397	3. 707	0. 049	0. 249	0. 275	0. 017
2018 年 12 月	0. 015	0. 380	3. 907	0. 058	0. 189	0. 295	0. 054
平均	0. 089	0. 242	3. 335	0. 091	0. 150	0. 183	0. 090

（二）预警模型的构建与结果分析

依据本文思路，银行系统性风险预警模型的主要任务是通过各指标的历史数据来判断未来一段时期是否有风险。这里主要有两个问题需要解决：一是构建指标；二是依据历史数据判断风险。本文之前依次解决了指标选取问题、银行系统性风险识别问题和指标预测问题，构建银行系统性风险预警模型的步骤则为：首先依据历史数据，采用 DeepAR 模型分别预测 4 个指标下一月的预测值，然后将预测值连同一段区间的历史数据输入到银行系统性风险识别模型中，由风险识别模型判断下一个月是否有银行系统性风险。

依据上述步骤，将 4 个指标测试集的预测值连同训练集数据逐步输入到预警模型，为了对比分析，将真实值也输入到同一模型，得到结果如表 10 所示。

表 10　　**银行系统性风险预警模型结果**

年月	真实值得分	有无风险	预测值得分	有无风险
2018 年 7 月	1. 201	无	1. 109	无
2018 年 8 月	0. 545	无	0. 591	无
2018 年 9 月	0. 962	无	0. 810	无
2018 年 10 月	1. 216	有	1. 165	有
2018 年 11 月	0. 623	无	0. 722	无
2018 年 12 月	0. 797	无	0. 928	无

通过表 10 结合表 5 可得出，虽然是逐月分别将真实值和预测值输入识别模型，但两者都识别出 2018 年 10 月这一风险相对较大的点，这与表 5 中一次性输入全部真实数据得到的结果是一致的。预警模型依据预测值，正确地识别出了有风险的数据，说明本文构建的银行系统性风险预警模型能够在一定程度上对银行系统性风险进行预警。

六、总结与建议

（一）总结

本文构建了银行系统性风险预警模型，在构建中主要解决了三个方面的问题：（1）预警指标的构建问题。在文献回顾的基础上，结合本文的研究内容，选取了宏观经济指标和银行业相关指标两方面的 24 个度量指标，然后采用方差阈值法筛选出 11 个指标，最后提出了基于孤立森林法的指标筛选方法，筛选出最终的 4 个度量指标：银行指数月波动率、银行间同业拆借加权平均利率、商业银行流动性比例以及银行指数月收益率，对筛选结果进行解释，并对入选指标与银行系统性风险的逻辑关系进行了分析。（2）银行系统性风险识别问题。在银行系统性风险预警模型构建中需要对银行系统性风险进行识别，此问题可以

转换为异常检测问题，即数据异常程度越大，说明风险越高。在识别模型的构建中首先采用 Z-score 法和 LLE 法对数据进行预处理，接着提出了一种基于角度的离群点检测法、孤立森林法和局部异常因子法相结合的异常检测方法，对银行系统性风险进行识别，并将识别结果与历史事件进行对照解释。（3）指标预测问题。本文将 4 个指标看作时间序列，采用 DeepAR 方法分别对指标进行预测，为了比较预测结果，还采用了 MQ-CNN、ARIMA 等 6 种方法分别对 4 个指标进行预测，并对预测结果进行比较和分析。

在以上三方面工作的基础上，构建了银行系统性风险预警模型，并将预测值的预警结果、真实值的预警结果以及一次性输入全部真实数据的风险识别结果相比较，三种情况下识别出的高风险点是一致的，证明银行系统性风险预警模型在一定程度上能够有效对风险进行预警。

（二）建议

1. 通过银行流动性风险的大小来窥测全面风险。

企业破产的重要原因之一是资金链断裂，只要资金链能够维持，则短期的盈利能力下降或营运能力降低等均不会危及企业的生存。银行也属于企业，同样具有该特性，虽然银行面临众多风险，但在各风险中流动性风险是重中之重。除流动性比例能够直接反映银行的流动性风险外，银行同业拆借利率由于能够直接反映银行获取资金的成本，间接反映银行获取资金的难易程度，其高低也可以衡量银行的流动性风险。

流动性是决定银行能否正常运转的重要因素，银行是否具有流动性取决于其是否能够随时以合理价格获得满足客户提取的资金。银行间同业业务（例如同业拆借）在流动性的不均衡中发挥了调节作用，使银行系统能够正常运转，而当系统中流动性不足，则会有部分银行流动性需求无法满足。当通过负债不能以适当成本获取资金时，解决方法只能是通过出售资产以满足流动性资金需求。银行同业直接和间接的关联使银行间流动性风险具有较强的传染性，银行解决流动性需求的各种行为在银行网络间互相影响，若没有强有力的第三方介入（例如中央银行等），资产折价抛售可能会演变为一致性行为，进而使资产迅速变现能力下降、资产减值。流动性短缺的加剧促使信用违约发生，负面效应外溢，进一步传染到资本市场，使投资者产生恐慌、信心受挫，对系统中的银行产生不信任，可能会采取提款、撤资等应对行为，当这一风险在各银行间同时发生，出现“挤兑”，流动性风险和传染性风险交织上升，进一步演化为流动性危机，极易引发银行系统性风险，因此系统性风险预警应该重点关注流动性风险以及与之相随的传染性风险。银行资金链的张弛有度，也即单个银行和整体银行系统的适度的流动性是持续发展的关键所在，因此可以重点关注流动性风险，通过流动性风险的变化及时预警银行系统性风险。

2. 通过上市银行市场表现及时预测风险可能性。

在我国的金融结构里，商业银行和证券市场（包括股市与债市）作为主要的融资机构，相互之间具有联动性，尤其是在现代资本市场联系越来越紧密的情况下，一方的变化会促进或制约另外一方。例如，当商业银行信贷扩张，则资本市场资金充裕，股市也会上涨，反过来股市的繁荣又能够给商业银行“补血”，商业银行获得充足的资本金后，又能够继续扩充信贷，循环往复使两者之间形成良好的共生关系。同时，二者之间又会相互制约，一方的不景气会传播与影响另一方，例如股市下跌会影响商业银行收益，商业银行信贷收缩会影响股市的上涨。

银行股的表现也是股市涨跌的风向标，银行指数直接反映银行业股票的整体价格水平，间接反映银行经营状况，而经营状况决定了银行的持续发展能力。因此，关注银行股指数的表现和变化，可以尽早发现风险的高低。

3. 通过内部风险管理和外部风险监管防范风险。

银行系统由各家银行共同组成，每家银行的健康发展对于整个银行系统的安全都有重要作用。推动银行承担起风险防控的主体责任，守好风险防控第一道防线非常重要。风险控制的程度对银行经营的短期利益和长期利益有着不同的影响，短期放松监管能加快银行扩大规模、增加利润，但长期却埋下了风险隐患，对银行长期发展不利。风险管理是商业银行安身立命的基石，它的作用尤其是在经济衰退时尤为凸显，因此商业银行应将银行各项主要风险全部纳入统一管理框架，进行积极主动的风险管理，加强风险监测、人员管理、合规教育，制定适度的风险承担目标。

银监会作为银行的主要监管机构，2000 年以后出台了多项监管指引，使银行的经营日益规范，但宏

观环境日趋复杂，监管当局仍须不断提高监管能力，注意风险的隐蔽性和滞后性，注重风险管理的前瞻性和计量化，通过银行系统性风险预警来及时发现危机，提高对风险的应对能力，以此来适应外部不断变化的环境，保障银行系统的稳健发展，维护国家金融安全。

参考文献

[1] Kaminsky G L，Reinhart C M. On Crises，Contagion and Confusion [J]. Journal of International Economics，2000，51 (1)：145-168.

[2] Demirgüç-Kunt A，Detragiache E. Does Deposit Insurance Increase Banking System Stability? An Empirical Investigation [J]. Journal of Monetary Fconomics，2002，49 (7)：1373-1406.

[3] 孙强，崔光华. 我国银行业系统性风险预警指标体系设计与实证分析 [J]. 中央财经大学学报，2017 (2)：43-51.

[4] 刘松林，王晓娟，王赛. 经济新常态下商业银行风险预警指标体系构建 [J]. 统计与决策，2018 (23)：160-163.

[5] 赵丹丹，丁建臣. 中国银行业系统性风险预警研究——基于 SVM 模型的建模分析 [J]. 国际商务 (对外经济贸易大学学报)，2019 (4)：100-113.

[6] 肖璞，刘轶，杨苏梅. 相互关联性、风险溢出与系统重要性银行识别 [J]. 金融研究，2012 (12)：96-106.

[7] 陈忠阳，刘志洋. 国有大型商业银行系统性风险贡献度真的高吗？——来自中国上市商业银行股票收益率的证据 [J]. 财贸经济，2013 (9)：57-66.

[8] 田海山，周玉琴，吴恒煜. 广义双曲线分布族下的动态系统性风险研究 [J]. 数理统计与管理，2017 (1)：59-72.

[9] 蒋海，张锦意. 商业银行尾部风险网络关联性与系统性风险——基于中国上市银行的实证检验 [J]. 财贸经济，2018 (8)：50-65.

[10] 吴敏灵. 我国上市银行系统性风险度量实证研究 [J]. 大连理工大学学报 (社会科学版)，2018 (2)：24-31.

[11] Patro D K，Qi M，Sun X，et al. A Simple Indicator of Systemic Risk [J]. Journal of Financial Stability，2013，9 (1)：105-116.

[12] 高国华，潘英丽. 基于动态相关性的我国银行系统性风险度量研究 [J]. 管理评论，2013 (1)：9-15.

[13] 王琳，沈沛龙. 基于 DCC-GARCH 模型的中国上市银行系统性风险研究 [J]. 财经理论与实践，2017 (1)：24-29.

[14] 毛昌梅，韩景倜，刘举胜. 基于复杂网络的银行波动溢出效应研究 [J]. 复杂系统与复杂性科学，2020 (2)：11-21.

[15] 陆静，王捷. 基于贝叶斯网络的商业银行全面风险预警系统 [J]. 系统工程理论与实践，2012 (2)：225-235.

[16] 汪冬华，黄康，龚朴. 我国商业银行整体风险度量及其敏感性分析——基于我国商业银行财务数据和金融市场公开数据 [J]. 系统工程理论与实践，2013 (2)：284-295.

[17] 方意. 系统性风险的传染渠道与度量研究——兼论宏观审慎政策实施 [J]. 管理世界，2016 (8)：32-57，187.

[18] 吴念鲁，徐丽丽，苗海宾. 我国银行同业之间流动性风险传染研究——基于复杂网络理论分析视角 [J]. 国际金融研究，2017 (7)：34-43.

[19] 赵越，韩永辉. 中国银行业系统性风险传染研究——基于涵盖银行间传染风险与流动性风险的量化模型 [J]. 金融学季刊，2016 (3)：43-64.

[20] 吴卫星，邵旭方，吴锟. 中国商业银行流动性风险传染特征分析——基于商业银行同业负债的时间序列数据 [J]. 国际商务 (对外经济贸易大学学报)，2016 (4)：81-92.

[21] 杨霞，吴林. 我国银行业系统性风险预警研究 [J]. 统计与决策，2015 (10)：147-150.

[22] 谢贤芬，王斌会. 银行宏观安全状态指数及预警监测分析 [J]. 数理统计与管理，2017 (5)：906-918.

[23] Li S，Wang M，He J. Prediction of Banking Systemic Risk Based on Support Vector Machine [J]. Mathematical Problems in Engineering，2013.

[24] Adrian T，Brunnermeier M K. CoVar [J]. American Economic Review，2016，106 (7)：1705-1741.

[25] Acharya V V，Pedersen L H，Philippon T，et al. Measuring Systemic Risk [J]. The Review of Financial Studies，2017，30 (1)：2-47.

[26] 李楠，樊鹏英，王艳，陈敏. 新资本协议下基于半参数估计的商业银行市场风险 VaR 研究 [J]. 数理统计与管理，2014 (6)：1080-1089.

[27] 戴丽娜. 商业银行操作风险的度量——基于非参数方法 [J]. 数理统计与管理，2017 (3)：541-549.

[28] Roweis S T，Saul L K. Nonlinear Dimensionality Reduction by Locally Linear Embedding [J]. Science，2000，290 (5500)：2323-2326.

[29] Kriegel H P，Schubert M，Zimek A. Angle-based Outlier Detection in High-dimensional Data [C]. The 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining，2008：444-452.

[30] Liu F T，Ting K M，Zhou Z H. Isolation Forest [C]. The 8th IEEE International Conference on Data Mining，IEEE，2008：413-422.

[31] Liu F T，Ting K M，Zhou Z H. Isolation-based Anomaly Detection [J]. ACM Transactions on Knowledge Discovery from Data (TKDD)，2012，6 (1)：1-39.

[32] Breunig M M，Kriegel H P，Ng R T，et al. LOF：Identifying Density-based Local Outliers [C]. ACM SIGMOD International Conference on Management of Data，2000：93-104.

[33] Salinas D, Flunkert V, Gasthaus J, et al. DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks [J]. International Journal of Forecasting, 2019.
[34] 胡历芳, 刘纪鹏. 扩大对外开放背景下中国银行业系统性风险的预警与防范 [J]. 经济与管理研究, 2019 (3): 21-28.
[35] Cuturi M, Blondel M. Soft-DTW: A Differentiable Loss Function for Time-series [C]. International Conference on Machine Learning, 2017: 894-903.
[36] 李智, 牛晓健. 基于内生网络的银行间传染风险特征 [J]. 北京理工大学学报 (社会科学版), 2017 (2): 67-71, 102.
[37] 邓晶, 曹诗男, 潘焕学, 秦涛. 基于银行间市场网络的系统性风险传染研究 [J]. 复杂系统与复杂性科学, 2013, 10 (4): 76-85.
[38] 方意, 郑子文. 系统性风险在银行间的传染路径研究——基于持有共同资产网络模型 [J]. 国际金融研究, 2016 (6): 61-72.
[39] 李晓伟, 宗计川. 金融稳定视角下的流动性风险传染研究新进展 [J]. 经济学动态, 2018 (4): 138-147.
[40] 沈沛龙, 李志楠. 投资者恐慌情绪对银行间风险传染影响研究 [J]. 财贸研究, 2020, 31 (3): 59-71, 110.
[41] Schölkopf B, Williamson R C, Smola A J, et al. Support Vector Method for Novelty Detection [C]. Advances in Neural Information Processing Systems, 2000: 582-588.
[42] Lazarevic A, Kumar V. Feature Bagging for Outlier Detection [C]. The 11th ACM SIGKDD International Conference on Knowledge Discovery in Data Mining, 2005: 157-166.
[43] Kalaycı İ, Ercan T. Anomaly Detection in Wireless Sensor Networks Data by Using Histogram Based Outlier Score Method [C]. the 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), IEEE, 2018: 1-6.
[44] Aggarwal C C, Sathe S. Theoretical Foundations and Algorithms for Outlier Ensembles [J]. ACM SIGKDD Explorations Newsletter, 2015, 17 (1): 24-47.
[45] Du-Mouchel W H. Estimating the Stable Index α in Order to Measure Tail Thickness: A Critique [J]. The Annals of Statistics, 1983: 1019-1031.
[46] 林宇, 黄迅, 淳伟德, 黄登仕. 基于 ODR-ADASYN-SVM 的极端金融风险预警研究 [J]. 管理科学学报, 2016 (5): 87-101.
[47] Hyndman R J, Koehler A B. Another Look at Measures of Forecast Accuracy [J]. International Journal of Forecasting, 2006, 22 (4): 679-688.
[48] Taylor S J, Letham B. Forecasting at Scale [J]. The American Statistician, 2018, 72 (1): 37-45.
[49] Huang N E, Shen Z, Long S R, et al. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis [J]. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 1998, 454 (1971): 903-995.
[50] Dragomiretskiy K, Zosso D. Variational Mode Decomposition [J]. IEEE Transactions on Signal Processing, 2013, 62 (3): 531-544.
[51] Chen T, Guestrin C. Xgboost: A Scalable Tree Boosting System [C]. The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016: 785-794.
[52] Hochreiter S, Schmidhuber J. Long Short-term Memory [J]. Neural Computation, 1997, 9 (8): 1735-1780.
[53] Bai S, Kolter J Z, Koltun V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling [J]. arXiv preprint arXiv: 1803.01271, 2018.
[54] Lai G, Chang W C, Yang Y, et al. Modeling Long and Short-term Temporal Patterns with Deep Neural Networks [C]. The 41st international ACM SIGIR Conference on Research & Development in Information Retrieval, 2018: 95-104.
[55] Shih S Y, Sun F K, Lee H. Temporal Pattern Attention for Multivariate Time Series Forecasting [J]. Machine Learning, 2019, 108 (8/9): 1421-1441.
[56] Chang Y Y, Sun F Y, Wu Y H, et al. A Memory-network Based Solution for Multivariate Time-series Forecasting [J]. arXiv preprint arXiv: 1809.02105, 2018.
[57] Wen R, Torkkola K, Narayanaswamy B, et al. A Multi-horizon Quantile Recurrent Forecaster [J]. arXiv preprint arXiv: 1711.11053, 2017.
[58] Gasthaus J. Non-parametric Time Series Forecaster [R]. Technical Report, Amazon, 2016.
[59] Hyndman R J, Khandakar Y. Automatic Time Series Forecasting: The Forecast Package for R [J]. Journal of Statistical Software, 2008, 27 (1): 1-22.
[60] Hyndman R, Koehler A B, Ord J K, et al. Forecasting with Exponential Smoothing: The State Space Approach [M]. Springer Science & Business Media, 2008.

(责任编辑: 韩 媛 张安平)

CEO与CFO的薪酬契约安排、权力配置对会计错报的影响研究

——基于CEO、CFO制衡的视角

The Effect of CEO and CFO Compensation Contract Arrangements, Power Difference on Accounting Misstatement: Based on Checks and Balances between CEO and CFO

马　晨　李晨溪　敬舒贻

MA Chen　LI Chen-xi　JING Shu-yi

［摘　要］本文从理论与实证角度检验了CEO、CFO薪酬契约安排对会计错报的作用机制，并考虑这种影响是否会受到CEO与CFO权力差异的调节，得到以下研究结论：第一，CEO与CFO薪酬契约安排差异均会正向影响公司会计错报；第二，二者之间的薪酬契约安排差异对公司发生会计错报具有正向的影响；第三，薪酬契约安排差异对会计错报的正向影响会受到CEO与CFO权力差异的负向调节，CEO权力越大于CFO时，CEO越能够通过其他渠道实现机会主义行为，而不仅仅限于会计操纵。在这种情形下，CEO与CFO薪酬契约安排差异对会计错报的正向影响作用反而会被弱化。此外，本文在进一步分析中还发现，随着CEO与CFO薪酬契约安排差异的增加，不但公司发生会计差错的概率会提升，在发生的差错中，更加可能是由舞弊而不是人为疏忽引起的。

［关键词］薪酬契约安排　权力配置　差异　会计错报

［中图分类号］F23　［文献标识码］A　［文章编号］1000-1549（2020）10-0054-16

Abstract: This paper studies the effect of CEO and CFO compensation contract arrangements (CA) on accounting misstatement, considering the moderating effect of CEO and CFO power difference. The results indicate that CEO and CFO CA positively affects accounting misstatement. Besides, the difference between CEO and CFO CA has positive effect on accounting misstatement, showing that as the difference between CEO and CFO CA increases, the incidence of misstatement increases. Moreover, the effect of the difference between CEO and CFO CA on misstatement is negatively moderated by the difference between CEO and CFO power, indicating that as the difference between CEO and CFO power increases, CEO is more likely to engage in opportunistic behavior through other channels, not only confined to accounting manipulations, in which the effect of the difference between CEO and CFO CA on misstatement is attenuated. Moreover, the additional analysis shows that, as the difference between CEO and CFO CA increases, not only the incidence of misstatement increases, but also the misstatement is much more related to fraud rather than negligence.

Key words: Compensation contract arrangement　Power allocation　Difference　Accounting-misstatements

［收稿日期］2019-02-05

［作者简介］马晨，男，1983年9月生，西北大学经济管理学院副教授，研究方向为审计与财务重述；李晨溪，女，1996年8月生，西北大学经济管理学院研究生，研究方向为审计与公司治理；敬舒贻，女，1996年11月生，西北大学经济管理学院研究生，研究方向为财务会计。

［基金项目］国家自然科学基金项目“会计师事务所异质性对财务重述的影响研究——基于媒介环境的调节效应分析”（项目编号：71502138）；国家自然科学基金项目“母子公司距离、财务报告质量与资源配置效率”（项目编号：71572144）；教育部人文社科青年基金项目“空间分布、关联交易对会计透明度影响研究”（项目编号：20YJC630098）；陕西省社会科学基金项目“陕西省公司空间分布、关联交易对会计信息质量影响研究”（项目编号：2019S015）；陕西省自然科学基金项目“外部审计、媒介环境对财务重述的影响研究”（项目编号：2015JQ7281）；陕西省教育厅项目“财务重述对年报披露及时性的影响及其经济后果研究”（项目编号：15JK1721）。

感谢匿名评审人提出的修改建议，笔者已做了相应修改，本文文责自负。

一、问题的提出

近些年，我国政府频繁出台国有企业限制高管薪酬的政策措施，2017 年 11 月，综艺限薪令的发布，又使薪酬话题成为当今社会的议论焦点。在高管薪酬的波及范围当中，CEO 与 CFO 受到的影响巨大，二者在财务报告的生成和披露方面都负有极其重要的责任。Twitter 于 2015 年向美国证券交易委员会（SEC）提交的信息表明，作为 CFO 的 Anthony Noto，在 2014 年收获了 7 280 万美元的薪酬，在 Twitter 所有高管中居于榜首，是作为 CEO 的 Dick Costolo 的5 240倍。这个例子在一定程度上表明 CFO 的角色和身份发生了转变，不再简单是 CEO 的代理人。也正是因为 CEO 与 CFO 在企业内部的核心职位，他们的薪酬契约安排动机应当如何设计、薪酬动机的差异如何安排，以此降低代理成本，从而提升财务报告质量也是学术界和实务界所关注的热点问题。

上市公司发生的财务丑闻可以被归结于高管薪酬契约的设计（Burns 和 Kedia，2006[1]）。大量研究证实了 CEO 对会计信息质量具有重大影响（Dow 和 Raposo，2005[2]；Jiang 等，2010[3]；Chava 和 Purnanandam，2010[4]；Feng 等，2011[5]；林大庞和苏冬蔚，2012[6]）。Jiang 等（2010）[3]认为 CFO 的薪酬激励可能会导致盈余管理。美国税务署（Internal Revenue Service）委员 Mark Everson 在 2006 年认为 CFO 应当履行监督糖罐的责任（minding the cookie jars），其薪酬不应当包括股票期权，而应当主要由可观的、固定的薪酬构成。CEO、CFO 的薪酬契约安排动机究竟是否能够起到应有的作用，亦或增加代理成本，降低会计信息质量，是值得学术界继续关注的热点话题。

另一方面，CEO、CFO 薪酬契约安排动机差异对会计信息质量影响的研究却鲜有人问津，而且相关研究更多地还是关注 CEO 而非 CFO 对会计信息质量的影响，尽管对 CFO 的薪酬契约安排动机也有考虑。这可能主要是由于 CEO 的薪酬契约安排动机要强于 CFO（Jiang 等，2010[3]）。此外，CFO 很多时候作为 CEO 的代理人，如果不遵循 CEO 的偏好，是有可能被有能力的 CEO 更换掉的。因此，CFO 可能不会对自身的薪酬激励作出反应，而更多的是针对 CEO 的薪酬激励作出反应。但是，Jiang 等（2010）[3]认为，检验 CFO 薪酬契约安排动机对财务报告的影响是非常有必要的，这是由于 CFO 在财务报告方面起着至关重要的作用。Feng 等（2011）[5]认为 CFO 在进行会计操纵方面有着非常独特的身份，比如重组交易、选择不合适的会计处理以及做假账。世界各国的财务丑闻，如安然、世通、Qwest 以及 Adelphi 等事件都表明 CFO 对会计信息质量的影响是至关重要的。世通的前 CFO Scott Sullivan 也承认了自己的会计造假行为。此外，2002 年的萨班斯-奥克斯利法案（the Sarbanes-Oxley Act）正式要求 CFO 与 CEO 对财务信息的准确性以及完整性均要承担个人责任。我国的会计相关法律也对此作出了相关规定。

鉴于以上，CEO 薪酬契约安排动机对财务报告质量可能存在负向影响，而 CFO 薪酬契约安排动机在对财务报告质量存在负向影响的同时也有着需要承担监督 CEO 特别是同财务报告方面相关的职责，即双向影响。CEO、CFO 在薪酬契约安排动机、权力大小方面都存在着差异，这种差异大小应当如何安排？本文从薪酬契约安排（compensation contract arrangements，简称 CA）与管理层权力理论视角层层递进地分析和检验并得出如下结论：第一，CEO 与 CFO 的薪酬契约安排差异的确会影响到会计错报；第二，二者之间的薪酬契约安排差异均会影响公司会计错报；第三，CEO 与 CFO 薪酬契约安排差异对会计错报的影响作用还会受到 CEO 与 CFO 权力差异的调节。

本文的贡献或创新之处有以下几点：第一，拓展了有关 CEO、CFO 薪酬问题的研究成果。以往都是分别关注 CEO、CFO 个人特征对财务报告质量的影响，本文将在分析 CEO、CFO 薪酬契约安排动机影响的基础上提供 CEO、CFO 薪酬契约安排差异对会计错报影响的证据，补充了 Jiang 等（2010）[3]，Feng 等（2011）[5]的研究。他们检验了 CEO 与 CFO 股权动机对盈余管理的影响差异，发现 CFO 的股权动机要大于 CEO。鉴于 CFO 在会计操纵方面的重要作用，本文进一步将 CEO 与 CFO 薪酬契约安排动机结合起来，深入分析二者薪酬契约安排差异对财务报告质量的影响。第二，进一步拓展了管理层权力的研究。本文的研究发现 CEO 与 CFO 权力差异的增加，反而会弱化 CEO、CFO 薪酬契约安排差异对会计错报的影响，表明当 CEO 权力相对于 CFO 越大时，

CEO 会通过其他渠道实现机会主义行为，而不仅仅围绕薪酬契约、会计报表操纵。

余文结构安排如下：第二部分为文献回顾与研究假设；第三部分为样本来源与研究设计；第四部分为实证检验与结果分析；第五部分为研究结论。

二、文献回顾与研究假设

（一）CEO、CFO 薪酬契约安排动机对会计信息质量的影响

已有研究证实了上市公司的会计错报或财务重述会给资本市场和广大投资者带来重大的不利影响（Burns 和 Kedia，2006[1]；马晨等，2015[7]）。一系列的文献试图寻找会计差错、财务重述发生的原因。这些研究主要关注 CEO 或整个高管团队的动机（Burn 和 Kedia，2006[1]；马晨和张俊瑞，2012[8]）。王福胜和程富（2014）[9]从高层梯队理论的视角进行分析，认为洞察能力、价值选择及认知层面是高管背景特征的重要表现，高管在管理工作中的偏好选择便是由这些因素决定。一个公司会计政策的制定者或决策者——CFO，其个人特征同样可能对公司的会计政策选择产生影响。然而，令人意外的是，研究 CFO 薪酬契约安排动机对财务报告质量影响的文献较少，毕竟 CFO 是作为具有专业能力的对财务报告质量负责的公司高管。且 CFO 的财务决策行为经常会受到 CEO 的不良干预，由于企业当中的代理冲突矛盾屡见不鲜，因而干预到其执行财务判断的事件时有发生（向锐，2015[10]）。同时，向锐（2015）[10]认为，若 CFO 与 CEO 合谋造假，会给 CFO 职业生涯带来很大污点。因此，从保护自身声誉角度出发，CFO 会及时与董事会内部成员沟通自身熟悉的公司财务战略信息，通过减少信息差的行为，以达到限制 CEO 的机会主义行为的目的。这点从经验数据上也是得到了一定的证实。邱昱芳等（2011）[11]根据证监会组织的对我国上市公司财务负责人的专项调查结果，发现财务负责人的专业能力对会计信息质量有一定的正向影响。刘永丽（2015）[12]发现 CFO 权力越大，越能够根据自身的专业背景对公司的发展和决策作出判断，不受高层关于一些政治任务及其他目的而作出对利润的粉饰要求的影响，会计稳健性应当越高。Ge 等（2011）[13]通过实证研究发现 CFO 个人特征会影响到会计选择。

CFO 通常要负责财务计划、预算、内部控制以及财务报告（Feng 等，2011[5]）。如果出现了会计操纵的情形，则一定意味着 CFO 未能起到应有的监督责任。另外，CFO 如果有意图参与会计操纵，更加能够利用自身的专长去虚增利润。比如，Stuart G. Lasher，Silk Greenhouse 的前 CFO，制造了一系列的虚假记录和工作底稿以作为推迟费用入账的借口。Feng 等（2011）[5]用薪酬契约安排作为薪酬动机的代理变量，分析了 CEO、CFO 自身的薪酬动机对财务报告质量的影响，他们发现 CFO 进行盈余管理不是因为自身的股权动机，而是由于受到具有股权动机的 CEO 的压力导致的。从相关文献可以看出，尽管国内外学者关注过 CEO、CFO 薪酬动机对会计错报的影响，但鲜有关注 CEO 与 CFO 薪酬契约安排的差异对会计信息质量的影响。已有研究主要停留在 CEO 或 CFO 的薪酬激励与盈余管理之间的关系上，而本文对此进行了一定的补充，综合 CEO 与 CFO 薪酬契约安排并关注二者差异对会计错报的影响作用。本文的结论表明 CEO 与 CFO 的薪酬契约安排共同影响到会计信息质量。鉴于上述分析，本文认为薪酬契约安排一方面能够激励 CEO 与 CFO 为了个人薪酬努力地去创造股东财富的增值，但同时也会使得 CEO 与 CFO 有动机操纵财务报表以实现个人利益。针对 CEO、CFO 个人薪酬契约安排动机，本文根据上述分析提出假设 1 和假设 2。

H1：CEO 薪酬契约安排正向影响公司的会计错报。

H2：CFO 薪酬契约安排正向影响公司的会计错报。

（二）CEO 与 CFO 薪酬契约安排动机差异对会计信息质量的影响

在我国，学术界和业界认为，一方面，CFO 应该在公司治理中扮演重要而独特的角色，其既应该负有监督责任，即履行对 CEO 监督特别是财务方面的监督职能，又应该对外客观、公允地披露财务报告，承担与 CEO 相同的法律责任，以切实维护财务报告使用者的合法权益；另一方面，由于公司各种经营决策都离不开财务的可行性分析，所以 CFO 应该参与公司决策，并履行相应的管理职能（林大庞和苏冬蔚，2012[6]）。但是由于 CFO 在履行公司治理职责的过程中尚不具有充分的实力担负起监督 CEO 的任务，

当 CEO 与 CFO 存在利益冲突时，中国现阶段的制度环境和公司治理结构尚不足以支持 CFO 抵御 CEO 的压力（孙光国和郭睿，2015[14]），使得我国 CFO 薪酬激励的公司治理效应可能显著区别于西方发达国家。此外，一般而言，CEO 由董事会聘任，而 CFO 则由 CEO 提名再由董事会通过，因此从理论上说 CFO 是经由 CEO 挑选，其法定性权力应当比 CEO 的更小，以致 CEO 获得公司的薪酬激励程度比 CFO 更大，进而使得 CEO 薪酬比 CFO 薪酬对业绩更加敏感。

然而，很多学者却持有不同的观点。林大庞和苏冬蔚（2012）[6]发现，尽管 CEO 股权激励对盈余管理的抑制作用显著大于 CFO，然而在诱发盈余管理行为方面两者却不存在显著差异。Chava 和 Purnanandam（2010）[4]发现，CFO 较高的薪酬契约安排会产生较高的应计管理，CFO 的薪酬契约安排对总应计项目的影响作用甚至大于 CEO。Jiang 等（2010）[3]发现，相对于 CEO 的权益动机，CFO 的权益动机会导致可操控应计程度更大幅度的增长。他们解释，这是由于 CFO 对财务系统的管理负有最终责任，包括财务报告的编制，这种影响超过了 CEO 薪酬契约安排动机对财务报告质量的影响。此外，盈余管理是一种非常重要的工具，CFO 可以利用自身的专业为其薪酬契约安排动机服务，而 CEO 则有许多其他方面的监督责任以及其他工具可以利用以实现自身利益。Indjejikian 和 Matejka（2009）[15]认为，尽管 CEO 的薪酬契约安排动机更强，然而 CFO 才是财务报告过程中的主要负责人，如果 CFO 薪酬契约安排动机较强，反而会削弱他们在降低利润中所含错报的程度的作用。也就是说，尽管薪酬契约安排动机强有助于刺激 CFO 的工作动力，但可能也会促成其机会主义行为。他们还发现董事会在萨班斯法案之后降低了 CFO 的薪酬契约安排动机，以抑制其会计造假行为。

已有研究更多单独研究 CEO、CFO 薪酬契约安排动机与会计信息质量之间关系，鲜有关注 CEO、CFO 的薪酬契约安排差异对财务报告质量的影响作用。Feng 等（2011）[5]从盈余管理发生的原因出发分析了 CEO、CFO 在这方面的区别。他们利用理论模型研究了 1982 年至 2005 年间共计 499 家存在会计操纵行为的公司，发现 CFO 进行盈余操纵并不一定能获得足够的自身利益，而且 CFO 进行会计操纵是由于 CEO 的压力而不是自身的薪酬动机。

薪酬动机越大，越能够降低代理成本，将高管个人利益同股东利益结合在一起（Jensen 和 Murphy，1990[16]），这点既适用于 CEO，也适用于 CFO。但另一方面，较高的薪酬动机会使得 CEO、CFO 个人为了追求个人薪酬而去操纵会计利润。CEO 薪酬动机更多的是一种机会主义行为，这点在已有文献中被广泛证实（Burns 和 Kedia，2006[1]）。而 CFO 的薪酬契约安排动机更多体现的是两面性，履行对 CEO 在公司财务方面的监督职责，同时具有潜在的机会主义行为（Indjejkian 和 Matejka，2009[15]）。因此，本文认为，考虑高管薪酬契约安排动机对财务报告质量影响时，应当对 CEO 与 CFO 的薪酬契约安排综合考虑，CEO 与 CFO 薪酬契约安排的差异体现了二者对财务报告质量的综合影响。当 CEO 薪酬契约安排高于 CFO 薪酬契约安排时，薪酬契约相对来说更加符合 CEO 而非 CFO 的个人意愿，CFO 监督意愿和能力会降低，监督职责自然难以有效发挥。而且随着二者间薪酬契约安排差异的增加，CFO 更加难以发挥对 CEO 的牵制和制衡作用，公司发生会计错报的概率随之提升。由此，本文提出假设 3。

H3：随着 CEO 与 CFO 薪酬契约安排差异的增加，公司发生会计错报的概率随之提升。

（三）管理层权力对高管薪酬契约安排、会计信息质量的影响研究

Bebchuk 等（2002）[17]提出了薪酬契约“管理层权力假说”，认为不断扩大的管理层权力影响了薪酬契约的激励效果，如果高管可以自定薪酬，契约就是失效的。管理层权力理论认为，薪酬契约安排的提高并不一定意味代理成本的减少，因为企业业绩包含着强烈的噪音，如管理层自身对公司业绩的盈余操纵。因此，在考察薪酬契约安排对财务报告质量影响时，不能忽视管理层权力的作用。在我国，管理层权利有时甚至能够直接干预管理层薪酬体系的设计。权小锋等（2010）[18]对此也进行了证实，他们发现管理层权力越大，薪酬与操纵性业绩之间的敏感性越大，表明随着权力的增长，管理层更倾向于利用盈余操纵获取绩效薪酬。还有很多学者也证实了管理层权力会影响到契约的制定。刘星和徐光伟（2012）[19]发现，高管利用手中权力影响自身契约，导致薪酬具有向下的刚性和向上的弹性。薪酬契约安排存在不对称现象，说明高管利用手中权力具有获取私利的动机。陈修德

（2012）[20]发现，当管理层相对于董事会的权力越大，高管薪酬将越高或是其薪酬与业绩之间的关联性会越小。

另一方面，王克敏和王志超（2007）[21]提出尽管高管权力的增加使得高管可以自定薪酬，但却降低了高管报酬诱发盈余管理的程度。他们认为，我国缺乏投资者保护机制，企业的外部监管机制薄弱，这会导致高管的激励约束机制失灵，进而使得高管利用控制权谋取私人利益的寻租渠道多元化。这样一来，实现个人利益最大化的渠道不单单是会计操纵，管理层操纵报表的动机会下降，转而通过其他渠道去寻求一己私利的实现。他们将这种高管控制权增加描述为从“遵守报酬游戏规则转变为制定报酬游戏规则”。Dow 和 Raposo（2005）[2]提出权力较大的管理层不会基于薪酬目的进行盈余管理，但可能基于其他目的进行盈余管理，且权力较小的管理层可能为提高自身薪酬而进行盈余管理。

以往研究 CEO 权力的文献得出的结论主要集中于 CEO 权力会提升自身薪酬（Bebchuk 等，2002[17]），不利于会计信息质量（Dow 和 Raposo，2005[2]），即管理层权力视角。从这样的视角来说，如果 CEO 权力远远大于 CFO 权力，则更能够有助于 CEO、CFO 薪酬契约安排差异对会计错报的影响作用。但如果 CEO 权力较大，难以被制衡，可能会打破原有的均衡或企业内部形成的制度安排，比如薪酬契约，进而使得薪酬契约难以有效发挥相应的作用。而且 CEO 权力过大时可以通过其他渠道而不仅仅局限于会计操纵去实现一己私利，比如直接追加年薪、奖金制度，减少了选择风险和成本较高的盈余管理行为（王克敏和王志超，2007[21]）。在这种情形下，CEO、CFO 权力差异可能会弱化 CEO、CFO 薪酬契约安排差异对会计错报的影响作用。由此，本文提出竞争性假设 4。

H4a：CEO、CFO 权力差异会强化 CEO、CFO 薪酬契约安排差异对会计错报的影响。

H4b：CEO、CFO 权力差异会弱化 CEO、CFO 薪酬契约安排差异对会计错报的影响。

基于假设 H1～H4，本文的理论检验模型如图 1。

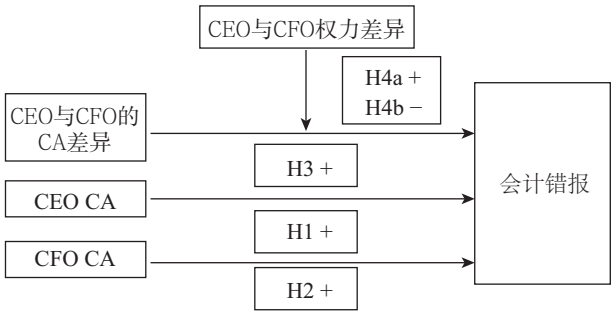

图 1　理论与假设模型

三、样本来源与研究设计

（一）样本来源

本文选取 2003—2015 年中国上市公司的年报数据，通过手工收集、整理，结合巨潮资讯网站搜索更正、差错、调整关键词，最终得到 2005—2014 年共 1 497 个发生错报的观测值①。本文错报样本的数据与会计账务处理直接挂钩，与年报补丁大相径庭，且不包括与股票拆分、发放股利、终止经营、会计政策变化、并购有关的事项。本文将初始样本从两方面进行筛选。首先，删除了 4 个金融保险类数据。原因如下：金融保险行业的会计处理方法不同于日常企业，其较高的薪酬体系受到银监会的监管。其次，剔除了 686 个 CEO、CFO 薪酬契约安排缺省的错报观测值。最终，本文结合其他控制变量得到 697 个错报观测值以及 12 507 个非错报观测值。薪酬契约安排、高管权力以及其他控制变量的数据来源于 CSMAR 数据库。在进行分析时，由于数据的可获得性，分析样本可能有差异。

（二）研究设计

按照 Jensen 和 Murphy（1990）[16]、Jayaraman 和 Milbourn（2012）[22]的研究，本文将薪酬契约安排度量为高管薪酬的变化/（股东财富）企业价值的变化，具体为 $CA = \Delta Comp/(\Delta Earnings)$②。其中，$\Delta Comp$③为高管薪酬的变化量，$\Delta Earnings$ 为净利润的变化量。因此，CEO 的薪酬契约安排 $CA_CEO = \Delta Comp_CEO/(\Delta Earnings)$，而 CFO 的薪酬契约安排为 $CA_CFO = \Delta Comp_CFO/(\Delta Earnings)$。

本文采用 Probit 回归模型，分别对 CEO、CFO 薪

① 披露的错报样本是根据财务重述样本获得的，因此，错报的发生期要早于财务重述的发生期。另外，鉴于 B 股类上市公司的特殊规定，本文的会计错报样本不包含 B 股类公司。

② 鉴于中国资本市场股票价格的波动性，本文未使用股票价值而是使用公司高管更为关注的业绩，即净利润衡量企业价值或股东财富。

③ 本文的高管薪酬主要是现金薪酬，包括基本工资、红利以及奖金。此外，不同于其他国家，股票期权的授予和行权在我国还较少，鉴于数据有限，本文不考虑股票期权。

酬契约安排与公司发生会计错报之间的影响进行了检验。以下为关于假设 H1、假设 H2 的回归模型：

$$Mis_{it}=\alpha_0+\alpha_1 CA_CEO_{it}+\alpha_2 CA_CFO_{it}+\alpha_3 Controls+\varepsilon \quad (1)$$

其中：Mis_{it}为公司 i 发生会计错报的情况，t 年发生会计错报时为 1，否则为 0。*Controls* 为公司层面的控制变量。为了控制公司的盈利能力以及成长情况（马晨等，2014[23]），本文控制了总资产收益率（*ROA*）以及总资产增长率（*Agrow*），还控制了公司的资产规模（Ln*a*）以及资产负债率（*Lev*）。另外，外部审计也同公司发生会计错报有关，因此对会计师事务所规模（*Big*4）以及审计意见类型（*AO*）进行了控制。此外，本文还借鉴 Lobo 和 Zhao（2013）[24]的研究，对市值规模（*MB*，market-to-book ratio）进行了控制。

本文还要通过 Probit 回归模型，以检验 CEO、CFO 薪酬契约安排差异对公司发生会计错报的影响。针对假设 H3 的回归模型如下：

$$Mis_{it}=\alpha_0+\alpha_1 DCA_{it}+\alpha_2 Controls+\varepsilon \quad (2)$$

其中：*DCA* 为 CEO 与 CFO 薪酬契约安排的差异，即 $DCA=CA_CEO-CA_CFO$；*Controls* 为公司层面的控制变量，如前所述。

本文还考虑 CEO 与 CFO 的权力差异是否会调节二者薪酬契约安排差异对会计错报的影响。因此，要对权力差异模型进行相应的设计。借鉴王克敏和王志超（2007）[21]、权小锋和吴世农（2010）[25]的方法，本文用 CEO 是否兼任董事长以及 CEO 是否是公司董事会成员两个虚拟变量来衡量 CEO 的组织权力强度；用 CEO 的任职年限来度量其专家权力；选择机构投资者的持股水平以及股权集中度来度量 CEO 的所有制权力强度。当 CEO 所在企业的机构投资者持股水平低于均值时，第一大股东持股比例与第二大股东持股比例的比值越大，则 CEO 受到外部监督的力量就会越弱。本文还使用虚拟变量 CEO 是否持股以及 CEO 年龄来度量其声誉权力。在这些测度指标的基础上通过直接求和的方式合成 CEO 权力强度的综合指标（*Power_CEO*）①。同理，针对 CFO 生成其权力强度的综合指标②。本文用虚拟变量 CFO 是否是公司董事会成员来衡量 CFO 的组织权力强度；用 CFO 的任职年限来度量其专家权力；还使用虚拟变量 CFO 是否持股以及 CFO 年龄来度量其声誉权力。在这些测度指标的基础上通过直接求和的方式合成 CFO 权力强度的综合指标（*Power_CFO*）。在此基础上，生成 CEO 与 CFO 权力差异指标，$DPower=Power_CEO-Power_CFO$。*DPower* 作为调节变量，本文用其与薪酬契约安排差异形成的交互项（$DCA\times DPower$）以检验假设 H4，回归模型如下：

$$Mis_{it}=\alpha_0+\alpha_1 DCA_{it}+\alpha_2 DCA_{it}\times DPower_{it}+\alpha_3 DPower_{it}+\alpha_4 Controls+\varepsilon \quad (3)$$

其他变量如前所述，具体变量情况见表 4。

表 1　　变量说明表

变量类型	变量名称	经济含义	解释	预测符号
被解释变量	*Mis*	会计错报	某公司发生会计错报为 1，否则为 0	
解释变量	*CA_CEO*	CEO 薪酬契约安排	$=\Delta Comp_CEO/(\Delta Earnings)$	+
	CA_CFO	CFO 薪酬契约安排	$=\Delta Comp_CFO/(\Delta Earnings)$	+
	DCA	CEO 与 CFO 薪酬契约安排的差异	$=CA_CEO-CA_CFO$	+
	DPower	CEO 与 CFO 权力的差异	$=Power_CEO-Power_CFO$	N/A
控制变量	*ROA*	总资产收益率	表示盈利能力	−
	Lev	资产负债率	表示偿债能力	+
	Ln*a*	总资产自然对数	衡量公司规模	−
	Agrow	成长情况	总资产增长率	+

① 没有考虑 CEO 的学历背景、政府层级等因素，因为缺省值较多。

② 本文没有对 CFO 生成所有制权力强度，这是由于不同于 CEO，CFO 无法达到对企业进行控制和操纵的能力。

续前表

变量类型	变量名称	经济含义	解释	预测符号
控制变量	*Big4*	会计师事务所规模	如果公司的财务报告经“四大”会计师事务所审计为 1，否则为 0	-
	AO	审计意见类型	如果公司被出具标准审计意见为 1，否则为 0	-
	MB	市值/账面价值比	Market-to-book ratio	-
	Σ *Indu dummy*	行业虚拟变量	当观测值为某一特定行业时取值 1，其他为 0，以控制行业影响	
	Σ *Year dummy*	年度虚拟变量	当观测值为某一特定年度时为 1，其他为 0，以控制年度影响	

四、实证检验与结果分析

（一）描述性统计

为更好地检验本文的假设，本文对连续变量在 1%的水平上进行了缩尾，并按照观测值是否发生错报分组进行了描述性统计分析，结果详见表 2。

在非错报组中（*Mis* = 0），CEO 薪酬契约安排变量（*CA_CEO*）的均值为 $0.728\ 0\times10^{-4}$，中位数为 0；而在错报组中（*Mis* = 1），CEO 薪酬契约安排变量（*CA_CEO*）的均值为 0.005 1，中位数为 0；*CA_CEO* 在两组间的均值差异为-0.005 1（*t* 值 = -3.458 1），表明错报组中 *CA_CEO* 的均值要显著高于非错报组中

表 2　　**描述性统计表**

变量	*Mis*	样本量	标准差	均值	中位数	最小值	最大值	均值差异
CA_CEO	0	12 507	0.022 9	$0.728\ 0\times10^{-4}$	0.000 0	-2.000 0	1.000 0	-0.005 1***
	1	697	0.132 1	0.005 1	0.000 0	0.000 0	3.000 0	(-3.458 1)
CA_CFO	0	12 507	0.008 0	$0.256\ 0\times10^{-4}$	0.000 0	0.000 0	0.000 0	-0.001 2***
	1	697	0.030 9	0.001 2	0.000 0	0.000 0	1.000 0	(-2.844 3)
DCA	0	12 507	0.023 5	$0.471\ 0\times10^{-4}$	0.000 0	-2.000 0	2.000 0	-0.003 9***
	1	697	0.101 4	0.004 0	0.000 0	0.000 0	3.000 0	(-3.070 1)
DPower	0	9 731	1.455 4	0.916 5	1.000 0	-13.000 0	14.000 0	0.152 4**
	1	534	1.435 5	0.764 0	1.000 0	-8.000 0	9.000 0	(2.357 7)
ROA	0	12 507	0.061 3	0.039 1	0.037 0	-0.349 5	0.210 0	0.025 6***
	1	697	0.085 7	0.013 5	0.022 9	-0.349 5	0.210 0	(10.468 6)
Lev	0	12 507	0.235 4	0.466 7	0.464 1	0.049 5	1.644 4	-0.078 6***
	1	697	0.280 6	0.545 3	0.549 2	0.049 5	1.644 4	(-8.486 9)
Ln*a*	0	12 507	1.365 2	21.813 6	21.631 3	14.941 6	30.449 1	0.392 3***
	1	697	1.182 2	21.421 3	21.350 4	17.048 7	26.220 7	(7.433 2)
Agrow	0	12 507	0.380 7	0.191 9	0.109 9	-0.438 1	2.889 0	0.028 6*
	1	697	0.451 9	0.163 3	0.066 1	-0.438 1	2.889 0	(1.906 5)
MB	0	12 507	1.731 5	1.985 0	1.488 2	0.189 5	10.289 0	0.039 1
	1	697	1.742 1	1.945 9	1.464 2	0.189 5	10.289 0	(0.579 7)
Big4	0	12 507	0.251 4	0.067 8	0.000 0	0.000 0	1.000 0	0.049 2***
	1	697	0.135 4	0.018 7	0.000 0	0.000 0	1.000 0	(5.119 9)
AO	0	12 507	0.186 6	0.963 9	1.000 0	0.000 0	1.000 0	0.088 7***
	1	697	0.330 8	0.875 2	1.000 0	0.000 0	1.000 0	(11.572 8)

注：括号内表示 t 值；***、**、*分别表示在 1%、5%、10%的水平上显著（双尾检验）。

CA_CEO 的均值。在非错报组中（*Mis* = 0），CFO 薪酬契约安排变量（*CA_CFO*）的均值为 0.256 0×10^{-4}，中位数为 0；而在错报组中（*Mis* = 1），CFO 薪酬契约安排变量（*CA_CFO*）的均值为 0.001 2，中位数为 0；*CA_CFO* 在两组间的均值差异为 -0.001 2（*t* 值 = -2.844 3），表明错报组中 *CA_CFO* 的均值要显著高于非错报组中 *CA_CFO* 的均值。在非错报组中（*Mis* = 0），CEO 与 CFO 薪酬契约安排差异变量（*DCA*）的均值和中位数分别为 0.471 0×10^{-4}和 0；而在错报组中（*Mis* = 1），薪酬契约安排差异变量（*DCA*）的均值和中位数分别为 0.004 0 和 0；*DCA* 在两组间的均值差异为-0.003 9（*t* 值 = -3.070 1），表明错报组中 *DCA* 的均值要显著高于非错报组中 *DCA* 的均值。在非错报组中（*Mis* = 0），CEO 与 CFO 权力差异变量（*DPower*）的均值和中位数分别为 0.916 5 和 1；而在错报组中（*Mis* = 1），权力差异变量（*DPower*）的均值和中位数分别为 0.764 0 和 1；*DPower* 在两组间的均值差异为 0.152 4（*t* 值 = 2.357 7），表明错报组中 *DPower* 的均值要显著低于非错报组中 *DPower* 的均值。控制变量方面，在非错报组中（*Mis* = 0），公司业绩变量（*ROA*）的均值和中位数为 0.039 1 与 0.037 0；而在错报组中（*Mis* = 1），公司业绩变量（*ROA*）的均值和中位数分别为 0.013 5 与 0.022 9；*ROA* 在两组间的均值差异为 0.025 6（*t* 值 = 10.468 6），表明错报组中 *ROA* 的均值要显著低于非错报组中 *ROA* 的均值。类似地，非错报组中资产负债率变量（*Lev*）要显著低于错报组，而公司规模（Ln*a*）、公司成长情况（*Agrow*）、由四大会计师事务所审计（*Big*4）以及被出具标准审计意见（*AO*）在非错报组中的均值显著高于在错报组中的均值。

（二）Probit 回归

本文首先针对假设 H1 与 H2，检验 CEO 与 CFO 薪酬契约安排对公司发生会计错报的影响，具体结果反映在表 3 中。列（1）与列（2）检验了 CEO 薪酬契约安排对会计错报的影响。在列（1）中，*CA_CEO* 的回归系数为 0.841 5（*z* 值 = 4.396 6），在 1% 的水平上显著，表明 CEO 薪酬契约安排正向影响公司发生会计错报的概率，假设 H1 得到了验证。类似地，在列（2）中，考虑了其他控制变量的影响，*CA_CEO* 的回归系数为 0.857 6（*z* 值 = 4.607 4），在 1%的水平上显著，表明在考虑相关控制变量的情况下，CEO 薪酬契约安排正向影响公司发生会计错报的概率，假设 H1 再次得到了验证。列（3）与列（4）检验了 CFO 薪酬契约安排对会计错报的影响。在列（3）中，*CA_CFO* 的回归系数为 2.558 3（*z* 值 = 2.483 1），在 5%的水平上显著，表明 CFO 薪酬契约安排正向影响公司发生会计错报的概率，假设 H2 得到了验证。类似地，在列（4）中，考虑了其他控制变量的影响，*CA_CFO* 的回归系数为 2.761 4（*z* 值 = 2.926 3），在 1%的水平上显著，表明在考虑相关控制变量的情况下，CFO 薪酬契约安排正向影响公司发生会计错报的概率，假设 H2 再次得到了验证。列（5）与列（6）将 CEO 与 CFO 薪酬契约安排放在一起进行回归。*CA_CEO* 的回归系数在列（5）与列（6）中均在 1% 的水平上显著为正，而 *CA_CFO* 的回归系数尽管为正，但不显著。整体上来看，CEO 与 CFO 薪酬契约安排对会计错报的正向影响得到了证实。进一步通过卡方检验（*chi*2 = 0.02）发现，在列（6）中，*CA_CEO* 与 *CA_CFO* 的回归系数不存在显著差异，未发现 CEO 与 CFO 薪酬契约安排对会计错报的影响大小孰高孰低。另外，在控制变量方面，盈利能力（*ROA*）、公司规模（Ln*a*）、账面市值比（*MB*）、会计师事务所规模（*Big*4）以及标准审计意见（*AO*）负向地影响公司发生会计错报的概率。

表 3　　CEO 与 CFO 薪酬契约安排对会计错报影响的回归结果

	（1） *Mis*	（2） *Mis*	（3） *Mis*	（4） *Mis*	（5） *Mis*	（6） *Mis*
CA_CEO	0.841 5*** （4.396 6）	0.857 6*** （4.607 4）			0.781 5*** （3.118 4）	0.774 2*** （3.322 0）
CA_CFO			2.558 3** （2.483 1）	2.761 4*** （2.926 3）	0.393 7 （0.386 3）	0.591 7 （0.588 3）
ROA		-1.281 0*** （-3.692 0）		-1.279 8*** （-3.688 4）		-1.280 9*** （-3.691 8）

续前表

	(1) *Mis*	(2) *Mis*	(3) *Mis*	(4) *Mis*	(5) *Mis*	(6) *Mis*
Lev		0.150 7 (1.443 7)		0.150 2 (1.437 8)		0.150 4 (1.440 3)
Ln*a*		−0.044 3* (−1.874 4)		−0.044 2* (−1.870 3)		−0.044 3* (−1.873 2)
Agrow		0.048 3 (0.834 4)		0.048 2 (0.832 6)		0.048 2 (0.834 0)
MB		−0.034 4** (−2.037 2)		−0.034 2** (−2.031 7)		−0.034 3** (−2.034 2)
*Big*4		−0.459 7*** (−2.660 3)		−0.462 7*** (−2.670 6)		−0.460 4*** (−2.662 5)
AO		−0.445 9*** (−4.541 7)		−0.447 0*** (−4.551 9)		−0.446 2*** (−4.544 3)
Σ *Indu dummy*	已控制	已控制	已控制	已控制	已控制	已控制
Σ *Year dummy*	已控制	已控制	已控制	已控制	已控制	已控制
常数项	−9.366 8*** (−20.277 8)	−8.371 1*** (−13.960 0)	−9.366 1*** (−37.896 0)	−8.372 8*** (−13.798 1)	−9.366 7*** (−86.313 1)	−8.371 7*** (−34.060 5)
CA_CEO = CA_CFO						chi^2 = 0.02
N	13 204	13 204	13 204	13 204	13 204	13 204
Pseudo R^2	0.090 3	0.116 5	0.090 1	0.116 3	0.090 4	0.116 5

注：对标准误按照公司代码和年度进行了聚类；括号内表示z值；***、**、*分别表示在1%，5%以及10%的水平上显著（双尾检验）。下同。

本文针对假设H3，检验CEO与CFO薪酬契约安排差异对公司发生会计错报的影响，具体结果反映在表4中。列（1）与列（2）检验了薪酬契约安排差异对会计错报的影响。在列（1）中，*DCA*的回归系数为0.882 9（*z*值=2.911 1），在1%的水平上显著，表明CEO与CFO薪酬契约安排差异正向影响公司发生会计错报的概率，假设H3得到了验证。类似地，在列（2）中，考虑了其他控制变量的影响，*DCA*的回归系数为0.882 9（*z*值=3.022 4），在1%的水平上显著，表明在考虑相关控制变量的情况下，CEO与CFO薪酬契约安排差异正向影响公司发生会计错报的概率，假设H3再次得到了验证。另外，在控制变量方面，盈利能力（*ROA*）、公司规模（Ln*a*）、账面市值比（*MB*）、会计师事务所规模（*Big*4）以及标准审计意见（*AO*）负向地影响公司发生会计错报的概率。

表4　　CEO与CFO薪酬契约安排差异对会计错报影响的回归结果

	(1) *Mis*	(2) *Mis*
DCA	0.882 9*** (2.911 1)	0.882 9*** (3.022 4)
ROA		−1.280 6*** (−3.691 2)
Lev		0.151 5 (1.450 8)
Ln*a*		−0.044 4* (−1.878 7)
Agrow	0.048 3 (0.834 4)	

续前表

	(1) *Mis*	(2) *Mis*
MB		−0.034 5** (−2.044 1)
*Big*4		−0.458 8*** (−2.655 1)
AO		−0.445 5*** (−4.537 7)
Σ *Indu dummy*	已控制	已控制
Σ *Year dummy*	已控制	已控制
常数项	−9.366 9*** (−12.904 3)	−8.368 8*** (−12.904 4)
N	13 204	13 204
Pseudo R^2	0.090 2	0.116 4

进一步地，本文还要检验 CEO 与 CFO 薪酬契约安排差异对会计错报的影响是否会受到 CEO 与 CFO 权力差异的调节。表 5 列示了薪酬契约安排差异、权力差异对会计错报影响的 Probit 回归结果。列（1）与列（2）检验了薪酬契约安排差异、权力差异对会计错报的影响。在列（1）中，*DCA* 的回归系数为 0.607 0（*z* 值 = 2.576 5），在 1% 的水平上显著，表明 CEO 与 CFO 薪酬契约安排差异正向影响公司发生会计错报的概率；而 *DCA* 与 *DPower* 的交互项（*DCA*×*DPower*）的回归系数为 −0.599 5（*z* 值 = −2.511 7），在 5% 的水平上显著，表明 CEO 与 CFO 权力差异负向调节薪酬契约安排差异对会计错报的影响，即弱化了薪酬契约安排对会计错报的正向影响，假设 H4b 得到了验证。类似地，在列（2）中，考虑了其他控制变量的影响，*DCA* 的回归系数为 0.625 1（*z* 值 = 2.742 2），在 1% 的水平上显著。而 *DCA* 与 *DPower* 的交互项（*DCA*×*DPower*）的回归系数为 −0.585 9（*z* 值 = −2.483 1），在 5% 的水平上显著。假设 H4b 再次得到了验证。另外，在控制变量方面，盈利能力（*ROA*）、账面市值比（*MB*）、会计师事务所规模（*Big*4）以及标准审计意见（*AO*）负向地影响公司发生会计错报的概率。

表 5　　CEO 与 CFO 权力差异的调节作用

	(1) *Mis*	(2) *Mis*
DCA	0.607 0*** (2.576 5)	0.625 1*** (2.742 2)
DCA×*DPower*	−0.599 5** (−2.511 7)	−0.585 9** (−2.483 1)
DPower	−0.022 2 (−1.322 5)	−0.018 2 (−1.090 1)
ROA		−0.834 5** (−1.991 3)
Lev		0.198 6 (1.613 9)
Ln*a*		−0.033 6 (−1.240 3)

续前表

	(1) *Mis*	(2) *Mis*
Agrow		0.062 6 (0.886 0)
MB		−0.035 9* (−1.903 0)
*Big*4		−0.395 6* (−1.953 0)
AO		−0.521 7*** (−4.435 0)
Σ *Indu dummy*	已控制	已控制
Σ *Year dummy*	已控制	已控制
常数项	−9.330 0*** (−3.441 8)	−5.492 8*** (−8.796 0)
N	10 265	10 265
Pseudo R^2	0.096 5	0.118 3

(三) 进一步分析

1. 对舞弊的考虑。

会计错报通常由两种情形引起，一种是错误，而另一种为舞弊，是一种管理层的机会主义行为。如果薪酬契约安排是造成CEO、CFO操纵报表的薪酬动机，那么在研究薪酬契约安排差异对会计错报影响时，也应当进一步考虑会计错报究竟是由舞弊还是差错引起的。舞弊引起的错报是一种蓄意的机会主义行为，同薪酬动机相关。借鉴马晨等(2015)[7]、Ma等(2015)[26]的做法，本文使用关键词查找会计信息，如“虚增”“虚减”等；另外，如果监管部门介入，如财政机关、证监会、审计署等认为公司会计信息披露不实，要求上市公司整改。符合以上两个条件之一的情形即被认定为由财务舞弊导致的会计错报。在此基础上生成因变量 *Misfraud*，公司未发生会计错报时 *Misfraud* 取值为0，发生由差错导致的会计错报时取值为1，发生由舞弊导致的会计错报时取值为2。*Misfraud* 的取值范围在［0，2］且为离散变量。因此，针对因变量 *Misfraud*，本文进行了 Ordered Probit 回归，表6列示了回归结果。

表6　　CEO与CFO薪酬契约安排对财务舞弊影响的回归结果

	(1) *Misfraud*	(2) *Misfraud*	(3) *Misfraud*	(4) *Misfraud*	(5) *Misfraud*	(6) *Misfraud*
CA_CEO	0.985 4*** (4.761 8)	1.006 5*** (5.011 0)			0.903 0*** (3.455 0)	0.898 5*** (3.713 4)
CA_CFO			3.097 4*** (2.762 8)	3.324 0*** (3.277 5)	0.523 6 (0.521 0)	0.727 9 (0.757 4)
ROA		−1.286 7*** (−3.804 3)		−1.284 8*** (−3.798 8)		−1.286 6*** (−3.803 9)
Lev		0.130 1 (1.273 9)		0.129 4 (1.265 9)		0.129 6 (1.268 9)
Ln*a*		−0.032 9 (−1.411 9)		−0.032 8 (−1.406 6)		−0.032 9 (−1.410 4)
Agrow		0.050 2 (0.845 2)		0.050 1 (0.843 1)		0.050 2 (0.844 7)

续前表

	(1) *Misfraud*	(2) *Misfraud*	(3) *Misfraud*	(4) *Misfraud*	(5) *Misfraud*	(6) *Misfraud*
MB		−0.030 5* (−1.843 3)		−0.030 3* (−1.836 9)		−0.030 4* (−1.839 4)
*Big*4		−0.452 9** (−2.412 0)		−0.456 9** (−2.424 4)		−0.453 8** (−2.414 8)
AO		−0.414 3*** (−4.297 5)		−0.415 8*** (−4.311 9)		−0.414 7*** (−4.301 1)
Σ *Indu dummy*	已控制	已控制	已控制	已控制	已控制	已控制
Σ *Year dummy*	已控制	已控制	已控制	已控制	已控制	已控制
CA_CEO=*CA_CFO*						chi^2=0.02
*cut*1						
常数项	9.365 2*** (36.036 7)	8.601 3*** (13.441 6)	9.364 6*** (36.032 0)	8.603 3*** (16.724 8)	9.365 1*** (36.036 7)	8.602 0*** (31.396 4)
*cut*2						
常数项	10.033 3*** (39.498 3)	9.284 0*** (14.387 9)	10.032 3*** (39.493 7)	9.285 7*** (18.103 2)	10.033 3*** (39.498 2)	9.284 8*** (33.130 6)
N	13 204	13 204	13 204	13 204	13 204	13 204
Pseudo R^2	0.078 9	0.100 0	0.078 6	0.099 8	0.078 9	0.100 1

表 6 针对假设 H1 与 H2，检验了 CEO 与 CFO 薪酬契约安排对财务舞弊的影响。其中，列（1）与列（2）检验了 CEO 薪酬契约安排对财务舞弊的影响。*CA_CEO* 的回归系数在列（1）与列（2）中均显著为正，且在 1%的水平上显著，表明随着 CEO 薪酬契约安排的增加，公司发生的会计错报更加可能是由于蓄意的财务舞弊而不是疏忽导致。表 6 中的列（3）与列（4）检验了 CFO 薪酬契约安排对财务舞弊的作用。*CA_CFO* 的回归系数在列（3）与列（4）中均显著为正，且在 1%的水平上显著，表明随着 CFO 薪酬契约安排的增加，公司发生的会计错报更加可能是由于蓄意的财务舞弊而不是疏忽导致。列（5）与列（6）将 CEO 与 CFO 薪酬契约安排放在一起进行回归。*CA_CEO* 的回归系数在列（5）与列（6）中均在 1%的水平上显著为正，而 *CA_CFO* 的回归系数尽管为正，但不显著。从整体上来看，CEO 与 CFO 薪酬契约安排对财务舞弊具有正向影响。进一步通过卡方检验（*chi*2=0.02）发现，在列（6）中，*CA_CEO* 与 *CA_CFO* 的回归系数不存在显著差异，未发现 CEO 与 CFO 薪酬契约安排对财务舞弊的影响作用孰大孰小。这些结果说明，CEO 与 CFO 薪酬契约安排对会计错报的影响不因错报原因而发生改变。

本文还对 CEO 与 CFO 薪酬契约安排差异、权力差异对财务舞弊的影响进行了 Ordered Probit 回归分析。表 7 中的列（1）与列（2）检验了 CEO 与 CFO 薪酬契约安排差异对财务舞弊的影响。*DCA* 的回归系数在列（1）与列（2）中均显著为正，且在 1%的水平上显著，表明随着 CEO 与 CFO 薪酬契约安排差异的增加，公司发生的会计错报更加可能是由于蓄意的财务舞弊而不是疏忽导致。表 7 中的列（3）与列（4）检验了 CEO 与 CFO 权力差异对薪酬契约安排差异与财务舞弊之间关系的调节作用。*DCA* 的回归系数在列（3）与列（4）中均显著为正，且在 1%的水平上显著；而 *DCA* 与 *DPower* 的交互项（*DCA*×*DPower*）的回归系数在 1%的水平上显著为负，表明 CEO 与 CFO 权力差异会弱化其薪酬契约安排差异与财务舞弊之间的关系。这说明，CEO 与 CFO 薪酬契约安排差异、权力差异对会计错报的影响不因错报原因而发生改变。

表7　　CEO与CFO权力差异的调节作用（考虑舞弊）

	(1) *Misfraud*	(2) *Misfraud*	(3) *Misfraud*	(4) *Misfraud*
DCA	1.037 9*** (3.031 6)	1.044 2*** (3.151 2)	0.681 8*** (2.770 7)	0.699 9*** (2.948 0)
DCA×*DPower*			−0.738 1*** (−3.014 9)	−0.732 7*** (−3.046 7)
DPower			−0.022 2 (−1.400 5)	−0.018 5 (−1.174 1)
ROA		−1.286 2*** (−3.803 2)		−0.809 4** (−1.996 1)
Lev		0.131 1 (1.283 9)		0.184 0 (1.532 0)
Ln*a*		−0.033 0 (−1.417 4)		−0.022 8 (−0.854 7)
Agrow		0.050 2 (0.845 2)		0.074 3 (1.020 1)
MB		−0.030 6* (−1.852 3)		−0.030 2 (−1.640 4)
*Big*4		−0.451 7** (−2.406 2)		−0.378 9* (−1.711 8)
AO		−0.413 7*** (−4.291 7)		−0.497 9*** (−4.325 9)
Σ *Indu dummy*	已控制	已控制	已控制	已控制
Σ *Year dummy*	已控制	已控制	已控制	已控制
*cut*1				
常数项	9.365 3*** (24.378 6)	8.598 5*** (25.448 4)	9.380 4*** (8.679 5)	5.759 8*** (9.361 3)
*cut*2				
常数项	10.033 2*** (24.809 2)	9.281 0*** (26.514 6)	10.063 9*** (9.582 1)	6.455 2*** (10.547 0)
N	13 204	13 204	10 265	10 265
Pseudo R^2	0.078 8	0.099 9	0.084 0	0.101 6

2. 对非错报样本的重新考虑。

按照马晨等（2014）[23]的做法，由于经历过错报的公司在其他未发生错报的年度可能在会计信息的生成方面存在缺陷，本文对这部分观测值进行了剔除。针对697个错报观测值以及8 001个非错报观测值再次进行了回归，结果如表8所示。

表 8　　CEO 与 CFO 权力差异的调节作用（对非错报样本的重新考虑）

	(1) *Misfraud*	(2) *Misfraud*	(3) *Misfraud*	(4) *Misfraud*	(5) *Misfraud*
CA_CEO	0.791 3*** (4.192 4)		0.723 1*** (3.212 9)		
CA_CFO		2.888 3*** (2.733 5)	0.682 0 (0.551 6)		
DCA				0.803 8*** (2.900 9)	0.632 1*** (2.666 7)
DCA×DPower					−0.451 2* (−1.927 5)
DPower					−0.020 7 (−1.105 0)
ROA	−1.948 0*** (−4.445 8)	−1.947 8*** (−4.445 7)	−1.948 2*** (−4.446 4)	−1.947 2*** (−4.443 8)	−1.462 4*** (−2.789 2)
Lev	0.549 1*** (4.080 4)	0.547 5*** (4.068 2)	0.548 5*** (4.075 3)	0.550 2*** (4.088 8)	0.595 5*** (3.754 0)
Ln*a*	−0.062 5** (−2.396 0)	−0.062 5** (−2.393 8)	−0.062 5** (−2.395 9)	−0.062 6** (−2.398 3)	−0.049 7* (−1.648 5)
Agrow	0.053 0 (0.759 6)	0.053 0 (0.759 3)	0.053 0 (0.759 3)	0.053 0 (0.759 3)	0.069 9 (0.867 1)
MB	−0.029 0 (−1.437 4)	−0.029 0 (−1.434 8)	−0.029 0 (−1.435 3)	−0.029 1 (−1.442 0)	−0.027 2 (−1.220 6)
*Big*4	−0.611 4*** (−3.214 9)	−0.611 4*** (−3.215 6)	−0.611 4*** (−3.215 1)	−0.611 3*** (−3.214 4)	−0.550 7** (−2.438 8)
AO	−0.569 5*** (−4.884 7)	−0.571 3*** (−4.901 0)	−0.569 8*** (−4.887 9)	−0.569 0*** (−4.880 7)	−0.667 9*** (−4.873 3)
Σ *Indu dummy*	已控制	已控制	已控制	已控制	已控制
Σ *Year dummy*	已控制	已控制	已控制	已控制	已控制
常数项	−7.697 3*** (−14.979 3)	−7.692 3*** (−18.667 7)	−7.696 0*** (−19.718 8)	−7.697 9*** (−15.389 8)	−5.147 1*** (−7.681 7)
N	8 698	8 698	8 698	8 698	6 854
Pseudo R^2	0.168 0	0.167 8	0.168 0	0.167 9	0.167 8

表 8 中的列（1）至列（3）分别检验了假设 H1 与假设 H2，即 CEO 与 CFO 薪酬契约安排对会计错报的影响。*CA_CEO* 与 *CA_CFO* 的回归系数均同之前保持一致，即正向影响公司的会计错报，表明本文的研究结论不受非错报样本选择方式的影响。针对假设 H3 与假设 H4，*DCA* 的回归系数在列（4）与列（5）中均在 1%的水平上显著为正，表明 CEO 与 CFO 薪酬契约安排差异正向影响公司发生会计错报的概率，假设 H1 再次得到验证；而 *DCA* 与 *DPower* 的交互项（*DCA×DPower*）的回归系数在 10%的水平上显著为负，表明 CEO 与 CFO 权力差异弱化 CEO 与 CFO 薪酬契约安排差异与会计错报之间的关系，假设 H4b

再次得到验证。

3. 配对样本分析。

为使得本研究结论更为可靠，本文借鉴马晨和张俊瑞（2012）[27]的做法，采用配对样本方法，按照年度、上市地、行业以及总资产规模（正负20%以内）四项标准进行分析。在此基础上，本文重新进行了回归分析，结果反映在表9中。

表9　　CEO与CFO薪酬契约安排及其差异对会计错报影响的回归结果

	(1) *Misfraud*	(2) *Misfraud*	(3) *Misfraud*	(4) *Misfraud*	(5) *Misfraud*	(6) *Misfraud*	(7) *Misfraud*	(8) *Misfraud*
CA_CEO	461.612***	543.220***			566.810***	640.125***		
	(4.040)	(4.605)			(4.839)	(5.046)		
CA_CFO			25.961***	25.066**	−174.810***	−157.277***		
			(2.716)	(2.533)	(−3.121)	(−2.893)		
DCA							234.064***	228.936***
							(4.357)	(4.262)
ROA		−1.423**		−1.321*		−1.403**		−1.326*
		(−2.003)		(−1.878)		(−1.979)		(−1.883)
Lev		0.110		0.105		0.113		0.101
		(0.516)		(0.498)		(0.530)		(0.476)
Ln*a*		−0.009		−0.019		−0.014		−0.019
		(−0.174)		(−0.376)		(−0.261)		(−0.369)
Agrow		0.143		0.133		0.148		0.141
		(1.088)		(1.016)		(1.126)		(1.072)
MB		−0.112***		−0.107***		−0.112***		−0.107***
		(−3.043)		(−2.921)		(−3.061)		(−2.938)
Big4		−0.918***		−0.916***		−0.893***		−0.886***
		(−3.083)		(−3.100)		(−3.007)		(−2.986)
AO		−0.436**		−0.410**		−0.427**		−0.408**
		(−2.555)		(−2.408)		(−2.503)		(−2.401)
Σ *Indu dummy*	已控制	已控制	已控制	已控制	已控制	已控制	已控制	已控制
Σ *Year dummy*	已控制	已控制	已控制	已控制	已控制	已控制	已控制	已控制
常数项	−0.578	0.131	−0.579	0.331	−0.554	0.244	−0.541	0.364
	(−1.305)	(0.104)	(−1.310)	(0.265)	(−1.247)	(0.193)	(−1.220)	(0.290)
N	910	910	910	910	910	910	910	910
Pseudo R^2	0.034	0.074	0.016	0.053	0.039	0.078	0.029	0.065

表9中的列（1）至列（6）关注CEO与CFO薪酬契约安排对会计错报的影响。从中可以看出，*CA_CEO*与*CA_CFO*的回归结果均显著为正，同之前的结果保持一致。再次表明CEO与CFO薪酬契约安排均正向影响公司发生会计错报的概率。列（7）和列（8）关注CEO与CFO薪酬契约安排差异对会计错报的影响。从中可以看出，*DCA*的回归系数均显著为正，同之前结果一致。再次表明CEO与CFO薪酬契约安排差异正向影响公司发生会计错报的概率。

五、研究结论

CEO与CFO均要对财务报告的真实性和完整性

负责。已有研究更多关注 CEO 薪酬动机对财务报告质量的影响，本文认为单纯关注 CEO 而忽视 CFO 的影响所得出的结论可能是有偏的。薪酬动机均会使得 CEO 与 CFO 有意愿去操纵报表以实现个人利益，薪酬动机还会使得 CFO 更努力履行自身作为财务监督人的职责，因此，有必要将 CEO 与 CFO 薪酬契约安排动机结合起来，即研究 CEO 与 CFO 薪酬契约安排动机差异对财务报告质量的影响。薪酬契约安排动机发挥作用的程度还离不开个人能力，因此本文对 CEO 与 CFO 权力差异也进行了考虑。

本文通过手工收集 2005—2014 年间上市公司发生的会计错报数据，用薪酬契约安排作为薪酬动机的代理变量，检验了 CEO、CFO 薪酬契约安排对会计错报的影响，并考虑这种影响是否会受到 CEO 与 CFO 权力差异的调节。本文发现：CEO 与 CFO 薪酬契约安排均会正向影响公司会计错报；CEO 与 CFO 薪酬契约安排差异对公司发生会计错报具有正向的影响，表明随着 CEO 与 CFO 薪酬契约安排差异的增加，公司发生会计错报的概率随之上升。进一步地，本文还发现 CEO 与 CFO 薪酬契约安排差异对会计错报的正向影响会受到 CEO 与 CFO 权力差异的负向调节，表明 CEO 权力越大于 CFO，CEO 越能够通过其他渠道实现机会主义行为，而不仅仅限于会计操纵，在这种情形下，CEO 与 CFO 薪酬契约安排差异对会计错报的正向影响作用反而会被弱化。此外，本文在进一步分析中还发现，随着 CEO 与 CFO 薪酬契约安排差异的增加，不但公司发生会计差错的概率会提升，在发生的差错中更加可能是由舞弊而不是人为疏忽引起的。

本文的研究结论为上市公司的公司治理提供以下建议：第一，适当降低 CEO 与 CFO 之间薪酬契约安排的差异，以避免 CFO 潜在的监督意愿的下降，这在一定程度上能降低会计错报发生的概率；第二，CEO 与 CFO 权力配置要适度，随着 CEO 权力与 CFO 权力差异的增加，虽然能够在一定程度上弱化 CEO 与 CFO 薪酬契约安排差异对会计错报的正向影响，但可能也是 CEO 能够通过其他渠道而不仅局限于会计操纵以实现一己私利的表现。

参考文献

[1] Burns N，Kedia S. The Impact of Performance-based Compensation on Misreporting [J]. Journal of Financial Economics，2006，79 (1)：35-67.

[2] Dow J，Raposo C C. CEO Compensation，Change，and Corporate Strategy [J]. The Journal of Finance，2005，60 (6)：2701-2727.

[3] Jiang J X，Petroni K R，Wang I Y. CFOs and CEOs：Who Have the Most Influence on Earnings Management? [J]. Journal of Financial Economics，2010，96 (3)：513-526.

[4] Chava S，Purnanandam A. CEOs Versus CFOs：Incentives and Corporate Policies [J]. Journal of Financial Economics，2010，97 (2)：263-278.

[5] Feng M，Ge W，Luo S，et al. Why Do CFOs Become Involved in Material Accounting Manipulations? [J]. Journal of Accounting and Economics，2011，51 (1)：21-36.

[6] 林大庞，苏冬蔚. CEO 与 CFO 股权激励的治理效应之比较：基于盈余管理的实证研究 [J]. 南方经济，2012 (6)：15-31.

[7] 马晨，程茂勇，张俊瑞. 外部审计、媒介环境对财务重述的影响研究 [J]. 管理工程学报，2015 (4)：65-75.

[8] 马晨，张俊瑞. 管理层持股，领导权结构与财务重述 [J]. 南开管理评论，2012 (2)：143-150.

[9] 王福胜，程富. 管理防御视角下的 CFO 背景特征与会计政策选择——来自资产减值计提的经验证据 [J]. 会计研究，2014 (12)：32-38，95.

[10] 向锐. CFO 财务执行力与企业过度投资——基于董事会视角的分析 [J]. 会计研究，2015 (7)：56-62，97.

[11] 邱昱芳，贾宁，吴少凡. 财务负责人的专业能力影响公司的会计信息质量吗？——基于中国上市公司财务负责人专项调查的实证研究 [J]. 会计研究，2011 (4)：61-67.

[12] 刘永丽. 财务总监权力对会计稳健性影响的实证研究 [J]. 中国软科学，2015 (4)：121-130.

[13] Ge W，Matsumoto D，Zhang J L. Do CFOs Have Style? An Empirical Investigation of the Effect of Individual CFOs on Accounting Practices [J]. Contemporary Accounting Research，2011，28 (4)：1141-1179.

[14] 孙光国，郭睿. CFO 内部董事有助于董事会履行监督职能吗？[J]. 会计研究，2015 (11)：27-33，96.

[15] Indjejikian R，Matějka M. CFO Fiduciary Responsibilities and Annual Bonus Incentives [J]. Journal of Accounting Research，2009，47 (4)：1061-1093.

[16] Jensen M C，Murphy K J. Performance Pay and Top-management Incentives [J]. Journal of Political Economy，1990，98 (2)：225-264.

(下转第 114 页)

委托方女性CFO更认可“保守”的评估结论吗？

Do the Entrusting Party's Female CFOs More Approve "Conservative" Appraisal Conclusion?

赵　毅　　张双鹏

ZHAO Yi　　ZHANG Shuang-peng

［摘　要］基于当前大力加强资产评估行业理论创新和专业创新的背景，本文深入考察委托方CFO性别差异对资产评估机构出具的评估报告结论（增值率）的价值判断影响。研究发现，委托方女性CFO所认可的被评估资产增值率显著低于男性CFO，其资产增值率平均比男性CFO至少低62%以上。进一步研究发现，委托方CFO性别差异对国有企业的被评估资产价值增值率的降低效应低于非国有企业，在相对低声誉的资产评估机构样本中委托方CFO对被评估资产价值的专业判断的性别效应能够发挥得更加明显；并且，委托方企业风险承担水平是CFO性别差异影响被评估资产价值增值率的稳定且重要的路径；另外，研究还表明，与董事长或首席执行官不同，CFO是委托方资产评估业务活动的真正执行者，会直接参与到具体的评估实践过程中。本文的研究不仅从资产评估视角拓展了CFO性别差异下的经济行为研究，还为评估机构探索影响资产价值评估的重要因素，完善现有的评估技术提供了有益的启示。

［关键词］资产评估　评估增值率　委托方CFO　性别效应

［中图分类号］F239　［文献标识码］A　［文章编号］1000-1549（2020）10-0070-18

Abstract: Based on the background of strengthening innovation between academics and practitioners of asset appraisal industry, this paper makes an in-depth study of the impact of gender difference of entrusting party's CFOs on the value judgment of the conclusion (value-added rate) of the appraisal report issued by asset appraisal intermediaries. The results show that the value-added rate of the assessed assets approved by female CFOs is significantly lower than that of male CFOs, and the average value-added rate is at least 62% lower than their male counterparts. Further research shows that the gender difference of entrusting party's CFOs has a lower effect on the appreciation rate of assessed assets of state-owned enterprises than that of non-state-owned enterprises. In the sample of asset appraisal institutions with relatively low reputation, the gender effect of entrusting party's CFOs on the professional judgment of assessed asset value is more obvious. Moreover, the risk level of the entrusting enterprise is a stable and important path for the gender difference of CFO to affect the comprehensive value of the assessed assets. In addition, the study also shows that the existence of female chairmen and female CEOs does not significantly influence the findings of this paper except for female CFOs. This paper not only promotes the financial behavior research under the gender difference of CFO from a new perspective, but also provides useful inspiration for business guidance of China Appraisal Society (CAS) and the specific practice of asset appraisal intermediaries.

Key words: Asset appraisal　Assessment of value added rate　Entrusting party's CFOs　Gender effect

［收稿日期］　2019-03-06

［作者简介］　赵毅，男，1985年1月生，山西财经大学会计学院讲师，管理学博士，研究方向为技术并购估值、企业价值评估等；张双鹏，男，1984年11月生，山东工商学院会计学院副教授，上海国家会计学院博士后，管理学博士，研究方向为公司治理与企业战略、公司治理与企业价值。

［基金项目］　山西省“1331工程重点创新团队建设计划”（项目编号：晋教科［2017］12号）。

感谢匿名评审人提出的修改建议，笔者已做了相应修改，本文文责自负。

一、引言

资本市场建设对资产评估服务有着高端、直接的需求，其中对资产价值的准确判断是市场参与者进行资产交易的基础。如何提高对交易对象价值判断的准确性，进而合理地评价、运用资产评估结论作为定价参考的依据就成为评估实践中亟待解决的重要问题。但纵观前期研究发现，实务界侧重从推动资产评估立法、完善评估准则、修正评估方法、提升评估机构独立性以及加强行业监管等制度层面展开有益探索（北京证监局评估监管课题组，2017[1]；张国春，2016[2]；纪益成，2016[3]；余炳文和张君来，2016[4]），但对个体微观层面关键因素的研究则缺乏足够的关注。

事实上，在资产评估实践中，委托方负责人与评估机构随时处于信息沟通之中，评估机构实际上处于现有设计的财务信息框架之中，因而在既定准则和程序下，委托方作为信息提供和判断框架的设计者，特别是对生成和提供公司财务信息肩负重要职责的 CFO，就很可能依据其自身特征对企业的未来预期收益与风险做出判断和决策，进而对拟评估资产的价值判断（差异）产生一条符合经济直觉的影响路径。进一步地，高管性别作为管理者最为重要的个体特征对企业财务实践的影响已日益受到理论界的关注。学者们普遍认为，高管性别特征在公司财务与公司治理等方面均具有重要影响，而公司财务与治理又与评估资产价值密切相关（李小荣等，2016[5]），那么高管性别是否会影响其对于资产价值的评估判断呢？也即可表述为：由资产评估机构出具的评估报告中对被评估资产的价值判断是否也会受到委托方 CFO 性别差异的影响呢？遗憾的是，尽管大量经验研究揭示了高管性别对财务与会计实践的诸多方面影响，但是研究重点仍集中于从股东或投资者视角去进行 CFO 性别差异的经济效果检验。值得注意的是，作为促进资本市场发展中的重要中介机构——资产评估机构，从该视角的研究还未得到学者们应有的重视（马海涛等，2017[6]；李小荣等，2016[5]）。此外，在现有实务界对资产价值的评定估算实践过程中，其评估机构所采用的评估方法都尚未对企业高管特征对资产价值评估的影响予以足够考虑。例如，市场法在选取参照物时通常注重选择交易因素、时间、行业、企业经营绩效、规模、成长性等要素相同或相似的交易案例，但缺乏考虑高管特征是否相同或类似。又如采用收益法对交易资产进行价值评估时，其管理层的高管特征异质性又会影响预期收益额和风险，而企业风险又与折现率的确定直接相关。因此，本文旨在研究委托方高管（CFO）对评估机构出具的评估报告结论（增值率）的价值判断中的性别效应，以弥补现有理论与实务两界的缺憾，并为资产评估机构如何进一步探索影响资产价值评估的重要因素，完善现有的评估技术提供合理的建议。

本文的边际贡献可以归结如下：第一，在理论层面上，先前研究对女性高管财务行为的考察仍然以企业股东视角或投资者视角为研究重点，其主题主要集中于女性 CFO 对公司治理与绩效、过度投资行为、税收规避行为、财务舞弊行为、盈余管理、并购决策等方面的影响（Adams 和 Ferreira，2009[7]；何威风和刘启亮，2010[8]；Srinidhi 等，2011[9]；Ge 等 2011[10]；李小荣等，2013[11]；Francis 等，2014[12]；Francis 等，2015[13]；Huang 和 Kisgen，2013[14]），而鲜有从资产评估中介视角着手的研究。本文首次揭示了委托方 CFO 性别特质在被评估资产价值判断中的实际影响，这为全面理解性别差异在资产评估中所产生的经济效果提供了有益的启示。第二，在实务层面上，本文紧密联系资产评估实践中的不同情境，将资产评估学的实务理论与高管特征理论相结合，丰富了评估领域的实务知识体系，也为深刻理解评估场景围绕 CFO 性别差异下资产价值判断的交互影响提供了崭新且切实可行的视角。第三，在研究方法层面上，目前学界普遍采用的倾向值匹配方法并不能完全保证试验组和对照组在消除偏误和减少方差两方面同时实现最优化。本文在传统的 PSM 分析的基础上，分别引入两种全新的优化分组匹配效率的算法，并对研究结果的可靠性进行再测试，这为今后学者运用较为新颖、前沿的应对选择性偏误的处理方法进行因果推断提供了有价值的参考。

二、理论分析与研究假设

一系列心理学、管理学等相关研究表明，性别特质作为决策者性别差异的重要解释，存在系统性差异。相较于男性高管，女性高管在财务与会计实践中具体表现出的个体、心理、人格等特质呈现出与男性高管不同的倾向归因（Khlif 和 Achek，2017[15]；Francis 等，2015[13]；Huang 和 Kisgen，2013[14]；

Figner 和 Weber，2011[16]；Borghans 等，2009[17]；Niederle 和 Vesterlund，2007[18]）。进一步地，由高管性别差异所反映出的系统性特质差异又会引致相应的管理行为呈现出明显的分化，进而对组织输出结果也将造成差异的影响。由于 CFO 负责生成和提供公司的财务信息，所产生信息的质量取决于 CFO 的特征，因而本文将 CFO 性别差异的内在倾向特质作为一种重要的解释，去剖析 CFO 系统性特质差异对被评估资产价值判断的潜在作用机理。

就女性管理者的心理特质而言，多数文献都为女性具有更高的风险规避程度提供了理论分析和经验证据的支持（Khlif 和 Achek，2017[15]；Huang 和 Kisgen，2013[14]；李小荣和刘行，2012[19]；周泽将，2012[20]；Borghans 等，2009[17]；Agnew 等，2003[21]；Carter 等，2017[22]；Bernasek 和 Shwiff，2001[23]）。女性高管的这种天生的风险偏好将导致其谨慎、敏感、保守及非过度自信等行为倾向。在财务与会计实践中，通常女性 CFO 与男性 CFO 的决策和效果差异也能够由基于性别差异的个体风险偏好特质所解释（Olsen 和 Cox，2001[24]；Krishnan 和 Parsons，2008[25]；Francis 等，2014[12]；Francis 等，2015[13]；Hwang 等，2017[26]）。就评估实践活动来说，一方面，如采用收益法对单项或整体类型的资产进行评定、估算时，准确预测资产的未来预期收益和确定折现率是估值的最为关键的两个要素。而高管特征会同时影响预期收益额和企业风险（李小荣等，2016[5]），并且企业风险又与折现率的确定直接相关。女性 CFO 所具有的更高风险规避程度使得其在财务决策中更加倾向于降低风险（Hwang 等，2017[26]），因而在对期望投资报酬率的确定上会与男性 CFO 存在系统性差异。此外，学者 Carter 等（2017）[22]发现，较低的风险厌恶被认为与过度自信相关，而高管的自信程度会对企业的财务决策产生重要影响，过度自信的管理者通常高估自身能力与决策收益，低估风险（姜付秀等，2009[27]），而女性 CFO 所具有的更高风险规避程度以及所体现出的相对不过度自信将影响其在财务行为、决策的选择上更为谨慎，风险承受水平也更低，这进而也会影响未来预期收益的预测和期望投资报酬率的确定。因而，在面对过高的资产增值率的情形时，女性 CFO 所具有的更高风险规避特质使其更倾向于从避免增大企业风险和谨慎决策角度考虑，接受认可相对较低或波动性较小的资产增值率结论。另一方面，按照资产评估程序要求，评估机构在评估报告内部审核之后，必须与委托人或相关当事人对评估报告的内容进行沟通，以使评估机构了解委托人或者相关当事人对评估结论的反馈意见。在这一需要委托人合理地理解评估结论，正确使用评估报告的关键环节中，通常而言，女性 CFO 相较于男性 CFO 更加谨慎、保守以及风险规避的特质要求对其财务信息负责，并且女性 CFO 通常希望以较小的法律与监管风险来实现职业安全风险的降低，从而在法律与监管风险、职业风险之间取得良好的平衡。这实际上表现出一种整体风险最小化行为（Francis 等，2015[13]），将使得其与评估机构交换意见时更倾向于选择相对稳健和可靠的价值评估信息，其对评估报告的内容也会表现出更高程度的谨慎性，对过高资产增值率的评估结论所可能引发的企业风险可能更为敏感，对风险的承受意愿也更低。就女性管理者的个体特质而言，女性高管可以利用其独特的职业沟通优势、敏锐的财经触觉以及法律规则赋予的权利来影响企业财务行为。已有经济学、心理学及管理学等研究表明，与男性相比，女性 CFO 更具有独特的沟通优势，并能够改善在沟通中可能存在的信息不对称等问题，进而有助于提供质量更好的财务信息（Hwang 等 2017[26]）。就评估实践活动来说，受托的评估机构在业务的开展过程中，如评估计划编制、现场业务勘查以及评估报告编制，均需要委托方及相关当事人的大力配合，而配合得是否顺利很大程度上又体现在熟悉企业财务状况的 CFO 的沟通能力上。CFO 既需要将委托方管理层的未来战略意图及资产评估的目的准确、细致地传递给评估机构，从而使评估机构能够更加科学和有效地开展后续业务，又需要负责与被评估单位、评估机构间保持实时的沟通和联络。而女性 CFO 所具有的独特沟通交流技能，能够更有效地在评估多方之间传递信息，降低可能存在的评估信息不对称等问题，进而在一定程度上提升和改善由评估机构出具的评估报告信息质量。此外，就女性管理者的人格特质而言，其自身表现出的职业道德操守在财务行为中亦发挥着重要作用。Labelle 等（2010）[28]、Hwang 等（2017）[26]发现，总体而言，女性 CFO 的职业道德操守优于男性 CFO，女性高管通常不太愿意从事不道德的行为，以获取自身的经济利益（Adams 和 Ferreira，2009[7]；Jurkus 等，2011[29]），并且这种道德水准在缓解管理者机会主义行为中同样发挥着重要

作用，例如通过影响评估机构的独立性来影响评估结论的客观性。因而，作为委托方熟悉企业财务状况的 CFO，特别是女性 CFO，在职业、法律及法规的约束下，当出现委托人、相关当事方与评估机构在对评估结论产生理解偏差时，这种相对于男性的更强道德敏感性的人格特质更有可能使其及时提出异议，在降低执业风险的同时也倾向于接受更加谨慎的、避免由过高资产增值率所引致的增大企业风险的资产价值评估结论。此外，从代理成本角度来看，过高的资产增值率可以被视作企业代理问题的一种表现形式，并因其隐秘性等优势日益成为“内部人”实施自利性寻租行为的良好途径。进一步地，相较于男性高管成员，女性进入管理层能够有效降低公司的代理成本（Jurkus 等，2011[29]），并且女性高管的管理层机会主义行为相对较弱，表现出更为节制的自利型行为（Adams 和 Ferreira，2009[7]；Krishnan 和 Parsons，2008[25]）。因此，女性高管（如 CFO）对代理冲突的缓和可能伴随着相对较低的资产增值率的价值判断行为，也即女性 CFO 能够在一定程度上抑制被评估资产的高估行为。综合以上分析，本文提出假设 1。

H1：委托方的 CFO 若由女性担任时，其接受、认可的被评估资产增值率将显著低于男性 CFO，也即女性 CFO 与被评估资产价值增值率负相关。

三、研究设计

（一）样本和数据来源

中国资产评估协会自 2011 年起正式发布资产评估机构综合评价前百家名单，同时鉴于研究需运用前期财务指标，因此本文选取 2010—2016 年主板全部上市公司作为初始样本，并按照如下标准进行筛选：（1）结合资产评估报告所涉及的相关当事人，本文选择以下三种情形的样本公司，一是委托方、被评估单位以及上市公司三者是同一公司；二是委托方与上市公司两者是同一公司，被评估单位是前者的子公司；三是上市公司与委托方和被评估单位三者属于其他关系情形，主要为上市公司是以控制后两者为基础来确定合并财务报表的合并范围，后者财务报表均纳入到上市公司中期与期末合并财务报表编制和列报的范畴。（2）评估对象聚焦于企业股东全部权益价值的评估事件。（3）选择遵循以交易假设、公开市场假设为基本假设前提且价值类型为“市场价值”或“投资价值”的资产评估事件，破产清算假设及其对应的价值类型不包括在本研究中。（4）剔除金融类、ST、*ST 和数据不完整的公司所发生的资产评估事件。本文的资产评估机构综合评价指标数据来源于中国资产评估协会网站，委托方企业特征和被评估对象特征数据均来源于国泰安 CSMAR 数据库。数据统计分析工具为 R version 3.5.1。

（二）变量定义与度量

结果变量。本文的结果变量为资产评估机构对被评估单位的企业股东全部权益价值的评估增值率（*Ratio*）。以评估机构的评估值和企业股东全部权益的账面价值（或经调整后的账面价值）之差与企业股东全部权益的账面价值（经调整后的账面价值）之比计算得出。

处理变量。本研究中处理变量为 CFO 性别（*CFO Gender*），若 CFO 职位由女性担任赋值为 1，否则为 0。由此将样本初步划分为试验组（*CFO Gender* = 1）和对照组（*CFO Gender* = 0）。

协变量。参考先前学者的研究（何威风和刘启亮，2010[8]；王霞等，2011[30]；Huang 和 Kisgen，2013[14]；Francis 等，2015[13]；李小荣等，2013[11]；马海涛等，2017[6]；Ge 等，2011[10]；Hwang，2017[26]），本文控制了倾向指数模型中的重要协变量。

具体变量符号和定义见表 1。

表 1　　主要变量定义

变量符号	变量描述	定义及度量方法
Ratio	资产评估增值率	以资产评估机构的评估值和企业股东全部权益的账面价值（或经调整后的账面价值）之差与企业股东全部权益的账面价值（经调整后的账面价值）之比度量
CFOGender	CFO 性别	委托方是否有女性 CFO，有为 1，否则为 0
Profession	CFO 职业背景	CFO 是否拥有注册执业资格或财会高级职称，有为 1，否则为 0
Diversity	高管性别多元化	企业年度女性高管占比

续前表

变量符号	变量描述	定义及度量方法
Chair	女性董事长	委托方中是否有女性CFO，有为1，否则为0
Nomc	提名委员会设立	委托方中是否设立提名委员会，有为1，否则为0
Duality	两职合一	董事长和总经理两职合一为1，否则为0
Board	董事会规模	Ln（董事会人数）
Growth	成长性	（营业收入本年本期金额-营业收入上年同期金额）/（营业收入上年同期金额）
ROEgrowrate	净资产收益率增长率	（本期单季度净资产收益率-上年同期净资产收益率）/上年同期净资产收益率
ABSDACC	会计信息质量	截面修正琼斯模型的操控性应计利润的绝对值
Lev	资产负债率	负债总额/资产总额
Size	公司规模	Ln（企业总资产）
SOE	企业产权	实际控制人性质为国有取值为1，其他取0
Valuetype	评估价值类型	价值类型若为市场价值取值为1，其他取0
Appmth	评估方法	评估结论最终采用的方法为市场法取值为1，其他取0
P/Bratio	评估价值比率	每股市价/每股净资产
Intassetratio	无形资产比率	无形资产净额/期末总资产
Goodwillratio	商誉比率	商誉/期末总资产
Reputation	资产评估机构综合评价指标	Ln（资产评估机构综合评价年度业务收入）

资料来源：作者整理。

（三）模型构建与方法

Francis等（2015）[13]指出，对于企业而言，CFO是由女性担任还是由男性任职可能并不是随机委派的。因而，采用匹配分析方法能够通过基于一组可观测的特征来对试验组和对照组的样本观察值进行匹配，进而在一定程度上减少样本选择偏差问题（Carter等，2017[22]），也同时能够缓解诸如采用多元回归分析时由模型设定偏误所产生的遗漏变量问题（Swanquist等，2017[31]）。据此，本文在检验研究假设H1时首先采用传统的PSM分析方法进行分析，探究与男性CFO（Control Group）相比，当CFO职位由女性担任（Treatment Group）时对资产评估机构出具的评估报告中被评估资产价值增值率影响所产生的净效应。

运用PSM分析方法对模型进行设定时，首先需要构建倾向得分的Logit模型，对倾向得分进行获取后进行样本匹配。本文将CFO性别作为二分类处理变量（后续模型计算中用D_i表示，当CFO职位由女性担任时取1，否则取0），将所有协变量（后续模型计算中用X_i表示，代表样本企业多个维度的匹配特征）纳入倾向得分模型中，模型构建如下：

$$p(X_i)=E(D_i|X_i)=\frac{\exp^{(\beta x1)}}{1+\exp^{(\beta x1)}}$$
$$=\beta_1+\beta_2 Profession+\beta_3 Diversity+\beta_4 Reputation$$
$$+\beta_5 Valuetype+\beta_6 Appmth+\beta_7\frac{P}{B}ratio$$
$$+\beta_8 Intassetratio+\beta_9 Goodwillratio+\beta_{10} Growth$$
$$+\beta_{11} ROEgrowrate+\beta_{12} ABSDACC+\beta_{13} Nomc$$
$$+\beta_{14} Chair+\beta_{15} Duality+\beta_{16} Board+\beta_{17} Lev$$
$$+\beta_{18} Size+\beta_{19} SOE+u_{i,t}$$

通过Logit模型获取倾向得分后，本文采用“被处理单位的平均处理效应”（*ATT*）来估计CFO性别差异对被评估资产价值增值率的潜在影响，该值若显著为负，则表示相较于对照组，CFO由女性担任时对被评估资产价值增值率有显著的负向影响。其中*ATT*的计算过程如下：

$$ATT=E[Ratio_{1,i}-Ratio_{0,i}|D_i=1]$$

$$=E\{E[Ratio_{1,i}-Ratio_{0,i}|D_i=1,p(X_i)]\}$$

也即：

$$ATT=E\{E[Ratio_{1,i}|D_i=1,p(X_i)]-E[Ratio_{0,i}|D_i=0,p(X_i)|D_i=1]\}$$

其中，$Ratio_{1,i}$、$Ratio_{0,i}$分别表示同一家企业 CFO 职务是由女性担任和不是由女性担任的情况下所对应的被评估资产的增值率。

四、实证结果分析

（一）描述性统计分析

表 2 列示了对各主要变量的描述性统计结果，由于财务数据存在少数异常值，因此在全样本水平对连续变量 1% 以下和 99% 以上的分位数进行 winsorize 处理，以消除异常值对分析结果造成的偏差。在 1 226 个初始观测值中（涉及 861 家企业），平均来说，样本公司聘请评估机构针对企业净资产价值评定、估算后得出的评估增值率（*Ratio*）为 144.350%，标准差为 353.758，反映出资产评估机构受托对不同企业净资产价值评定估算的增值率存在较大差异，具有充分的变异性，这为本文提供了一个良好的研究契机。全样本中 31% 的 CFO 职位由女性担任（*CFO Gender*）。Panel C 列示了各协变量的基本统计指标。高管团队性别多样化变量（*Diversity*）的均值和中位数分别为 0.171 和 0.158，说明上市公司中女性高管成员的参与比例还比较低。从资产评估机构综合评价指标（*Reputation*）来看，其均值和中位数分别为 8.707 和 8.496，样本企业都选择了具有从事证券评估资格业务的资产评估机构，而且基本在中国资产评估协会发布的年度评估机构百强名单中上榜。从价值类型指标（*Valuetype*）来看，有近 97%的样本委托方选择市场价值。评估方法中选择市场法和收益法的占样本比例的 24.2%。市净率指标（*P/BRatio*）的均值和中位数分别为 4.622 和 2.931，标准差为 6.010，说明不同委托方的价值比率存在一定差异。无形资产比率指标（*Intassetratio*）的均值和中位数分别为 0.054 和 0.037。商誉比率指标（*Goodwillratio*）的均值和中位数分别为 0.018 和 0.000。本文还对 CFO 性别差异的资产评估增值率差异性进行初步分析，结果显示，总体上女性 CFO 组的资产评估增值率均值为 114.20%，男性 CFO 组的资产评估增值率均值为 158.10%，均值和中位数检验均表明，委托方 CFO 由女性担任时所认可的评估报告的增值率显著低于男性 CFO 组，并且在 5%的水平上显著。此外，针对相关性检验，Pearson 相关系数检验结果显示，各主要变量的相关系数都在 0.5 以下，因此不会产生明显的多重共线性问题。限于篇幅，相关系数检验结果未在文中列示。

表 2　　主要变量的描述性统计

Panel A：结果变量描述性统计								
Variable	N	Mean	Sd	Min	Q1	Median	P75	Max
Ratio	1 226	144.350	353.758	−592.000	4.283	34.663	154.996	1 703.000
Panel B：处理变量描述性统计								
CFOGender	1 226	0.313	0.464	0.000	0.000	0.000	1.000	1.000
Panel C：协变量描述性统计								
Profession	1 226	0.259	0.438	0.000	0.000	0.000	1.000	1.000
Diversityfemale	1 226	0.171	0.104	0.000	0.100	0.158	0.235	0.563
Reputation	1 226	8.707	1.173	6.477	7.775	8.496	9.724	10.976
Valuetype	1 226	0.962	0.192	0.000	1.000	1.000	1.000	1.000
AppmthIA	1 226	0.223	0.417	0.000	0.000	0.000	0.000	1.000
AppmthMA	1 226	0.019	0.136	0.000	0.000	0.000	0.000	1.000
P/Bratio	1 226	4.622	6.010	0.307	1.818	2.931	5.117	48.468
Intassetratio	1 226	0.054	0.067	0.000	0.015	0.037	0.064	0.664

续前表

Panel C：协变量描述性统计							
Goodwillratio	1 226	0.018	0.057	0.000	0.000	0.005	0.536
Nomc	1 226	0.857	0.350	0.000	1.000	1.000	1.000
Chair	1 226	0.040	0.196	0.000	0.000	0.000	1.000
Board	1 226	2.149	0.213	1.609	2.079	2.197	2.890
Duality	1 226	0.200	0.400	0.000	0.000	0.000	1.000
Growth	1 226	0.569	8.401	−0.991	−0.088	0.057	251.211
ROEgrowrate	1 226	−1.868	6.995	−37.644	−0.810	−0.213	0.157	6.545
ABSDACC	1 226	0.071	0.073	0.001	0.022	0.050	0.097	0.398
Lev	1 226	0.528	0.220	0.078	0.371	0.546	0.689	0.957
Size	1 226	22.136	1.415	16.757	21.124	21.986	22.998	27.955
SOE	1 226	0.484	0.500	0.000	0.000	0.000	1.000	1.000

资料来源：作者使用中国资产评估协会网站数据、国泰安CSMAR数据库并用R软件计算整理，下同。

（二）样本匹配效果分析

表3报告了以CFO性别为分类变量的PSM分析第一阶段Logit模型的估计结果。结果显示，模型整体非常显著（p值为0.000）。由于CFO Gender是二值变量，因而对于衡量分类变量模型设定的精确性和拟合优度好坏的判断方法通常有两种：一种是直接观察Logit模型的伪R^2，本文的伪R^2为0.111。另一种更为常用的是通过计算模型的正确预测比例，也即通过Logit模型计算绘制出ROC曲线下方的面积，得出AUC的值，以此来评价判定模型效果的好坏。本文的Logit模型计算绘制出的ROC曲线下面积（AUC）的值为0.729，也即协变量对处理变量的准确预测比率为72.9%。该模型结果表明，本研究的协变量能在一定程度上进行准确预测，因此模型中协变量的选择是较为合理的，构建的倾向得分模型也比较适用于估计研究样本的倾向值。

表3　　PSM第一阶段的Logit模型估计结果

处理变量：CFO Gender=1							
变量	*Diversity*	*Nomc*	*Lev*	*Reputation*	*Appmth*	*P/Bratio*	*Others*
系数	7.670***	0.513**	−0.868**	−0.134**	−1.555**	0.022**	Control
z值	10.42	2.47	−2.31	−2.29	−2.32	1.84	Control
Pseudo R^2	0.111	*LR chi2*	169.37	*Prob*	0.000	*N*	1 226

***、**、*分别表示在1%、5%、10%的水平上显著。下同。

进一步地，通过Logit模型计算倾向得分后，本文将采用常用的三种方法对试验组和对照组进行匹配。本文为比较匹配前后试验组和对照组倾向值差异性，还同步绘制出匹配前后相应的核密度函数曲线。结果显示，样本匹配前，试验组和对照组均具有左偏倾向，且样本中的两组数据倾向值的概率密度分布存在明显差异，如果不进行匹配，直接选取对照组全部样本作为比较对象将会产生严重的估计偏误。在实施匹配后，两组样本的倾向得分值的概率密度分布明显趋于一致，这表明两组样本经匹配后的多维度层面的特征已趋于一致，样本的选择性偏差已消除。表4报告了倾向值匹配分析的平衡性检验结果。结果表明，匹配后各协变量的标准化偏差均低于4%，远低于10%的平行假设规定的经验法则，且试验组和对照组的变量均值差异在匹配后均不显著。因此，平衡性假设得到满足。

表 4 **变量匹配前后差异对比**

变量	样本	均值		偏差%	减幅%	*t* 值
		试验组	对照组			
Profession	匹配前	0. 263	0. 257	1. 5	19. 20	0. 24
	匹配后	0. 264	0. 258	1. 2		0. 16
Diversityfemale	匹配前	0. 222	0. 147	75. 3	96. 2	12. 29
	匹配后	0. 221	0. 218	2. 8		0. 38
ABSDACC	匹配前	0. 070	0. 072	-4	46. 8	-0. 63
	匹配后	0. 069	0. 071	-2. 1		-0. 29
Growth	匹配前	0. 182	0. 745	-7. 8	95. 9	-1. 09
	匹配后	0. 182	0. 205	-0. 3		-0. 28
ROEgrowrate	匹配前	-1. 952	-1. 829	-1. 7	81. 4	-0. 28
	匹配后	-1. 957	-1. 934	-0. 3		-0. 04
Lev	匹配前	0. 501	0. 540	-17. 6	79. 6	-2. 88
	匹配后	0. 501	0. 508	-3. 6		-0. 48
Size	匹配前	21. 891	22. 248	-25. 9	91. 3	-4. 13
	匹配后	21. 891	21. 860	2. 3		0. 32
SOE	匹配前	0. 440	0. 504	-12. 7	95	-2. 06
	匹配后	0. 439	0. 435	0. 6		0. 09
Nomc	匹配前	0. 891	0. 842	14. 3	90	2. 26
	匹配后	0. 890	0. 895	-1. 4		-0. 22
Chair	匹配前	0. 049	0. 036	6. 9	79. 5	1. 15
	匹配后	0. 050	0. 047	1. 4		0. 18
Duality	匹配前	0. 221	0. 190	7. 7	87. 4	1. 27
	匹配后	0. 222	0. 226	-1		-0. 13
Board	匹配前	2. 143	2. 151	-4. 1	86. 1	-0. 67
	匹配后	2. 142	2. 141	0. 6		0. 08
Reputation	匹配前	8. 564	8. 772	-17. 8	77. 7	-2. 89
	匹配后	8. 562	8. 515	4		0. 56
Valuetype	匹配前	0. 966	0. 960	3. 4	8. 4	0. 55
	匹配后	0. 966	0. 972	-3. 2		-0. 48
Appmth	匹配前	0. 008	0. 024	-12. 8	80. 7	-1. 91
	匹配后	0. 008	0. 005	2. 5		0. 54

续前表

变量	样本	均值		偏差%	减幅%	t值
		试验组	对照组			
P/Bratio	匹配前	5.163	4.375	12.4	94	2.13
	匹配后	5.149	5.196	-0.7		-0.09
Intassetratio	匹配前	0.052	0.055	-4.2	42.1	-0.7
	匹配后	0.052	0.053	-2.4		-0.35
Goodwillrate	匹配前	0.019	0.017	2.5	52.4	0.4
	匹配后	0.019	0.019	-1.2		-0.15

（三）女性CFO组平均处理效应的结果分析

表5报告了CFO性别差异对被评估资产增值率的影响结果。以核匹配列示结果为例，匹配后女性CFO的资产评估报告增值率为114.815%，而作为对照组的男性CFO样本企业的评估报告增值率为176.977%，*ATT*值在1%水平上显著为负。此外，采用最近邻一对一匹配和半径匹配进行的配对结果显示，试验组的平均处理效应*ATT*值分别为-67.381和-73.690，在5%的水平上显著。在设定的半径匹配规则中，试验组有6例女性CFO未能匹配到合适的对照，同时整体结果也与核匹配基本保持一致。实证结果表明，如果委托方的CFO由女性任职，则由资产评估机构出具的资产评估报告的增值率无论在经济意义还是统计意义上均显著较小，也即与男性CFO组相比，女性CFO组的资产评估报告增值率平均降低幅度超过了62%。因此，总体而言，与男性CFO相比，女性CFO将会对被评估资产价值表现出更多的谨慎性、敏感性，其认可的评估报告的增值率也更低，以上结果与假设的推断相符。

表5　　CFO性别差异对被评估资产增值率的影响

变量	样本匹配		试验组	对照组	ATT	标准误	t值		Off support	On support
Ratio	核匹配	匹配前	114.199	158.099	-43.900	21.756	-2.02**	对照组	0	842
		匹配后	114.815	176.997	-62.182	23.598	-2.64***	试验组	1	383
	最近邻匹配	匹配前	114.199	158.099	-43.900	21.756	-2.02**	对照组	0	842
		匹配后	114.815	182.196	-67.381	26.416	-2.55**	试验组	1	383
	半径匹配	匹配前	114.199	158.099	-43.900	21.756	-2.02**	对照组	0	842
		匹配后	116.105	189.795	-73.690	34.876	-2.11**	试验组	6	378

最近邻匹配的配对比例为1∶1，控制组个体不会重复匹配，下同；半径匹配的卡尺范围采用最为普遍使用的设定数值0.01。

（四）稳健性测试

1. 基于倾向得分加权匹配分析的稳健性检验。

本部分我们采用Hirano和Imbens（2003）[32]提出的倾向得分加权匹配分析方法，对样本数据进行重新分析。从表6列示的结果看，通过该方法所进行的匹配变量前后差异明显减小，各协变量在匹配后的标准化平均差均大幅度下降，其标准化平均偏差均控制在10%的经验法则之内，因而通过了平衡性假设检验。此外，本文还对相应的倾向得分模型设定进行了检验，通过对倾向得分模型的拟合优度进行检验后，所得到的模型检验的p值为0.543，说明本文的模型设定不拒绝原假设，也即本研究的倾向得分模型适合在试验组和对照组之间进行协变量匹配。进一步地，采用倾向得分加权方法进行的稳健估计中ATT的值为-71.743，且在5%水平上显著为负，研究结果与前文保持一致。

表 6 **Propensity Score Weighting 分析**

倾向得分加权匹配分析——变量匹配前后差异对比						
变量	匹配前后	均值		标准差		标准化平均差（%）
		试验组	对照组	试验组	对照组	
Profession	U	0. 263	0. 257	0. 441	0. 437	1. 48
	M	0. 263	0. 263	0. 441	0. 441	0. 07
Diversityfemale	U	0. 222	0. 147	0. 100	0. 098	75. 26
	M	0. 222	0. 237	0. 100	0. 123	-13. 10
Reputation	U	8. 564	8. 772	1. 164	1. 171	-17. 80
	M	8. 564	8. 506	1. 164	1. 128	5. 01
Valuetype	U	0. 966	0. 960	0. 181	0. 197	3. 45
	M	0. 966	0. 971	0. 181	0. 168	-2. 81
Appmth	U	0. 008	0. 024	0. 088	0. 152	-12. 81
	M	0. 008	0. 006	0. 088	0. 076	2. 48
P/Bratio	U	5. 163	4. 375	7. 166	5. 388	12. 44
	M	5. 163	5. 083	7. 166	6. 543	1. 17
Intassetratio	U	0. 052	0. 055	0. 070	0. 065	-4. 23
	M	0. 052	0. 053	0. 070	0. 060	-1. 72
Goodwillratio	U	0. 019	0. 017	0. 058	0. 057	2. 45
	M	0. 019	0. 017	0. 058	0. 060	1. 81
ABSDACC	U	0. 070	0. 072	0. 067	0. 075	-3. 99
	M	0. 070	0. 069	0. 067	0. 073	0. 26
Nomc	U	0. 891	0. 842	0. 313	0. 365	14. 30
	M	0. 891	0. 896	0. 313	0. 305	-1. 87
Chair	U	0. 049	0. 036	0. 217	0. 185	6. 86
	M	0. 049	0. 046	0. 217	0. 209	1. 82
Board	U	2. 143	2. 151	0. 216	0. 211	-4. 08
	M	2. 143	2. 141	0. 216	0. 196	0. 92
Duality	U	0. 221	0. 190	0. 416	0. 393	7. 75
	M	0. 221	0. 221	0. 416	0. 415	0. 18
Growth	U	0. 182	0. 745	1. 288	10. 097	-7. 83
	M	0. 182	0. 186	1. 288	1. 137	-0. 34
ROEgrowrate	U	-1. 952	-1. 829	7. 458	6. 778	-1. 72
	M	-1. 952	-1. 835	7. 458	6. 720	-1. 65

续前表

倾向得分加权匹配分析——变量匹配前后差异对比						
变量	匹配前后	均值		标准差		标准化平均差（%）
		试验组	对照组	试验组	对照组	
Lev	U	0.501	0.540	0.227	0.216	−17.56
	M	0.501	0.507	0.227	0.234	−2.52
Size	U	21.891	22.248	1.298	1.452	−25.94
	M	21.891	21.813	1.298	1.428	5.67
SOE	U	0.440	0.504	0.497	0.500	−12.73
	M	0.440	0.418	0.497	0.494	4.47
Propensity Score Weighting 分析——CFO 性别差异对被评估资产增值率的影响						
变量	样本匹配	试验组	对照组	ATT	标准误	*t* 值
Ratio	PSW 分析	114.199	185.943	−71.744	28.47	−2.52**

注：采用倾向得分加权匹配方法进行估计时，试验组均能匹配到赋予不同权数的对照。

2. 基于 Matching Frontier 权匹配分析的稳健性检验。

在匹配方法的前沿动态研究中，King 等（2017）[33] 指出，目前流行的倾向值匹配（PSM）方法仅仅通过倾向得分本身进行匹配，在试验组与对照组的相似度、模型依赖度、统计效率、偏误等评价指标上仍有优化和改进的空间，因而本文也采用 King 等学者提出的一种前沿的匹配算法 Matching Frontier，通过该算法对样本联合优化分组匹配，得到经消除偏误和减少方差这两个维度同步达到最优化后的匹配样本，然后采用最小二乘法（OLS）对匹配样本进行估计检验。

（1）基于 Matching Frontier 算法的最优化匹配样本的构建。好的匹配方法应该同时在试验组与对照组的相似度上（消除偏误）和样本规模上（减少方差）同步实现最大化（King 等，2017[33]）。而目前流行的倾向值匹配分析缺乏一种原则性的解决方案去处理偏误——方差之间如何权衡的问题。也即现有的匹配方法是通过选择性地修剪观测数据来降低样本数据的不平衡，从而降低模型的依赖度。然而，修剪数据同时又会减少样本量，因而反过来又可能会增大最终估计值的方差。因此，研究者长期面临着偏误——方差的权衡困境。鉴于此，上述学者设计的 Matching Frontier 算法很好地解决了以上问题。该方法的直观思想是要求在每一次剔除观测单位时，要确保所获得的观测样本在所有可能的同等规模的观测样本中拥有最大的试验组与对照组的相似度。Matching Frontier 算法实现了对不平衡（在试验组和控制组之间）和样本大小的同步联合优化与精简，优化了分组匹配效率。该算法首先需要构造匹配边界，通过联合优化，构造出边界中每个点的子样本，从而形成平均马氏不平衡边界（Average Mahalanobis Imbalance）。本文将通过研究设计部分中的协变量去构造马氏不平衡边界。从计算结果来看，消除偏误和减少方差这两个维度同步达到最优化条件的边界曲线所相应计算得出与每种可能的样本量的最大平衡匹配的观测值为 165 个（包括试验组的 76 个观测值和对照组的 89 个观测值），也即在联合优化分组匹配过程中，修剪了 1 061 个观测值后，消除偏误和减少方差这两个维度同步达到最优化。结果进一步显示，大量对照组的观测并没有能够很好地匹配试验组观测，至少在资产负债率、企业规模价值比率等匹配维度方面。因此，通过 Matching Frontier 算法还需修剪更多的观测值（在试验组和对照组之间），以同步达到平衡性和样本量的联合最优。算法结果显示，基于优化后的最终匹配样本，经修剪后的剩余观测值为 165 个。因此，在得到经消除偏误和减少方差这两个维度同步达到最优化后的匹配样本后，本文将基于此进行后续的稳健性检验。

（2）基于 Matching Frontier 算法的最优化匹配样

本的稳健性检验。通过 Matching Frontier 匹配进行样本匹配处理后，本文采用 OLS 方法对结果进行估计。结果显示，委托方 CFO 若由女性担任时，其认可的被评估资产增值率在 10% 水平上仍显著低于男性 CFO（Gender 的系数为-70.573）。因而本文的结论保持稳健。

3. 其他稳健性测试。

本文还进行了以下系列稳健性测试：第一，上文对被评估资产增值率变量按照上下 1% 进行缩尾处理后其最大值和最小值仍然相差甚远，本文还分别按照上下 2% 和 5% 对被评估资产增值率进行缩尾后重新估计，结果保持稳健。第二，根据资产评估准则规定，有关评估机构原则上需要采取两种及以上评估方法对股东全部权益价值评估进行评定估算，并确定其中一种评估结论作为评估结果，因而本文出于谨慎性考虑，选取样本中资产评估机构目前最为普遍运用的成本法所确定的评估结论作为稳健测试样本（929 例观测值），剔除了易受主观影响较大的收益法和易受外部市场条件影响的市场比较法的样本，重新进行倾向得分匹配估计。从表 7 Panel A 列示的结果来看，研究结果保持稳健。第三，本文对被评估资产增值率结果变量进行替换，以资产评估机构的评估值与企业股东全部权益的账面价值（或经调整后的账面价值）的偏离程度来度量差异，也即以评估值和账面价值（或经调整后的账面价值）之差与账面价值（经调整后的账面价值）之比的绝对值计算得出新的因变量指标后，重新进行倾向得分匹配估计。从 Panel B 列示的结果来看，结果与前文结论一致，表现为委托方 CFO 若由女性担任时，其资产评估机构出具报告的评估值与被评估资产账面值差异更小，反映出 CFO 性别差异的确显著影响了对被评估资产的价值判断。第四，考虑到董事长和首席执行官是公司重要事务的最终决策者，本文借鉴 Xu 等（2016）[34] 和 Friedman 等（2014）[35] 的研究思路，通过加入更多的控制变量，选择董事长性别、首席执行官性别及 CEO 权力变量引入假定的模型中，其中 CEO 权力指标测度方法参考张祥建等（2015）[36]、权小锋和吴世农（2010）[37] 的研究，并重新进行估计。从 Panel C 列示的结果来看，我们并没有发现女性董事长和女性 CEO 的存在以及公司 CEO 的权力会显著影响到本文的研究结论。第五，本文还对由女性董事长领导或女性 CEO 领导或者由两职合一的女性领导的企业样本进行了剔除，研究结论仍未发生改变。上述测试结果表明，与董事长或首席执行官不同，CFO 是委托方资产评估业务活动的真正执行者，会直接参与到具体的评估实践过程中。因而，资产评估中介机构将首席财务官 CFO 而非首席执行官或董事长视为评估业务信息沟通和财务信息提供的主要高管，由此并最终引致其对委托评估的交易资产的价值判断产生影响。综上，以上结果均与前文保持一致，提升了研究结论的稳健性。

表 7　　其他稳健性检验结果

样本匹配		试验组	对照组	*ATT*	标准误	*t* 值		Common support	
Panel A	基于资产基础法所确定的样本							Off	On
核匹配	匹配前	37.492	69.522	-32.030	15.222	-2.10**	对照组	0	619
	匹配后	37.492	86.386	-48.894	16.316	-3.00***	试验组	0	310
最近邻匹配	匹配前	37.492	69.522	-32.030	15.222	-2.10**	对照组	0	619
	匹配后	37.492	69.296	-31.804	17.968	-1.77*	试验组	0	310
半径匹配	匹配前	37.492	69.522	-32.030	15.222	-2.10**	对照组	0	610
	匹配后	37.755	70.019	-32.264	19.794	-1.62	试验组	6	304
Panel B	基于资产增值率的偏离程度来度量的样本								
核匹配	匹配前	160.030	199.378	-39.347	20.491	-1.92*	对照组	0	842
	匹配后	160.030	230.440	-70.410	22.028	-3.20***	试验组	0	348
最近邻匹配	匹配前	160.030	199.378	-39.347	20.491	-1.92*	对照组	0	842
	匹配后	160.030	236.517	-76.487	25.131	-3.00***	试验组	0	348
半径匹配	匹配前	160.030	199.378	-39.347	20.491	-1.92*	对照组	0	842
	匹配后	162.088	231.486	-69.398	35.066	-1.98**	试验组	6	378

续前表

样本匹配		试验组	对照组	*ATT*	标准误	*t* 值		Common support	
Panel C	加入更多控制变量后的样本（CEO 性别和 CEO 权力）								
核匹配	匹配前	114.199	158.099	-43.900	21.756	-2.02**	对照组	0	842
	匹配后	114.816	179.140	-64.325	23.652	-2.72***	试验组	1	383
最近邻匹配	匹配前	114.199	158.099	-43.900	21.756	-2.02**	对照组	0	842
	匹配后	114.816	177.721	-62.906	26.037	-2.42**	试验组	1	383
半径匹配	匹配前	114.199	158.099	-43.900	21.756	-2.02**	对照组	0	842
	匹配后	115.050	185.375	-70.324	34.666	-2.03**	试验组	12	372

注：最近邻匹配的配对比例为 1∶1，控制组个体不会重复匹配；半径匹配的卡尺范围采用最为普遍使用的设定数值 0.01。

（五）内生性问题的识别与处理

根据本文的模型设定，要想获得 CFO 性别差异对被评估资产增值率的因果识别效应，还需要处理可能存在的内生性问题。Swanquist 等（2017）[31]指出，尽管传统的 PSM 经常会作为缓解内生性问题的一种方法，但是该方法仅能从降低因可观测的变量模型设定偏误而产生的遗漏变量问题，故并不能代替内生性检验来处理遗漏不可观测因素的问题。此外，Adams 和 Ferreira（2009）[7]发现，在高管性别（性别差异）对公司决策与公司财务行为影响的研究主题中，其内生性问题主要体现在遗漏了不可观测的企业特征因素（unobservable firm characteristics）。从某种程度来说，大多数不可观测的企业特征（特定异质性）通常在一定时期内并不会随时间发生变化或变化很慢（Carter 等，2017[22]）。因而，可以通过个体固定效应或 GMM 来处理该类问题（Adams 和 Ferreira，2009[7]；Carter 等，2017[22]；吕英和王正斌，2017[38]）。遗憾的是，由于本文所关注的资产评估事件为典型的混合数据，故无法直接采用面板数据的估计方法进行缓解。纵观前期高管性别对公司决策与公司财务行为影响的主题文献，我们发现，高管性别差异主题下对内生性问题探讨的焦点在于：遗漏变量的存在将可能会影响企业对于女性高管的雇用，也即核心解释变量（虚拟变量）与随机干扰项 ε 相关，那么可能会由此引发样本选择偏误这一经典内生性问题（Certo 等，2016[39]；Carter 等，2017[22]；Huang 和 Kisgen，2015[14]；Adams 和 Ferreira，2009[7]）。因此，针对可能由于样本选择所带来的内生性问题，本文将参考 Certo 等（2016）[39]、Carter 等（2017）[22]的做法，采用 IV+Treatment effects 模型来克服企业特定异质性对女性进入高管职位的内生自选择所造成的估计偏差的影响，以确保研究结论的可靠性。识别策略具体表现为：在处理效应模型中的第一阶段 Probit 选择方程中通过一组企业特征并引入两个外生工具（Instrument）来考察企业特定异质性对女性高管被雇用的可能性。对于外生变量的寻找，本文借鉴 Carter 等（2017）[22]、Xu 等（2016）[34]、Huang 和 Kisgen（2015）[14]的思路，选取 2009—2016 年度中国市级（分省份汇总）主要领导人（市、区长和市、区党委书记）的性别构成比例（Cityleaderratio）作为外生变量之一，其数据来源于 CSMAR 中国市级领导人数据库和人民网地方领导资料库进行收集整理。从性别不平等的相关文献来看，在一个州（市、区）中选出的女性主要领导人比例与当地女性担任高管职位的比例相关，因而一家企业总部位于该州（市、区）的公司雇用女性高管的可能性很大（Carter 等，2017[22]；Huang 和 Kisgen，2015[14]），但是企业所在州（市、区）的政府女性主要领导人比例与任何特定的本地企业的资产评估实践活动并不直接相关。故满足工具的相关性和外生性条件。此外，性别歧视比较严重的省份，女性高管较少（Xu 等，2016[34]），但是人口出生性别比与任何特定的本地企业的资产评估实践活动也并不直接相关。因而，本文选择公司总部所在省份的年度出生性别比例（性别比例）作为另一个重要的工具变量，其数据来源于 EPS 数据平台的中国卫生数据库。综上，第一阶段的 Probit 模型构建如下：

$$\begin{aligned} Pr(CFOGender=1|X_{i,t}) = \phi(&a_0+\beta_1 Birthsexratio \\ &+\beta_2 Cityleaderratio+\beta_3 Diversity \\ &+\beta_4 Nomc+\beta_5 Chair+\beta_6 Duality \\ &+\beta_7 Board+\beta_8 Lev+\beta_9 risk \\ &+\beta_{10} Firmage+\beta_{11} Size+\beta_{12} SOE) \end{aligned}$$

其中：被解释变量为是否有企业女性 CFO 虚拟变量，也即若企业聘任的 CFO 为女性时取值为 1，否则为 0。针对影响雇用女性高管的其他主要因素，借鉴已有文献对相关因素进行了控制（吕英和王正斌，2017[38]；Carter 等，2017[22]；Huang 和 Kisgen，2015[14]；李小荣等，2013[11]；Adams 和 Ferreira，2009[7]）。通过第一阶段模型，计算出风险比率（*Hazard Ratio*），然后代入原模型，进行第二阶段回归，以修正自选择偏差。表 8 列示的处理效应模型的两阶段回归结果显示，在第一阶段回归中，引入的外生变量出生性别比（*Birthsexratio*）前面系数为 -0.023（$z=-2.33$，$p<0.05$），说明使用的工具变量能够显著影响企业是否聘任女性 CFO 的决定，这和 Xu（2016）[34] 等的研究结论是一致。此外，外生变量中国市级主要领导人的性别构成比例与企业聘任女性 CFO 正相关，但不显著。在第二阶段回归中，女性 CFO 虚拟变量的系数为 -231.732（$z=-2.94$，$p<0.01$），说明考虑了企业特定异质性遗漏变量所可能产生的样本自选择问题后，CFO 性别差异对被评估资产价值增值率的影响仍然显著。

表 8　　内生性检验-Treatment effects 模型检验结果

变量	第一阶段 IV Probit		第二阶段	
	被解释变量：*CFOGender*		被解释变量：*Ratio*	
	系数	*z* 值	系数	*z* 值
Birthsexratio	-0.023	-2.33**		
Cityleaderfemaleratio	1.421	1.57		
Controls	Yes			
CFOGender			-231.732***	-2.94
Controls			Yes	
Constant			-190.336	-0.6
HazardRatio			0.298（0.129）	
Observations	766		766	

注：系数采用 Treatment effects 模型极大似然法进行估计。标准差按公司聚类和异方差调整，括号内数值表示对应系数的 *z* 统计量，*HazardRatio* 统计量括号内的结果以标准误值列示。

（六）影响机制分析

本部分将进一步分析女性 CFO 对被评估资产价值增值率的影响机制，对理论假设 1 进行实证检验。本文的假设 1 指出，除女性相比于男性拥有更强的沟通能力、更强的职业操守外，对风险承受的意愿可能构成 CFO 性别差异影响被评估资产价值增值率的重要渠道。如假设 1 所述，在遵循高管性别特质存在系统性差异的假定下，女性 CFO 所具有的更高风险规避程度将使其在财务行为与决策上更加谨慎，相比于男性具有更低的风险承受水平，其对风险承担的意愿将影响未来的预期收益的预测和期望投资报酬率的判断，进而影响其接受或认可的资产增值率的大小。本部分将对此影响机制予以检验，如果该机制成立，那么相比于男性 CFO，女性 CFO 与企业风险承担水平之间的关系将显著负相关，而风险承担的意愿将同步影响未来的预期收益的预测和期望投资报酬率，进而最终降低了资产评估增值率。

本文使用如下两个回归模型，其中，*RiskTake* 是企业风险承担变量（毛其淋和许家云，2016[40]；余明桂等，2013[41]），以企业在每一个观测时段内经行业调整的 *ROA* 的标准差来衡量。其余变量与前文保持一致。

$$Risktake=\alpha_0+\alpha_1 CFOGender+\theta_1 X+\varepsilon$$

$$Ratio=\beta_0+\gamma_1 RiskTake+\beta_2 CFOGender+\theta_2 X+\mu$$

表 9 列（1）和列（2）的结果显示，与男性 CFO 相比，女性 CFO 具有更低的企业风险承担水平，并在 1%的水平上显著。从列（3）~列（5）的结果来看，企业风险承担水平与被评估资产的增值率在 10%的水平上显著正相关，说明企业风险承担机制可能是重要

的。上述结果已经论证了企业风险承担是CFO性别差异影响被评估资产增值率的机制，那么企业风险承担能多大程度地解释CFO性别差异的影响作用呢？我们将CFO性别差异、企业风险承担同时放入回归，关注CFO性别差异的影响系数的变化程度，从而分析企业风险承担机制的解释程度，表9列（4）和列（5）报告的回归结果显示，当放入企业风险承担变量之后，女性CFO对被评估资产价值增值率的影响系数有一定程度的下降，但是显著性变化不大（相较于表5，仍在10%的水平上显著）。同时，企业风险承担对于被评估资产增值率的影响作用表现出稳健性和一致的正向显著影响，说明企业风险承担水平是稳定且重要的影响途径，支持了企业风险承担水平是CFO性别差异影响被评估资产价值增值率的机制。

表9　　影响机制探究：企业风险承担水平

变量	（1）	（2）	（3）	（4）	（5）
	RiskTake		*Ratio*		
Gender	−0.046*** (−7.27)	−0.045*** (−7.79)		−47.314* (−1.70)	−57.809** (−2.00)
RiskTake	—	—	504.34*** (3.45)	437.592*** (3.08)	265.084* (1.81)
Controls	no	yes	no	no	yes
Constant	0.020*** (4.95)	−0.318*** (−3.17)	150.067*** (11.39)	173.518*** (8.45)	−160.111 (−0.52)
Observations	766	766	766	766	766

注：系数采用最小二乘法进行估计。标准误按公司聚类和异方差调整。

五、拓展性分析：基于资产评估实践的不同情境

上述结果显示了CFO性别差异对评估机构出具的评估报告结论（增值率）的价值判断的显著影响，并且从多个角度进行稳健性检验后，结论依然成立。本部分分别从产权性质和资产评估机构综合评价指标出发，对主要结果分情境进行分析，进一步观察其效果的差异。

（一）企业所有权性质差异的影响

基于我国的制度背景，所有权性质是企业最为重要的一项微观制度安排。由于存在行政干预的因素，企业高管依赖自身特征对被评估资产进行价值判断的影响在国有企业和非国有企业之间可能会表现出一定的差异。本文认为，一方面，随着市场化改革推进的加快，政府对国有企业的行政干预不断减少，但是就现状而言，国有企业尚未完全做到真正的政企分开，最终控制人的重要角色仍由政府所扮演，对于企业的日常经营管理仍在一定程度上进行参与，对企业的干预也比较容易，企业的行为也较多地体现政府的意志（缪毅和胡奕明，2014[42]）。特别是从以往的资产评估业务来源来看，针对国有企业的评估服务是资产评估的主要业务来源，在重要的经济活动中，如兼并重组过程中的资产评估业务，地方政府对于资产的交易定价存在着直接干预的冲动，可能扭曲资产的评估价值（马海涛等，2017[6]）。另一方面，我国的资产评估机构从早期依附于政府部门逐渐向独立、客观、公正的中介服务机构过渡中，也仍会受到政府干预的影响（马海涛等，2017[6]）。综上分析，由于评估机构和评估对象（国有企业）均会在一定程度上受到来自行政干预的影响，因而，国有企业管理层完全依赖自身的特征对被评估资产进行价值判断的影响就难以充分发挥，会受到一定程度的约束，从而削弱了高管特征对其被评估资产价值专业判断的影响。相反，非国有企业的所有权相对比较清晰和简单，企业的管理者具有较大的自主性，由此使得相对于国有企业，非国有企业的高管特征对被评估资产价值的专业判断的影响能够发挥得更加充分。综上，在CFO性别差异与被评估资产价值增值率的关系上，国有企业与非国有企业的表现很可能会有显著差异。表10中Panel A的结果显示，CFO性别特征对国有企业的被评估资产价值增值率的降低效应低于非国有企业。也即相对于国有企业，女性CFO对被评估资产增值率的负向影响在

非国有企业更为显著。

（二）资产评估机构综合评价差异的影响

资产评估机构声誉机制在评估中是否有效发挥作用是另一个有趣而需要探索的问题。从目前的文献检索来看，有关对资产评估机构声誉机制作用探讨的实证研究较为鲜见（马海涛等，2017[6]），因而基于研究主题，本部分针对资产评估机构综合评价差异对 CFO 性别与被评估资产增值率两者关系的影响展开探索性分析。马海涛等（2017）[6]研究发现，资产评估机构的声誉越好，评估值越会被作为并购重组定价的依据，其表现出成交价与评估值差异越小。本文遵循先前学者的逻辑，首先，综合评价排名越高的评估机构在组织和人力资本中通常越高于中小机构，从专业性角度来说，因其具有更高的专业性，其做出的评估值也就更为准确，也更容易得到委托方或当事人的认可。其次，综合评价排名越高的机构也越容易获得业界认可，其更有可能传递出高品质信号，也更容易得到双方信赖。此外，高声誉机构也更加注重自己的声誉、形象、保持其独立性，其做出的评估值也能够得到业务双方的认可。由此推测，相对于排名高于等于中位数（高声誉）的机构，CFO 性别差异对资产价值判断的影响在排名低于中位数的机构样本中将表现得更加明显。

表 10 中 Panel B 的结果显示，在排名低于中位数的机构样本中，委托方 CFO 若由女性担任时，其认可的被评估资产增值率在 5%的水平上仍显著低于男性 CFO。在排名高于等于中位数（高声誉）的机构样本中，CFO 性别差异与被评估资产价值增值率仍呈现负相关关系，但在统计意义上不显著。这也即表明相对于排名高于等于中位数（高声誉）的机构，女性 CFO 对被评估资产增值率的负向影响在排名低于中位数的机构样本中更为显著。

表 10　　基于资产评估业务情境差异的影响

Panel A：国有企业和非国有企业子样本分组检验				
变量	A. 国有企业组（*N*=340）		B. 非国有企业组（*N*=425）	
Ratio	*ATT*	*t* 值	*ATT*	*t* 值
核匹配	−44.472	−1.66*	−89.651	−2.28**
最近邻匹配	−35.515	−1.43	−97.602	−2.55**
Panel B：资产评估机构综合评价子样本分组检验				
变量	A. 机构排名高于等于中位数（*N*=379）		B. 机构排名低于中位数（*N*=386）	
Ratio	*ATT*	*t* 值	*ATT*	*t* 值
核匹配	−59.939	−1.64	−68.709	−1.70*
最近邻匹配	−47.674	−1.40	−82.328	−2.08**

注：最近邻匹配的配对比例为 1∶1，控制组个体不会重复匹配。在国有企业子样本中，根据设定的核匹配和最近邻匹配规则，试验组有 1 例未能匹配到合适的对照；在非国有企业子样本中，根据设定的核匹配和最近邻匹配规则，试验组有 4 例未能匹配到合适的对照；在机构排名高于等于中位数子样本中，根据设定的核匹配和最近邻匹配规则，试验组有 4 例未能匹配到合适的对照；在机构排名低于中位数子样本中，根据设定的核匹配和最近邻匹配规则，试验组有 8 例未能匹配到合适的对照。

六、研究结论与讨论

本文首次揭示了委托方 CFO 性别差异对资产评估机构出具的评估报告结论（增值率）的价值判断影响。研究发现，总体而言，委托方女性 CFO 所认可的被评估资产增值率显著低于男性 CFO，其资产增值率平均比男性 CFO 至少低 62%以上。研究还发现，委托方 CFO 性别差异对国有企业的被评估资产价值增值率的降低效应低于非国有企业，在相对低声誉的资产评估机构样本中委托方 CFO 对被评估资产价值的专业判断的性别效应更加显著。此外，委托方企业风险承担水平是 CFO 性别差异影响被评估资产价值增值率的稳定且重要的路径。

本文的研究在理论层面上响应了 Francis 等学者有关在企业经济行为中性别效应这一开放性问题值得研究的呼吁，突破了以往从股东或投资者视角考察 CFO 性别差异对财务行为影响的研究局限，为学界从资产评估角度对被评估资产价值判断中的性别问题

提供了更深入的理解。在政策层面，本文研究的启示意义在于：第一，在委托方企业里，与董事长或首席执行官不同，CFO 是委托方资产评估业务活动的真正执行者，资产评估机构也将 CFO 视为直接参与到评估活动中的业务信息沟通和财务信息提供的主要高管，在资产评估机构对被评估资产的价值提供评定、估算服务的过程中，委托方 CFO 对被评估资产价值的专业判断的性别效应是存在的且具有显著差异，这将使资产评估机构对相关当事方的交易所提供的定价参考依据产生重要影响。第二，与男性相比，女性 CFO 与评估机构展开业务合作过程中，对评估结论在一定程度上表现得更为谨慎、敏感及保守，其认可的评估增值率和实际资产账面价值的偏离程度均显著低于男性 CFO，这表明 CFO 的性别效应在一定程度上有利于抑制被评估资产的虚增和高估，特别是针对在当前资本市场上企业资产价值评估业务中存在不同程度的被评估资产虚高现象，女性 CFO 可以在这一个过程中充分发挥其治理效应，对企业资产价值起到监督作用。第三，在资产评估实践活动中，CFO 性别会影响其决策行为和效果，是影响资产价值评定、估算的重要因素之一。而从目前行业机构的具体业务实践开展情况来看，现有实务界对资产价值的评定估算实践过程中，其评估机构所采用的评估方法（市场比较法、收益法、成本法）都对评估对象和评估方法中各参数的选择给予足够的重视，但对委托方企业特征（如高管特征）对资产价值评估的影响缺乏足够考虑，而对这一微观重要影响因素的考虑将有利于提高资产评估机构的评估技术方面的创新性和先进性，也有利于更好地发现资产价值，提升自身的执业水平。此外，对于中国资产评估行业协会而言，应当引导和鼓励更多的评估机构探索实践中可能影响资产价值评估的重要因素，丰富现有的评估技术并完善相应的业务指导。

参考文献

[1] 北京证监局评估监管课题组. 上市公司使用资产评估报告存在的问题及解决对策 [J]. 中国资产评估，2017（2）：36-39.

[2] 张国春. 资产评估与供给侧结构性改革 [J]. 中国资产评估，2016（4）：6-12.

[3] 纪益成. 论资产评估的基本功能 [J]. 中国资产评估，2016（1）：19-25.

[4] 余炳文，张君来. 行业分析在企业价值评估中应用的再探析 [J]. 中国资产评估，2016（10）：32-38.

[5] 李小荣，薛艾珂，张静静. 管理者特征对企业价值评估中收益法运用的影响 [J]. 中国资产评估，2016（9）：29-33.

[6] 马海涛，李小荣，张帆. 资产评估机构声誉与公司并购重组定价 [J]. 中国软科学，2017（5）：101-118.

[7] Adams R B，Ferreira D. Women in the Boardroom and Their Impact on Governance and Performance [J]. Journal of Financial Economics，2009，94（2）：291-309.

[8] 何威风，刘启亮. 我国上市公司高管背景特征与财务重述行为研究 [J]. 管理世界，2010（7）：144-155.

[9] Srinidhi B，Gul F A，Tsui J. Female Directors and Earnings Quality [J]. Contemporary Accounting Research，2011，28（5）：1610-1644.

[10] Ge W，Matsumoto D，Zhang J L. Do CFOs Have Style? An Empirical Investigation of the Effect of Individual CFOs on Accounting Practices [J]. Contemporary Accounting Research，2011，28（4）：1141-1179.

[11] 李小荣，刘行，傅代国. 女性 CFO 与资金配置 [J]. 经济管理，2013（12）：100-110.

[12] Francis B B，Hasan I，Wu Q，et al. Are Female CFOs Less Tax Aggressive? Evidence from TaxAggressiveness [J]. Journal of the American Taxation Association，2014，36（2）：171-202.

[13] Francis B，Hasan I，Park J C，et al. Gender Differences in Financial Reporting Decision Making：Evidence from Accounting Conservatism [J]. Contemporary Accounting Research，2015，32（3）：1285-1318.

[14] Huang J，Kisgen D J. Gender and Corporate Finance：Are Male Executives Overconfident Relative to Female Executives? [J]. Journal of Financial Economics，2013，108（3）：822-839.

[15] Khlif H，Achek I. Gender in Accounting Research：A Review [J]. Managerial Auditing Journal，2017，32（6）：627-655.

[16] Figner B，Weber E U. Who Takes Risks When and Why Determinants of Risk Taking [J]. Current Directions in Psychological Science，2011，20（4）：211-216.

[17] Borghans L，Bart H H G，Heckman J J，et al. Gender Differences in Risk Aversion and Ambiguity Aversion [J]. Journal of the European Economic Association，2009，7（2）：649-658.

[18] Niederle M，Vesterlund L. Do Women Shy Away from Competition? Do Men Compete Too Much? [J]. Quarterly Journal of Economics，2007，122（3）：1067-1101.

[19] 李小荣，刘行. CEO vs CFO：性别与股价崩盘风险 [J]. 世界经济，2012 (12)：102-129.
[20] 周泽将，刘文惠，刘中燕. 女性高管对公司财务行为与公司价值的影响研究述评 [J]. 外国经济与管理，2012 (2)：73-81.
[21] Agnew J，Balduzzi P，Sundén A. Portfolio Choice and Trading in a Large 401 (k) Plan [J]. American Economic Review，2003，93 (1)：193-215.
[22] Carter M E，Franco F，Gine M. Executive Gender Pay Gaps：The Roles of Female Risk Aversion and Board Representation [J]. Contemporary Accounting Research，2017，34 (2)：1232-1264.
[23] Bernasek A，Shwiff S. Gender，Risk，and Retirement [J]. Journal of Economic Issues，2001，35 (2)：345-356.
[24] Olsen R A，Cox C M. The Influence of Gender on the Perception and Response to Investment Risk：The Case of Professional Investors [J]. Journal of Psychology & Financial Markets，2001，2 (1)：29-36.
[25] Krishnan G V，Parsons L M. Getting to the Bottom Line：An Exploration of Gender and Earnings Quality [J]. Journal of Business Ethics，2008，78 (1/2)：65-76.
[26] Hwang I，Kim H T T，Pae S S. Do the Properties of Analyst Earnings Forecasts Improve When Firms Are Managed by Female CFOs? [J]. Quarterly Journal of Finance and Accounting，2017，55 (2)：27-51.
[27] 姜付秀，伊志宏，苏飞，等. 管理者背景特征与企业过度投资行为 [J]. 管理世界，2009 (1)：130-139.
[28] Labelle R，Gargouri R M，Francoeur C. Ethics，Diversity Management，and Financial Reporting Quality [J]. Journal of Business Ethics，2010，93 (2)：335-353.
[29] Jurkus A F，Park J C，Woodard L S. Women in Top Management and Agency Costs [J]. Journal of Business Research，2011，64 (2)：180-186.
[30] 王霞，薛跃，于学强. CFO 的背景特征与会计信息质量——基于中国财务重述公司的经验证据 [J]. 财经研究，2011 (9)：123-133.
[31] Swanquist Q T，Shipman J，Whited R L. Propensity Score Matching in Accounting Research [J]. The Accounting Review，2017，92 (1)：213-244.
[32] Hirano K，Imbens G W，Ridder G. Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score [J]. Econometrica，2003，71 (4)：1161-1189.
[33] King G，Lucas C，Nielsen R. The Balance-sample Size Frontier in Matching Methods for CausalInference [J]. American Journal of Political Science，2017，61 (2)：473-489.
[34] Xu X，Li Y，Chang M. Female CFOs and Loan Contracting：Financial Conservatism or Gender Discrimination? —An Empirical Test Based on Collateral Clauses [J]. China Journal of Accounting Research，2016，9 (2)：153-173.
[35] Friedman H L. Implications of Power：When the CEO Can Pressure the CFO to Bias Reports [J]. Journal of Accounting and Economics，2014，58 (1)：117-141.
[36] 张祥建，徐晋，徐龙炳. 高管精英治理模式能够提升企业绩效吗？——基于社会连带关系调节效应的研究 [J]. 经济研究，2015 (3)：100-114.
[37] 权小锋，吴世农. CEO 权力强度、信息披露质量与公司业绩的波动性——基于深交所上市公司的实证研究 [J]. 南开管理评论，2010 (4)：142-153.
[38] 吕英，王正斌. 同性相吸还是同性相斥——中国上市公司女性高管任命性别溢出效应研究 [J]. 外国经济与管理，2017 (12)：84-99.
[39] Certo S T，Busenbark J R，Woo H，et al. Sample Selection Bias and Heckman Models in Strategic Management Research [J]. Strategic Management Journal，2016，37 (13)：2639-2657.
[40] 毛其淋，许家云. 政府补贴、异质性与企业风险承担 [J]. 经济学 (季刊)，2016 (4)：1533-1562.
[41] 余明桂，李文贵，潘红波. 民营化、产权保护与企业风险承担 [J]. 经济研究，2013 (9)：112-124.
[42] 缪毅，胡奕明. 产权性质、薪酬差距与晋升激励 [J]. 南开管理评论，2014 (4)：4-12.

(学术顾问：吴　溪，责任编辑：王克方　张安平)

财政分权与政府承诺力：基于信息传递的视角

Fiscal Decentralization and Government Commitment: From the Perspective of Information Transmission

尹训东　胡思平

YIN Xun-dong　HU Si-ping

［摘　要］财政分权是指中央政府与地方政府在财政收入和支出方面的权责划分，中央政府将部分财政决策权交由地方政府，允许地方政府在一定程度上决定其税收收入和财政支出。与中央政府相比，地方政府更加了解当地居民和企业的真实信息，因此地方政府拥有一定的信息优势，能更有效地制定财政支出计划并进行资源配置。本文采用委托代理的方法来分析财政分权是如何使得中央政府有效利用地方政府的信息，我们首先基于信息传递的视角来构建不同财政体制下的单期理论模型，包括中央政府集权、廉价磋商和财政分权。在单期模型的基础上，本文重点研究了两期模型下的最优财政支出决策权的下放问题。两期模型的复杂之处在于我们考虑中央政府在不同承诺能力下的最优机制，即中央政府有完全的承诺能力、有限的承诺能力和没有任何承诺能力的情况，并重点探讨在有限承诺能力下中央政府的最优机制设计问题。

［关键词］财政分权　中央集权　廉价磋商　有限承诺

［中图分类号］F812.2　［文献标识码］A　［文章编号］1000-1549（2020）10-0088-14

Abstract: Fiscal decentralization refers to the division of labor and cooperation between the central government and local governments in fiscal revenue and expenditure. The central government has delegated some fiscal decision-making power to local governments, allowing it to make some independent decisions on their tax revenues and fiscal expenditures. Compared with the central government, local governments have a better understanding of local residents' preferences and utility and other information, so local governments have certain information advantages and can more effectively formulate fiscal expenditure plans and allocate resources. In this paper, we use the principal-agent approach to analyze how fiscal decentralization enables the central government to make effective use of local government information. Based on the perspective of information transmission, we construct a theoretical model of fiscal system, including centralization of central government, cheap talk and fiscal decentralization. Based on the single-period model, this paper focuses on the decentralization of optimal financial expenditure decision-making power under the two-period model. The complexity of the two-phase model is that we consider the optimal mechanism of the central government under different commitment capacities, which are cases under which the central government has full commitment capacity, limited commitment capacity, and no commitment capacity.

Key words: Fiscal decentralization　Centralization　Cheap talk　Limited commitment

［收稿日期］　2020-03-10

［作者简介］　尹训东（通讯作者），男，1980 年 11 月生，中央财经大学中国公共财政与政策研究院副教授，研究方向为激励理论、博弈论、产业组织；胡思平，女，1994 年 8 月生，就职于平安银行北京分行，研究方向为微观经济理论。

［基金项目］　中央高校基本科研业务费专项“政府监管体制的最优组织结构选择”（项目编号：020550319001）。

感谢匿名评审人提出的修改建议，笔者已做了相应修改，本文文责自负。

一、前言

本文中的财政分权主要是从信息传递的角度出发来研究财政支出分权的好处。在财政收入方面，主要涉及的是中央政府对地方政府的激励问题，这一激励方式大致可以分为两类。一类是官员晋升制，通过对地方政府官员提供职位晋升的渠道来激励地方官员提升 GDP，从而增加税收收入；另一类是财政分成制，通过给予地方政府一定比例的税收收入，中央政府达到激励地方政府收税的目的。然而在财政支出方面，与中央政府相比，地方政府更加了解当地居民和企业的真实信息，拥有一定的信息优势，能更有效地制定财政支出计划并进行资源配置。从信息传递的角度来看，中央政府如何有效地进行财政支出分权以利用地方政府的私有信息是决定财政支出配置效率的关键。目前来说，从信息传递的角度来探讨中央政府和地方政府在财政支出方面的分权问题的文献较少，本文在一定程度上弥补了在这一领域的文献空白。

中央政府在设计最优分权机制时，必须考虑自身的承诺能力。首先考虑单期的情况，如果中央政府没有承诺能力，那么中央政府和地方政府之间进行的就是廉价磋商（cheap-talk）的博弈。地方政府有地方上的信息优势，而缺乏地方信息的中央政府掌握决策权，中央政府和地方政府两者的偏好有差异。廉价磋商博弈的均衡说明：双方之间的信息传递是有限的，存在模糊信息传递的均衡，并且双方的偏好差异越大，信息传递的越少。然后，我们考虑单期时中央政府有承诺力的情形，在双方的偏好差异不大的情况下，中央政府最优的机制就是给地方政府规定一个财政支出上限，不超过该上限，决策权完全让渡给地方政府。

在单期模型的基础上，本文重点研究了两期模型下的最优财政支出决策权的下放问题。两期模型的复杂之处在于我们考虑中央政府在不同承诺能力（commitment power）下的最优机制，即中央政府有完全的承诺能力、有限的承诺能力和没有任何承诺能力的情况。完全的承诺能力指中央政府可以在第一期承诺两期的合同；有限的承诺能力指中央政府在第一期只能承诺第一期的合同，而不能承诺第二期的合同，只能到第二期的期初制定第二期的合同；没有任何承诺能力指双方之间每期只能进行廉价磋商的博弈。本文假设两期下的真实状态信息不随时期改变。

二、文献综述

总的来说，我们可以将财政分权理论的发展分为以下两个时期：第一代财政分权理论和第二代财政分权理论。第一代财政分权理论的开始以 Tiebout 于 1956 年发表的《地方公共支出的纯理论》[1] 为标志，随后经过了 Musgrave（1959）[2] 和 Oates（1972）[3] 等人的发展而成。第一代财政分权理论是从地方政府的信息优势和中央与地方的职能分工角度来阐述财政分权理论。这一理论认为地方政府间的竞争会促使地方政府提供帕累托最优的公共产品，但是却没有分析政府官员为什么会存在提供最优公共产品和提升社会福利水平的动机，即缺乏对可实施的分权机制的说明。

第二代财政分权理论将微观经济学的激励相容和机制设计理论引入了财政分权领域。该理论认为，政府官员不是完美的和无私的，作为经济个体，政府官员也会寻求自身的利益最大化，也需要外在的约束来规范其行为。一旦分权机制未给予政府足够的激励，将会滋生其寻租行为，从而降低社会福利水平。这一理论的代表性人物为 Qian 和 Roland（1998）[4]、Qian 和 Weingast（1997）[5] 等。

在两代财政分权理论的基础上，国内外学者进行了许多财政分权方面的实证研究。Akai 和 Sakata（2002）[6] 利用美国 50 个州 1988—1992 年的数据分析了这一影响，他们发现当选用不同的指标来表示财政分权时，所得出的结论不同，因此 Akai 和 Sakata 并未明确阐述财政分权与经济增长之间的关系。通过用地方政府的边际留成比例来表示财政分权，林毅夫等（2000）[7] 发现财政分权与经济增长之间存在正相关的关系，对此的解释是在财政分权制度下，地方资源配置得到帕累托改善，从而促进了经济增长。

在此之后，张晏和龚六堂（2005）[8] 发现财政分权能促进经济增长，但这一效应存在时间上和地区间的差异，文章表明财政分权的影响在经济发达的中东部地区和经济欠发达的西部地区是不同的，同时还发现财政分权主要是通过提高各级政府之间的沟通协作能力来促进经济增长。之后，周业安和章泉（2008）[9] 的研究中还讨论了财政分权对经济波动的影响，发现财政分权是导致经济波动的重要原因。

委托代理问题的研究源于 20 世纪 70 年代，在最开始的研究中，学者们假设委托人可以通过转移支付来激励代理人，根据代理人的业绩表现给予不同的报酬。其基本的结论是委托人为了获取代理人的私人信

息，必须给予代理人一定的信息租金。还有一类激励机制，委托人不能使用转移支付来激励代理人，而是通过让渡决策权来激励代理人，比如政府上下级之间的决策问题，很多时候并不涉及用薪酬来激励下级的行为，而是上级衡量是否把决策权下放给下级，以便让下级更有效地利用当地的信息。如果不下放决策权，下级可能会为了自身的利益谎报（或瞒报）其私人信息，而如果下放决策权，下级根据自己的私人信息所选择的行为通常并不是上级最偏好的。这种决策权的分配和信息传递的矛盾构成了是否下放决策权的关键。本文基于没有转移支付的激励理论，从信息传递的角度，研究中央政府最优财政决策权的下放机制，即最优财政分权的理论。

一般的委托代理模型中，委托人可以给予代理人转移支付来激励代理人按照自己的偏好行动，根据不同的结果，给予不同的报酬，以便激励代理人进行合意的行动。但 Holmstrom（1977）[10] 开创了另外一类的代理关系，在这种代理关系中，委托人由于某种限制而不能使用转移支付对代理人进行激励，为了利用代理人的私人信息，委托人不得不让渡自己的决策权给代理人，让代理人根据自己掌握的信息自由选择行动。本文研究的代理关系属于后面这种情形。

当只存在一期的代理时，Melumad 和 Shibano（1991）[11] 以及 Martimort 和 Semenov（2006）[12] 求出了静态最优代理机制，也就是给出了在一定条件下为代理人的决策设置一个上限是最优的。之后，Martimort 和 Koessler（2012）[13] 将这一研究延伸至多维代理问题。之后的文献包括有 Dessein（2002）[14]，Alonso 和 Matouschek（2008）[15]，Ambrus 和 Egorov（2017）[16]。在本文的分析中，我们假设在任何时候，代理人的最理想的决策总是与委托人的最理想的决策不同，也就是代理人和委托人之间总是存在偏好上的偏差。上述文献都只研究了一期下的委托代理问题，然而在现实生活中，这种代理关系通常是多期的，本文的目的也是将政府在不同承诺力下的两期代理问题引入财政分权问题。

在有限承诺力的两期代理问题下，我们假设委托人不能对自己第二期的行为做出承诺，委托人在第二期将根据在第一期获得的信息做出自身最优的决策，从而可能损害代理人的效用，这将导致代理人在第一期不愿意暴露自己的信息。Raphael 和 Tracy（2011）[17] 也分析了动态代理问题，但他们假设在每一期，真实信息是独立且均匀分布的，代理人知道真实信息的分布状况但不知道具体数值，而委托人连分布状况也不可知。他们认为在这种情形下，学习的能力是至关重要的。与上述文章不同的是，本文假设真实信息在两期是相同的，且都是均匀分布，委托人想获取真实信息的唯一途径就是给予代理人足够的激励，促使其揭露真实信息。另一个不同之处在于他们的文章只考虑了委托人有完全承诺和完全没有承诺的两种情形，而本文考虑的是委托人的有限承诺。本文的模型借鉴了 Yin（2012）[18] 的研究，需要指出的是，本文的模型并不仅限于运用在财政分权领域中，也可以运用于其他类似的委托代理情况，因此具有一定的通用性。但在财政分权中，本文所采用的模型具有针对性，在中央政府拥有决策权的情形下，中央政府的主要目标是充分利用地方政府的信息优势，因而从信息传递的视角来研究财政分权就非常恰当。

三、理论模型

本文从信息传递的角度构建模型，在以下所有模型中，我们假设地方政府更了解当地的具体信息，比如当地居民对各种公共物品的真实偏好、当地的风俗习惯和当地的经济水平等等。根据这些信息，可以得到当地的最优财政支出结构，比如地方政府的财政支出在经济发展和文化娱乐发展两方面的比例。一般来说，一个地方的发展要同时兼顾物质和精神两方面，物质方面就是指经济水平，精神方面包括文化娱乐设施和文化娱乐活动等。我们假设当地的这些信息是地方政府的私人信息，而中央政府只知道信息的概率分布。

同时，中央政府和地方政府的偏好不同，也就是各自对于当地的最优财政支出结构的偏好不同。对于地方政府而言，当其自身激励与居民福利相契合时，其最优支出结构就是最大化居民效用的支出比例；但是中央政府要考虑的不仅是某个区域的利益，更是全国整体的利益，因此中央政府需要统一协调全国各地的公共产品供给，实现资源在全国范围内的最优配置。正是由于两者的角色不同，中央政府与地方政府在当地财政支出水平上必然存在不同的偏好，双方的效用函数存在差异。当决策权分别掌握在中央政府和地方政府手中时，两者将根据各自获得的信息和偏好决定当地的财政支出结构，也就是决策 y。

地方政府观测到当地的真实信息为 t，这一真实

信息只有地方政府了解，中央政府只知道 t 均匀分布于［0，1］区间。地方政府根据真实状态 t 向中央政府传递信息 r，r 同样是位于［0，1］区间，中央政府根据传递的信息做出决策 y。此时中央政府的收益函数为 $-(y-t)^2$，地方政府的收益函数为 $-(y-(t+b))^2$，$b>0$ 表示中央政府与地方政府之间的收益偏差，两者的收益都受到真实状态 t 和实际决策 y 的影响。当真实状态为 t 时，中央政府的最优决策为 $y=t$，而地方政府的最优决策水平为 $y=t+b$。正是由于地方政府的最优决策水平与中央政府不同，所以当最终的决策权掌握在中央政府手中时，地方政府才有动机使传递的信息 r 偏离真实状态 t，从而使中央政府的决策偏向于 $y=t+b$。

（一）单期情况下的模型

本文先讨论单期情况下的模型，包括没有信息传递的中央集权模型、中央政府完全没有承诺能力的廉价磋商模型、中央政府有承诺能力的财政分权模型。在单期模型的基础上，本文重点研究了两期模型下的最优财政支出决策权的下放问题。

1. 无信息传递：中央集权模型。

在中央政府集权的情况下，中央政府和地方政府之间不存在信息传递，此时中央政府唯一已知的信息是真实状态 t 均匀分布在［0，1］区间，给定中央政府关于 t 的分布信息和其收益函数的形式，我们可以证明其最优选择为 $y=1/2$。证明如下：

将 t 在［0，1］区间上的中央政府的收益进行积分：

$$\int_0^1 -(y-t)^2\mathrm{d}t=-\left(y^2-y+\frac{1}{3}\right)$$

一阶条件为：

$$\mathrm{d}\left[-\left(y^2-y+\frac{1}{3}\right)\right]\Big/\mathrm{d}y=-2y+1=0 \tag{1}$$

对于中央政府而言，在没有任何信息的情况下，最优的决策是 $y=1/2$。总的来说，我们发现，在中央集权模型中，当完全由中央政府决定财政支出水平时，由于中央政府不清楚地方的真实信息，只知居民的真实信息均匀分布于［0，1］区间，此时中央政府的最优选择为区间的中间值 1/2。当中央政府与地方政府之间的偏好差异较大时，对于中央政府而言，$y=1/2$ 的决策是最优的。但当两者之间的偏好差异较小时，信息传递的缺乏就会导致双方的福利损失，且随着偏差越小，造成损失的可能性就越大。

2. 完全无承诺能力：廉价磋商模型。

现在我们考虑中央政府根据地方政府传递的信息进行决策的情况。我们假设中央政府保留决策的权力，并要求地方政府汇报其关于地方上的私有信息，然后中央政府再根据此汇报的信息选择地方上的财政支出水平。我们可以应用 Crowford 和 Sobel（1982）[19] 的模型来求解这种情况下的信息传递和最优决策，其基本结论是当中央和地方的偏好差异比较小的时候，在均衡时，存在模糊的信息传递或者部分信息传递。简言之，在均衡的时候，地方政府会根据真实的 t 值所在的不同区间，传递不同的信息。真实状态 t 均匀分布在［0，1］区间，地方政府会将该区间划分成 K 个不相交的区间。令 $t_0=0$，$t_K=1$，当 $t_0\leqslant t<t_1$ 时，地方政府传递信息 r_1；当 $t_1\leqslant t<t_2$ 时，地方政府传递信息 r_2；以此类推，当 $t_k\leqslant t<t_{k+1}$ 时，地方政府传递信息 r_{k+1}；最终，当 $t_{K-1}\leqslant t\leqslant t_K$ 时，传递信息为 r_K。当中央政府接收到信息 r_k 时，便能判断出真实的 t 值位于区间［t_{k-1}，t_k］内，给定中央政府认为 t 值总是在区间内均匀分布，因此最优选择为 $y=(t_{k-1}+t_k)/2$。而对地方政府来说，中央政府的决策也成为其最优选择必须满足如下条件：$-[(t_{k-1}+t_k)/2-(t_k+b)]^2=-[(t_k+t_{k+1})/2-(t_k+b)]^2$。也就是说当 $t=t_k$ 时，地方政府传递信息 r_k 和 r_{k+1} 是无差异的。因此 t_k+b 应为 $(t_{k-1}+t_k)/2$ 和 $(t_k+t_{k+1})/2$ 的中间值，即 $t_k+b=[(t_{k-1}+t_k)/2+(t_k+t_{k+1})/2]/2$，由此可得 $t_{k+1}-t_k=t_k-t_{k-1}+4b$。由该式可知，第 $k+1$ 个区间比第 k 个区间长 $4b$，已知第一个区间的长度为 t_1 且所有区间长度相加为 1，则有 $t_1+(t_1+4b)+\cdots+(t_1+(K-1)4b)=1$，化简可得：

$$Kt_1+2bK(K-1)=1 \tag{2}$$

因为要求 $t_1>0$，即 $2bK(K-1)<1$，只有当 b 足够小时，才能满足这一条件。若 $t_1>0$ 且 $2bK(K-1)<1$ 条件得到满足，则存在多个均衡。假设 $2bK(K-1)=1$，则最大的 K 值为 $K=(1+\sqrt{1+2/b})/2$。b 值越小，K 值越大，也就是中央政府与地方政府的收益偏差越小，则信息的传递就越精准。

举例来说，如图 1 所示，这是当 $1/24\leqslant b<1/12$，$K=3$ 时中央政府的决策。当 $0\leqslant r<t_1$ 时，中央政府选择 $y=1/2t_1$；当 $t_1\leqslant r<t_2$ 时，中央政府选择 $y=1/2(t_1+t_2)$；当 $t_2\leqslant r\leqslant 1$ 时，中央政府选择 $y=1/2\ (t_2+1)$。

我们可以很容易证明，廉价磋商下的中央政府收益高于中央集权下的收益。当 K 越大时，区间的划

分程度越高，地方政府与中央政府之间的信息传递也越精确，此时中央政府在廉价磋商下的收益更高于中央集权下的收益。

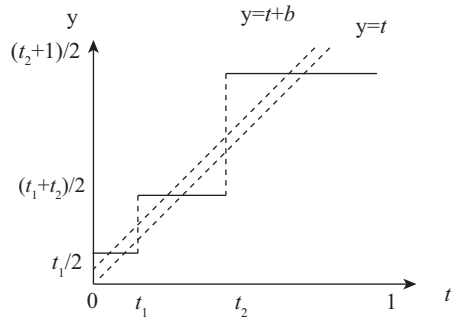

图 1　廉价磋商的解（分三段的情况）

3. 完全承诺能力：财政分权模型。

对于中央政府来说，在单期下有更好的决策机制吗？现在我们考虑财政分权模型：中央政府给予地方政府完全的自由选择权，不需要向上报告，不需要隐瞒，此时地方政府将选择对自身来说最优的决策，即当真实状态为 t 时，其决策为 $y=t+b$，而中央政府的收益始终为 $-b^2$。

我们可以将这一结果与上述的中央政府集权和廉价磋商模型的结果进行对比，我们可以先考虑最简单的两阶段廉价磋商模型，具体来说，图 2 描绘了当 $1/12\leqslant b<1/4$ 时廉价磋商和财政分权两种模型下中央政府的收益。接近横轴的水平直线为 $-b^2$，表示财政分权模型中每种状态下中央政府的收益，扇形曲线表示廉价磋商模型中每种状态下中央政府的收益。由于真实状态 t 均匀分布于［0，1］区间，因此财政分权模型下中央政府的期望收益为横轴与水平直线之间矩形面积的负数，廉价磋商模型下中央政府的期望收益为横轴与抛物线之间面积的负数。

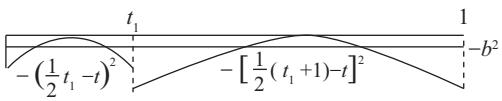

图 2　廉价磋商下中央政府的期望收益与财政分权下中央政府的期望收益

两阶段廉价磋商模型下中央政府的期望收益为 $\int_0^{\frac{1}{2}-2b}-\left[\frac{1}{2}\left(\frac{1}{2}-2b\right)-t\right]^2\mathrm{d}t+\int_{\frac{1}{2}-2b}^{1}-\left[\frac{1}{2}\left(\frac{3}{2}-2b\right)-t\right]^2\mathrm{d}t=-b^2-\frac{1}{48}$，财政分权模型下中央政府的期望收益为 $-b^2$，显然财政分权下的结果更优。实际上，Dessein（2002）[14] 证明了对于任意的 b，决策权下放给代理人都要好于廉价磋商博弈下的均衡解。

因此，当中央政府与地方政府的偏好差异不是太大时，中央政府将财政决策权交由地方政府，允许地方政府自主决定其财政支出，将有效地利用本地的信息，此时我们发现对于中央政府来说，财政分权模型总是优于中央集权模型和廉价磋商模型，从而证明了财政分权体制的合理性和优越性。

以上我们考虑的是完全财政分权下的情况，也就是说不管地方真实信息如何，中央政府总是将决策权完全下放给地方政府。然而实际上，我们发现，中央政府可以采取有条件的财政分权，在满足对地方政府的激励相容的情况下最大化自身的效用。这一有条件的财政分权机制就是静态最优代理机制：如果 $t\leqslant 1-2b$，则 $y_1=t+b$；如果 $1-2b\leqslant t\leqslant 1$，则 $y_1=1-b$。证明如下：

地方政府的真实信息为 t，向中央政府传递的信息为 r，中央政府根据传递的信息做出的决策为 $y(r)$。我们知道中央政府的目标为最大化自身的收益，即 $\mathrm{Max}_{y(r)}-[y(r)-t]^2$，地方政府的目标也是最大化自身的收益，即 $\mathrm{Max}_r-[y(r)-t-b]^2$。此时对于地方政府来说，说实话的策略是激励相容的，因此激励相容的一阶条件为 $2[y(r)-t-b]\cdot y'(r)|_{r=t}=0$，具体来说有 $[y(r)-t-b]|_{r=t}=0$ 或 $y'(r)|_{r=t}=0$。由此我们可以得到中央政府的最优机制必须满足的条件为 $y(t)=t+b$ 或 $y(t)=C$，也就是说中央政府要么任由地方政府选择其自身的最优水平，要么采取一个常数值。

在满足条件的机制中，我们发现最优的代理机制结构如图 3 所示：粗线条表示中央政府的最优机制设计，具体来说，当地方政府的最优水平足够低时，即 $t\leqslant 1-2b$ 时，中央政府将下放决策权，任由地方政府选择其最优水平；反之，当地方政府的最优水平超过一定的数值时，即 $t>1-2b$，中央政府将收回决策权，独立做出选择。这一结果可以由简单的代理集合求得，中央政府不需要直接运用显示机制或与地方政府沟通，而是可以向地方政府提供一系列选择并让地方政府在这一范围内自由决策。当地方政府的决策水平未达到中央政府设置的边界时，地方政府拥有完全的自由选择权，可以选择自己的最优水平；当触及中央政府的边界时，地方政府的自由选择权将被剥夺，中央政府将收回这一决策权并做出自身最优的选择。

与之前所述的完全由地方政府采取自身最优水平的财政分权相比，这个最优代理机制反映了利用地方

信息（即地方政府根据真实状态 t 做决策）和控制权丧失（地方政府获取决策权）之间的权衡。当决策权完全掌握在地方政府手中时，地方政府有足够的动机总是采取策略 $y=t+b$，而边界的设置限制了地方政府的自由选择权并减少了控制权的丧失，但也就没办法利用地方政府的私有信息了。显然，b 越小就意味着中央政府和地方政府之间的利益冲突越小，图 3 中粗横线部分就越短，即地方政府的自由选择权越大。

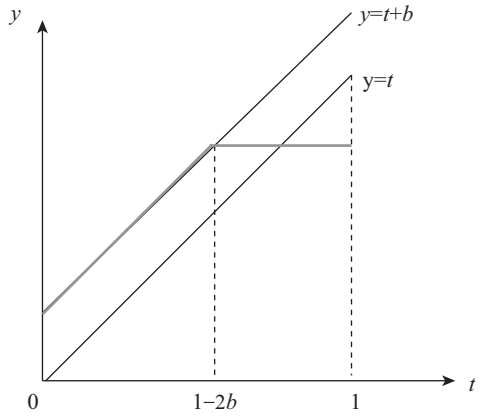

图 3　静态下的最优代理机制

然而，如果我们考虑中央政府与地方政府的两期代理问题，我们会发现，最优静态代理规则将不再适用于这一情形。在两期博弈中，最优静态机制下地方政府有动机在第一期高报其真实信息，从而降低自身第二期的效用损失。

（二）两期情况下的模型

现在我们考虑一个更具现实意义的两期代理模型，在这个模型中，我们仍然假设只有地方政府了解当地的真实信息。跟之前模型的设置一样，我们假设中央政府和地方政府的收益函数不同，且两期下两者的收益函数不变，两者的偏好差异不变，地方的真实状态 t 也不变。

正如之前所述，在两期模型中，我们仍然假设中央政府和地方政府的偏好不同，也就是各自对于当地政府的最优财政支出水平的偏好不同。因为对于地方政府而言，当其自身激励与居民福利相契合时，其最优支出水平就是最大化居民效用的公共产品供给水平；但是中央政府要考虑的不仅是某个区域的利益，而是全国整体的利益，因此中央政府需要统一协调全国各地的公共产品供给，实现资源在全国范围内的最优配置。正是由于两者的角色不同，中央政府与地方政府在当地财政支出水平上必然存在不同的偏好，双方的效用函数存在差异。

在第一期，地方政府了解到地方真实信息 t，中央政府只知 t 值均匀分布在［0，1］区间；在第二期，中央政府将根据地方政府在第一期的决策行为形成对真实信息 t 的推断。此时中央政府的最大化目标为：

$$\mathrm{Max}_{y_1(t),y_2(t)}[-(y_1(t)-t)^2-\delta(y_2(t)-t)^2] \qquad (8)$$

其中：$y_1(t)$ 表示第一期的决策，$y_2(t)$ 表示第二期的决策，有 $y_i \in Y$，其中 Y 是 R 的紧空间；δ 表示对第二期的折现。对于中央政府来说，给定地方真实状态 t，其最优选择为 $y_1(t)=y_2(t)=t$。

地方政府的最大化目标为：

$$\mathrm{Max}_{y_1(t),y_2(t)}\{-[y_1(t)-(t+b)]^2-\delta[y_2(t)-(t+b)]^2\} \qquad (9)$$

从上式我们可以看出，对于地方政府来说，最优决策水平为 $y_1(t)=y_2(t)=t+b$，其中 b 值仍表示中央政府与地方政府之间的偏好偏差。正如之前所述，这种二次型收益函数广泛应用于许多学者的经典文章中，比如 Crawford 和 Sobel（1982）[19]、Holmstrom（1977）[10]，再从 2000 年开始到现在的文献，都是采取的这一经典形式。此外，在本文中，二次型的收益函数形式也符合我们的模型要求，同时也能求出显示解，因此我们依然沿用这一函数形式。

1. 完全无承诺能力和完全承诺下的情况。

由之前的证明可知，在单期的情况下，中央政府完全没有承诺能力的结果就是廉价磋商均衡。在廉价磋商模型中，有 $2bK(K-1)<1$，即地方政府传递信息的精确度（也就是中央政府划分信息区间的个数 K 值）取决于两者的收益偏差 b 值。总的来说，b 值越小，K 值越大，也就是中央政府与地方政府的收益偏差越小，则信息的传递就越精准。

当中央政府有完全的承诺能力时，只要 b 满足一定条件，即 $b<\sqrt{3}/6$，财政分权模型同时优于廉价磋商模型和中央集权模型。也就是说，如果地方政府的偏好与中央政府的偏好较为接近，给予地方政府完全自由选择权对于中央政府来说更有利。此时在单期的代理模型下，财政支出的均衡结果最为接近居民真实状态，当地的资源配置最为有效。

当中央政府两期均没有承诺能力时，两期下的情况实际上就是单期的廉价磋商模型的两期重复。同样地，当中央政府两期均有完全承诺能力时，两期下的情况就是单期的财政分权模型的两期重复。由之前的

证明可知，在单期下，对于中央政府来说，财政分权下的结果总是优于廉价磋商的结果，那么两期下的情形也依然如此。

2. 有限承诺下的情况。

本文的重点就是分析中央政府有限承诺下的情况。与之前模型不同的是，在中央政府只有有限承诺能力的情况下，中央政府在第一期放权，但在第二期收权。也就是说，在第一期时，由地方政府独立决定其财政支出水平；在第二期，中央政府根据地方政府第一期的决策推断地方真实信息并独立做出决策。此时由于地方政府预期到中央政府的这一行为，为了避免中央政府第二期的决策损害其利益，地方政府有动机在第一期选择偏离其最优的支出水平，避免揭露其真实信息。而考虑到地方政府的反应，中央政府将寻求两期下激励相容的最优合同设计。

与传统静态模型不同的是，在本文的有限承诺下的两期代理问题中，中央政府不能在第一期决定 $y_2(t)$，只能在第二期的开始决定。如果地方政府在第一期揭露其真实信息 t，那么中央政府便会在第二期做出决策 $y_2=t$，而预期到中央政府在第二期的行为，即使地方政府在第一期有决策权，也不会做出揭露其真实状态的决策。这个博弈的时间顺序如图 4 所示：在 0 期仅有地方政府获知真实状态 t；在第 1 期，地方政府根据其收益函数做出决策，向中央政府传递信息 r_1；在第 2 期，中央政府根据第 1 期地方政府的行为和贝叶斯法则更新其信念并做出决策。

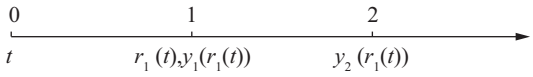

图 4　两期代理下的时序

接下来的定理 1 证明，在两期有限承诺的动态博弈下，静态最优代理机制将不再是激励相容的，因为此时地方政府有动机在第一期报告错误的信息。定理 2 与定理 3 证明存在连续的完全分离均衡可以使得地方政府在第一期揭露其真实信息，并且提升中央政府的效用。但是对于中央政府来说，这一完全分离策略虽然能在完全混同策略的基础上提高自身效用，但并不是最优的。因为在这一策略下，中央政府为了让地方政府在第一期揭露其真实信息所付出的代价太大。

因此，在完全分离策略的基础上，定理 5 证明了存在单侧不连续代理策略能进一步提高中央政府的效用。同时，定理 6 证明区间的右侧也存在一个临界值 $\tilde{t}$，使得中央政府能通过采取两侧不连续代理策略来提高自身效用。当中央政府与地方政府的收益偏差足够小时，最优代理策略中包含以下两种形式：一种是完全混同代理策略，在左右两侧的每个区间上，中央政府两期都采用完全混同策略，也就是说两期的决策水平分别都是一个固定值；另一种是完全分离策略，且存在完全分离策略的区间是连通的。最后定理 8 还证明了最优临界值所需满足的条件，给出了中央政府的最优代理机制。

定理 1：对于任何 $\delta>0$ 的两期代理来说，静态最优代理机制是激励不相容的。

证明：假设最优静态代理机制是可行的，我们取子区间 $t\in[t_1,\ t_2]\subset[0,\ 1-2b]$，此时若地方政府在第一期选择其最优水平 $y_1=t+b$，则中央政府可以通过这一行为了解到真实信息为 t，因此在第二期中央政府的决策为 $y_2=t$，此时地方政府的效用为 $U=-\delta b^2$。现在我们考虑在第一期时地方政府不揭露其真实信息，而是向中央政府报告 $t'=t+\varepsilon$，$\varepsilon>0$ 时的情况。在这种情况下，中央政府的决策为 $y_1=t+b+\varepsilon$，$y_2=t+\varepsilon$，此时地方政府的效用为 $U'=-\varepsilon^2-\delta(b-\varepsilon)^2$。通过简单的计算可知，如果 $\varepsilon<2b\delta/(1+\delta)$，则有 $U'>U$，已知 $\delta>0$，$b>0$，则有 $2b\delta/(1+\delta)>0$，即 $\varepsilon>0$ 存在。也就是说，当 $t\in[0,\ 1-2b]$ 时，静态最优代理机制下地方政府的决策为 $y_1=t+b$，但此时地方政府可以通过在第一期偏离其真实信息来提高收益。也正是因为地方政府不会在第一期揭露其真实信息，我们才认为静态最优代理机制是激励不相容的。

上述结论成立的前提是我们假设地方政府的收益函数 $-(y_1(t)-t-b)^2-\delta(y_2(t)-t-b)^2$ 是严格凸函数。在这种二次函数形式下，通过向中央政府传递信息 $t+\varepsilon$，其一阶效应为正，二阶效应为负，且一阶效应占优二阶效应，因此偏离真实值的信息可以提高地方政府的效用。

现在我们想知道当中央政府了解到在这种两期的有限承诺代理模型下，地方政府在第一期的信息传递是有偏差的，中央政府是否还能通过机制的设计，在第一期知悉地方政府的真实信息，若中央政府能了解到地方政府的真实信息，那么在第二期就能采取最优决策 $y_2(t)=t$。在接下来的定理中我们给出了中央政府能在第一期知悉地方政府真实信息的代理机制，且这一代理机制在整个信息区间是连续的。

定理 2：激励相容的全分离连续代理机制为：$y_1(t)=t+b+\delta b$，$y_2(t)=t$。

证明：如果地方政府在第一期的决策 $y_1=y_1(t)$

揭露了真实信息，则中央政府第二期的决策为 $y_2(t)=t$。一个激励相容的代理机制应该满足以下条件：对于地方政府来说，任意偏离真实值的 t' 带来的收益都小于传递真实值 t 带来的收益，即：

$$-[y_1(t')-t-b]^2-\delta(t'-t-b)^2\leqslant-[y_1(t)-t-b]^2-\delta b^2 \tag{10}$$

则我们有：$t=\operatorname{argmax}_{t'}-[y_1(t')-t-b]^2-\delta(t'-t-b)^2$。

通过对 t' 求导可得：$-[2(y_1(t')-t-b)\dot{y}_1(t')+2\delta(t'-t-b)]_{t'=t}=0$。

代入 $t'=t$ 可得：$[y_1(t)-t-b]\dot{y}_1(t)-\delta b=0$。

为了解这个微分方程，我们可以令 $s(t)=y_1(t)-t-b$，由此可得 $s(t)[\dot{s}(t)+1]=\delta b$，也可写为 $sds+(s-\delta b)\mathrm{d}t=0$。

如果 $s-\delta b\neq0$，则有 $[1+\delta b/(s-\delta b)]\mathrm{d}s+\mathrm{d}t=0$，该方程的一般解为 $s+\delta b\ln|s-\delta b|+t=C$，其中 C 为常数。因为 $s(t)=y_1(t)-t-b$，所以此解也可以用 $y_1(t)$ 表示为 $y_1(t)+\delta b\ln|y_1(t)-\delta b|-b=C$。

如果 $s-\delta b=0$，则 s 为常数，此时方程有特殊解为 $y_1(t)=t+b+\delta b$。因为此解是一个闭合解且更容易执行，所以我们主要关注这个特殊解。

由此我们得到了一个第一期的均衡结果，即 $y_1(t)=t+b+\delta b$，当中央政府设计并执行这一机制时，地方政府有动机在第一期揭露其真实信息。此时对于地方政府而言，一阶效应消失，只剩下负的二阶效应，因此传递偏离真实值的信息不再有利可图且不再是激励相容的。同时，上述定理告诉我们如果存在使中央政府能知悉地方政府真实信息 t 的子区间，在这个区间内，有 $y_1(t)=t+b+\delta b$。

我们将这一均衡结果称为完全分离均衡，因为中央政府可以通过这一机制知悉地方政府的真实信息，将真实信息进行完全分离识别。在单期代理模型中，静态最优代理机制下中央政府无法识别地方的真实信息，只能为地方政府的决策设置一个上限，避免地方政府的决策水平过于偏离中央政府的最优水平。在两期代理模型中，我们可以找到这样一个完全分离均衡，使得地方政府有足够的动机揭露其真实信息，从而让中央政府在第二期能采取最优决策。当然，在两期代理下，中央政府也可以选择在第一期采取完全混同，在第二期采取静态最优代理，在之后的定理中，我们将比较这种策略与完全分离策略下中央政府的收益，从而明确完全分离均衡是否能提升中央政府的效用。

Laffont 和 Tirole（1988）[20] 认为在有限承诺的两期委托代理模型中，任何一个区间都不存在完全分离均衡，在他们的文章中，如果高类型的代理人在第一期故意传递一个向下偏离 $\varepsilon\in(0,1)$ 的信息，那么他只需在第一期承受一个二阶损失 ε^2，若委托人认定他为低类型的人，那么就会为他在第二期带来一个一阶收益 ε，因此高类型的代理人可以传递错误的自身类型信息，从而获得一个正的租金。然而对于低类型的代理人来说，这个正的租金是不存在的，因此低类型的代理人将会在第一期谎称自己是高类型，并在第二期结束与委托人的关系，这就是“拿了钱就跑”（take-the-money-and-run）的策略。在 Laffont 和 Tirole 的文章中，参与约束允许这种“拿了钱就跑”的策略的实现，但是在本文中，考虑到中央政府对地方政府的强力约束，我们认为地方政府很难在第二期“拿了钱就跑”，所以没有加入这个参与约束，因此这种策略是不可行的。

到目前为止，我们已经求出了一个完全分离均衡 $y_1(t)=t+b+\delta b$，$y_2(t)=t$。当中央政府采取这一策略的时候，地方政府有动机在第一期揭露其真实信息。但是我们也发现，在完全分离均衡下，中央政府在第一期的决策水平较大地偏离了其最优水平，也就是说，中央政府为了在第二期做出最优决策，在第一期要付出较大的成本。事实上，我们可以找到在完全分离均衡基础上进一步提升中央政府效用的代理机制，本文将在之后的定理中具体分析这一改进方式。

由之前的定理结论可知除了上述的完全分离均衡之外，中央政府也可以采取完全混同策略：在第一期选择混同策略 $y_1=1/2$；在第二期选择静态最优代理机制（当 $t\leqslant1-2b$ 时，则 $y_2=t+b$；当 $1-2b\leqslant t\leqslant1$ 时，则 $y_2=1-b$）。接下来我们将给出第一期时完全分离均衡优于完全混同均衡的条件。

定理 3：如果 $(1+\delta+\delta^2)b^2+4\delta b^2/3<1/12$，则在第一期时有完全分离均衡优于完全混同均衡。

证明：在完全分离均衡中，中央政府在两期的决策为 $y_1(t)=t+b+\delta b$，$y_2(t)=t$，此时中央政府的效用为：

$$U_S=-\int_0^1(t+b+\delta b-t)^2\mathrm{d}t=-(b+\delta b)^2 \tag{11}$$

在完全混同均衡中，中央政府在两期的决策分别为 $y_1(t)=1/2$，如果 $t\leqslant1-2b$，则 $y_2=t+b$；如果

$1-2b\leqslant t\leqslant 1$，则 $y_2=1-b$。此时中央政府的效用为：

$$U_P=-\int_0^1\left(\frac{1}{2}-t\right)^2\mathrm{d}t-\delta\left[\int_0^{1-2b}(t+b-t)^2\mathrm{d}t+\int_{1-2b}^1(1-b-t)^2\mathrm{d}t\right]=-\frac{1}{12}-\delta\left(b^2-\frac{4b^3}{3}\right)\tag{12}$$

因此，如果 $(1+\delta+\delta^2)b^2+4\delta b^3/3<1/12$，并将这一临界值定为 $\hat{b}$，则有 $U_S>U_P$，也就是说完全分离均衡优于完全混同均衡，证毕。

从上述结果可知，当中央政府与地方政府的收益偏差 b 足够小的时候，完全分离均衡是占优的。但是与静态代理或 $\delta=0$ 的情况相比，在两期动态代理模型中，要使完全分离均衡优于完全混同均衡，所需的临界值 $\hat{b}$ 更小。因为两期模型下第一期的 $y_1(t)=t+b+\delta b$ 对中央政府的最优水平 $y=t$ 的偏离程度越高，也就是说中央政府在第一期将信息进行分离的代价更高，因此只有当中央政府与地方政府的收益偏差更小时，才能使得完全分离策略占优。此外，我们可以发现，折现因子 δ 越大，完全分离均衡优于完全混同均衡所需的临界值 $\hat{b}$ 也越小，因为 δ 越大，表示地方政府越在乎第二期的效用，就更不愿意在第一期披露其真实信息，所以只有 $\hat{b}$ 更小才能使中央政府在第一期采取完全分离策略是可行的。

这个定理给出了在第一期完全分离策略优于完全混同策略的条件，然而在两期代理模型中，中央政府除了可以采取上述的完全分离策略和完全混同策略，还可以采取两区间代理策略。两区间代理策略是指中央政府在第一期将真实信息 t 划分为两个区间，分别在两个区间上采取完全混同策略。接下来我们可以证明，当 b 满足一定条件时，中央政府可以通过将第一期的信息类型划分成两个区间来提高自身效用。下面的定理证明了在两期代理模型下，两区间均衡优于完全混同均衡却仍被占优于完全分离均衡的条件。

定理 4：两区间代理策略优于完全混同策略却仍被占优于完全分离策略的充分条件为：b 值满足 $(1+\delta+\delta^2)b^2+8/3\delta b^3<1/48$。

证明：已知完全分离均衡下中央政府的效用为 $U_S=-(b+\delta b)^2$，现在我们考虑在第一期的两区间代理策略，此时我们在求解时不考虑地方政府的的激励相容约束条件，因为以此求得的完全分离均衡优于两区间均衡的 b 值的范围更小。如果我们将地方政府的激励相容约束条件纳入考虑，b 值的范围将会更大。两区间策略是指中央政府在第一期将真实信息 t 划分为两个区间，分别在两个区间上采取完全混同策略。在这一策略下，我们可以证明对于中央政府来说，均匀划分区间是最优的，因此取 t 的中间值，定义为 $\bar{t}=1/2$，并有 $y_1(\bar{t}_-)=1/4$ 以及 $y_1(\bar{t}_+)=3/4$。此时两区间策略为：当 $t\in[0,1/2]$ 时，有 $y_1=1/4$；当 $0\leqslant t\leqslant 1/2-2b$ 时，有 $y_2=t+b$；当 $1/2-2b\leqslant t\leqslant 1/2$ 时，有 $y_2=1-b$；当 $t\in[1/2,1]$ 时，有 $y_1=3/4$；当 $1/2\leqslant t\leqslant 1-2b$ 时，有 $y_2=t+b$；当 $1-2b\leqslant t\leqslant 1$ 时，有 $y_2=1-b$。因此两区间策略下中央政府的效用为：

$$U_t=-\int_0^{\frac{1}{2}}\left(\frac{1}{4}-t\right)^2\mathrm{d}t-\int_{\frac{1}{2}}^1\left(\frac{3}{4}-t\right)^2\mathrm{d}t-\delta\left[\int_0^{\frac{1}{2}-2b}b^2\mathrm{d}t+\int_{\frac{1}{2}-2b}^{\frac{1}{2}}\left(\frac{1}{2}-b-t\right)^2\mathrm{d}t+\int_{\frac{1}{2}}^{1-2b}b^2\mathrm{d}t+\int_{1-2b}^1\left(\frac{1}{2}-b-t\right)^2\mathrm{d}t\right]=-\left[\frac{1}{48}+\delta b^2\left(1-\frac{8}{3}b\right)\right]\tag{13}$$

令 $U_S>U_t$，可得 $(1+\delta+\delta^2)b^2+8/3\delta b^3<1/48$，此时我们可以将这一临界值定义为 b'，当 $b<b'$ 时，完全分离均衡将占优于两区间均衡。从之前的定理我们可知：如果 $(1+\delta+\delta^2)b^2+4\delta b^3/3<1/12$，则有 $U_S>U_P$，且由上述定理可知，该临界值为 $\hat{b}$，当 $b<\hat{b}$ 时，完全分离均衡占优于完全混同均衡。我们也可以证明同样在这个条件下，两区间均衡优于完全混同均衡。已知 $U_P=-1/12-\delta(b^2-4/3b^3)$，则有 $U_t-U_P=1/16-8/3\delta b^3$，当 $(1+\delta+\delta^2)b^2+4\delta b^3/3<1/12$ 时，肯定有 $8/3\delta b^3<1/24$，即 $U_t-U_P>0$。也就是说当 $b<\hat{b}$ 时，完全分离均衡和两区间均衡均占优于完全混同均衡。同时从上面的式子中我们可以得到 $b'<\hat{b}$，所以当 b 取更小值，即 $b<b'$ 时，完全分离均衡占优于两区间均衡。由此我们可以得出当 $(1+\delta+\delta^2)b^2+8/3\delta b^3<1/48$ 时，两区间均衡占优于完全混同均衡却仍被占优于完全分离均衡，证毕。

由上述结果可知，当 b 值足够小时，对于中央政府来说，完全分离策略占优于完全混同策略和两区间策略。并且随着 b 值越来越小，在第一期可划分的区间数量就更多，多区间代理策略的效果就越来越好。那么，当 b 值足够小时，会使多区间策略趋近于完全分离策略吗？然而即便如此，我们也可以证明对于任意的 $b>0$，对于中央政府来说，在整个信息区间都采

取完全分离策略并不是最优的选择。中央政府总是可以通过以下方式来提高自身的效用：存在一个临界值 $\hat{t}$，当 t 低于 $\hat{t}$ 时，两期均采取混同策略；当 t 高于 $\hat{t}$ 时，采取完全分离策略；当 $t=\hat{t}$ 时，完全混同策略和完全分离策略带给地方政府的效用是相同的。也就是说，中央政府对于低类型的信息采取混同策略，对于高类型的信息采取分离策略。

对这一策略的直观解释是：对于中央政府来说，第一期的信息分离之所以能提高效用，是因为能使其在第二期做出最优决策，但是这一分离的成本是第一期的决策极大地偏离最优水平，且这一偏离程度越高，成本就越高。通过混同低类型的信息，中央政府可以降低第一期决策的偏离程度。因为在完全分离策略中，第一期的决策水平为 $y_1=t+b+\delta b$，当 t 足够小时，对于中央政府来说，在第一期采取比 $y_1=t+b+\delta b$ 更接近其最优水平 $y=t$ 的策略 $y_1(t)=y_1$，此时提升的第一期的效用足以弥补第二期混同策略下的效用损失。

现在我们考虑一个在整个信息区间上策略不连续的两期代理模型，跟之前一样，我们假设这个断点为 $\hat{t}$，两期代理策略为：当 $t\in[0,\hat{t}]$ 时，有 $y_1(t)=y_1$ 以及 $y_2(\hat{t})=\hat{t}/2$；当 $t\in[\hat{t},1]$ 时，有 $y_1(t)=t+b+\delta b$，$y_2(t)=t$。令 $x=\hat{t}+b-y_1$。为了使该模型激励相容，$\hat{t}$ 必须在左右两个区间上的策略带给地方政府的效用是相同的，也就是说，当 $t=\hat{t}$ 时，左侧的完全混同策略和右侧的完全分离策略带给地方政府的效用是相同的。通过对条件的整理，我们可得 $\hat{t}$ 满足以下条件：$x^2+\delta(\hat{t}/2+b)^2=(\delta b)^2+\delta b^2$。

定理 5：总是存在激励相容的不连续策略——在临界值 $\hat{t}$ 的左侧混同，在临界值 $\hat{t}$ 的右侧分离——占优于完全分离策略。

证明：上述不连续模型的策略为：存在一个临界值 $\hat{t}$，当 t 低于 $\hat{t}$ 时，中央政府两期均选择混同策略，即 $y_1(t)=y_1$，$y_2(\hat{t})=\hat{t}/2$；当 t 高于 $\hat{t}$ 时，中央政府将选择完全分离均衡，即 $y_1(t)=t+b+\delta b$，$y_2(t)=t$。

在我们的模型中，我们假设在上述临界值的左侧，中央政府在两期均采用完全混同策略。事实上，如果中央政府在第二期采取静态最优代理策略，可以进一步提高自身效用，但也会使得我们的计算更加复杂。本文重点在于说明存在不连续的代理策略优于完全分离策略，即使在左侧只是最简单的两期完全混同策略也能优于完全分离策略，那么更复杂形式下的不连续代理策略就更占优了。

首先我们需要先证明这个代理策略是激励相容的。我们可知任何 $0\leqslant t<\hat{t}$ 不会模仿 $0\leqslant t'<\hat{t}$，因为此时中央政府采取的是完全混同策略，无论地方政府传递该区间内的任何一个信息，中央政府都始终做出决策 $y_1(t)=y_1$。同时，任何 $\hat{t}<t\leqslant 1$ 不会模仿 $\hat{t}<t'\leqslant 1$，因为正如定理 3 证明的，此时中央政府采取完全分离策略，对于地方政府来说，说实话的策略是激励相容的，地方政府有动机在第一期揭露其真实信息。

因此，我们只需要证明（1）任何 $0\leqslant t<\hat{t}$ 不会模仿 $\hat{t}<t'\leqslant 1$，（2）任何 $\hat{t}<t\leqslant 1$ 不会模仿 $0\leqslant t'<\hat{t}$。证明如下：

（1）证明任何 $0\leqslant t<\hat{t}$ 不会模仿 $\hat{t}<t'\leqslant 1$。令 $U_t(t')$ 表示当地方政府的真实信息为 t 却传递信息 t' 时地方政府的效用，有 $U_t(t')=-(t'+b+\delta b-t-b)^2-\delta(t'-t-b)^2$，对 t' 求导可得 $U_t'(t')=-2(1+\delta)(t'-t)<0$，因为 $t'>t$，因此 $U_t(t')$ 是关于 t' 的减函数。此时若地方政府的真实信息为 $0\leqslant t<\hat{t}$ 却传递信息 $\hat{t}<t'\leqslant 1$，则他将传递信息 $\hat{t}_+$。现在我们要证明对于任意的 $0\leqslant t<\hat{t}$，有 $U_t(t)>U_t(\hat{t}_+)$。令 $\varepsilon=\hat{t}-t$，如上所述，$x=\hat{t}+b-y_1$，当 $\varepsilon<x$ 时，有 $U_t(t)=-(x-\varepsilon)^2-\delta(\hat{t}/2+b-\varepsilon)^2$，$U_t(\hat{t}_+)=-(\delta b+\varepsilon)^2-\delta(b-\varepsilon)^2$；当 $\varepsilon>x$ 时，有 $U_t(t)=-(\varepsilon-x)^2-\delta(\hat{t}/2+b-\varepsilon)^2$，$U_t(\hat{t}_+)=-(\delta b+\varepsilon)^2-\delta(b-\varepsilon)^2$。因此，在上述两种情况下，我们均可以得到 $U_t(t)-U_t(\hat{t}_+)=(\delta b+\varepsilon)^2+\delta(b-\varepsilon)^2-(\varepsilon-x)^2-\delta(\hat{t}/2+b-\varepsilon)^2$，从之前的定理已知临界值 $\hat{t}$ 满足条件 $x^2+\delta(\hat{t}/2+b)^2=(\delta b)^2+\delta b^2$，因此可以整理得到 $U_t(t)-U_t(\hat{t}_+)=2x\varepsilon+2\delta\varepsilon(\hat{t}/2+b)>0$。由此我们证明了 $U_t(t)>U_t(\hat{t}_+)$，也就是说当地方政府的真实信息为 $0\leqslant t<\hat{t}$ 时，他没有动机传递错误的信息 $\hat{t}<t'\leqslant 1$。

（2）证明任何 $\hat{t}<t\leqslant 1$ 不会模仿 $0\leqslant t'<\hat{t}$。此时我们有 $U_t(t)=(\delta b)^2-\delta b^2=-x^2-\delta(\hat{t}/2+b)^2$，$U_t(t')=-(x+t-\hat{t})^2-\delta(\hat{t}/2+b+t-\hat{t})^2$，因为此时 $t>\hat{t}$，我们可以很直接地得到 $U_t(t)>U_t(t')$，也就是说当地方政府的真实信息为 $\hat{t}<t\leqslant 1$ 时，他没有动机传递错误的信息 $0<t'\leqslant\hat{t}$。

接下来我们需要证明这种不连续代理均衡占优于完全分离均衡，也就是中央政府可以通过采取这种策略来提高自身的效用。令 $\eta=\hat{t}+b+\delta b-y_1$，之前 $\hat{t}$ 激励相容的约束条件可以重新整理为 $(\delta b)^2+\delta b^2=(\delta b-\eta)^2+\delta(b+\hat{t}/2)^2$。在完全分离均衡下，中央政府的效用为 $U_S=-\int_0^1(t+b+\delta b-t)^2\mathrm{d}t$，在不连续代理均衡中，中央政府的效用为：

$$U_D=-\int_0^{\hat{t}}(y_1-t)^2\mathrm{d}t-\delta\int_0^{\hat{t}}\left(\frac{\hat{t}}{2}-t\right)^2\mathrm{d}t-\int_{\hat{t}}^1(t+b+\delta b-t)^2\mathrm{d}t \tag{14}$$

因此有：

$$U_D-U_S=\int_0^{\hat{t}}(b+\delta b)^2\mathrm{d}t-\int_0^{\hat{t}}(y_1-t)^2\mathrm{d}t-\delta\int_0^{\hat{t}}\left(\frac{\hat{t}}{2}-t\right)^2\mathrm{d}t=(b+\delta b)^2\hat{t}-y_1^2\hat{t}+y_1\hat{t}^2-\frac{1}{3}\left(1+\frac{\delta}{4}\right)\hat{t}^3 \tag{15}$$

通过整理，我们可以将 $U_D-U_S>0$ 写成如下形式：

$$3y_1^2-3\hat{t}y_1+\left(1+\frac{\delta}{4}\right)\hat{t}^2<3(b+\delta b)^2 \tag{16}$$

由之前的定理可知，$\hat{t}$ 激励相容的约束条件为：

$$(\delta b)^2+\delta b^2=(\hat{t}+b-y_1)^2+\delta\left(b+\frac{\hat{t}}{2}\right)^2 \tag{17}$$

如果我们能找到 $\hat{t}$ 和 y_1 同时满足式（16）和式（17），我们就可以说该不连续代理策略占优于完全分离策略。事实上，我们可以很容易地找到符合条件的 $\hat{t}$ 和 y_1，举例来说，我们可以选择 $y_1=b+\delta b$，将其代入式（17）得到 $\hat{t}=4\delta b/(4+\delta)$，将其代入式（16）可得$\hat{t}<12(b+\delta b)/(4+\delta)$，由此可见 $\hat{t}=4\delta b/4+\delta$ 且 $\hat{t}<2b$ 时，同时满足上述两个条件。因此，我们发现总是存在 $\hat{t}$ 和 y_1 使得不连续代理策略优于完全分离策略，证毕。

对于上述单侧不连续代理策略占优于完全分离策略，我们有较为直观的解释。在完全分离策略下，中央政府为了让地方政府揭露其真实信息，第一期的决策水平过于偏离其最优水平，由此导致了过多的效用损失。而在单侧不连续代理策略下，中央政府在左侧信息区间内采取两期的完全混同策略，虽然这将导致第二期的效用损失，但也将大量减少第一期的效用损失，最终仍然能提升中央政府的效用。

此外，我们可以证明在这一临界值上，中央政府的这两个策略在第一期的决策水平是不一样的。如图 6 所示，左侧的粗点线表示混同策略下的决策水平，右侧的粗点线表示完全分离策略下的决策水平，细线表示第一期的决策，粗线表示第二期的决策。已知地方政府两期内最优的决策水平均为$y=t+b$。假设临界值的右极限为 $\hat{t}_+$，此时中央政府采取完全分离策略，根据之前的定理结论可知中央政府的决策为 $y_1(\hat{t}_+)=\hat{t}_++b+\delta b$，$y_2(\hat{t}_+)=\hat{t}_+$。同理，假设临界值的左极限为 $\hat{t}_-$，此时中央政府在两期均采取完全混同策略，第一期的决策为 $y_1(\hat{t}_-)$，第二期的决策为$y_2(\hat{t}_-)=\hat{t}_-/2$。如上所述，当 $t=\hat{t}$ 时，两个策略下地方政府的效用是相同的，从图 5 我们可以看出，有$\hat{t}_-/2<\hat{t}_+<\hat{t}+b$，因此为了使地方政府报告 $\hat{t}_-$和 $\hat{t}_+$无差异，必须满足条件 $y_1(\hat{t}_-)<\hat{t}_++b+\delta b$，由此推出$y_1(\cdot)$在 $\hat{t}$ 上不连续。

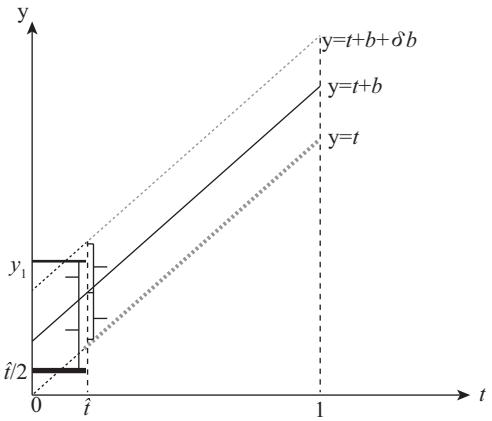

图 5　左侧不连续下的两期代理机制

我们同样可以证明在最优情况下，完全混同均衡也可以出现在信息区间的右侧。如果存在信息区间右侧的混同策略，我们假设这一临界值为 $\tilde{t}$，在临界值的左侧，中央政府采取完全分离策略，则有 $y_1(\tilde{t}_-)=\tilde{t}_-+b+\delta b$，$y_2(\tilde{t}_-)=\tilde{t}_-$；在临界值的右侧，中央政府采取完全混同策略，则有 $y_1(\tilde{t}_+)=y_1'$，$y_2(\tilde{t}_+)=(\tilde{t}_++1)/2$。此时地方政府在第二期可以通过传递信息 $\tilde{t}_+$来提高自身效用，为了使地方政府报告 $\tilde{t}_-$和 $\tilde{t}_+$无差异，中央政府在第一期的决策必须满足条件 $y_1(\tilde{t}_+)>\tilde{t}_-+b+\delta b$。之后我们可以证明这种不连续的代理模型占优于完全分离模型。

如果中央政府要在第一期采取混同策略，那么就必须接受第二期的效用损失，不管混同均衡发生在信息区间左侧还是区间右侧，第二期的效用损失都是不可避免的。但是对于地方政府来说，左右两侧第一期的混同均衡导致的效用收益却是不一样的。如图 6 所示，当混同均衡发生在临界值 $\hat{t}$ 的左侧时，有 $\hat{t}/2<\hat{t}<\hat{t}+b$。也就是说，在第二期，混同均衡下的决策比分离均衡的决策更偏离地方政府的最优水平，而当 $t=\hat{t}$ 时，为了使地方政府报告 $\hat{t}_+$和 $\hat{t}_-$无差异，中央政府在第一期的决策必须满足条件 $y_1(\hat{t}_-)<\hat{t}_-+b+\delta b$。当混同均衡发生在临界值的右侧时，有 $\tilde{t}<(\tilde{t}+1)/2<\tilde{t}+b$，分离均衡的第二期决策比混同均衡的第二期决策更偏离地方政府的最优水平，为了使地方政府的选择无差异，

中央政府在第一期的决策必须满足条件 $y_1(\tilde{t}_+)>\tilde{t}_++b+\delta b$。

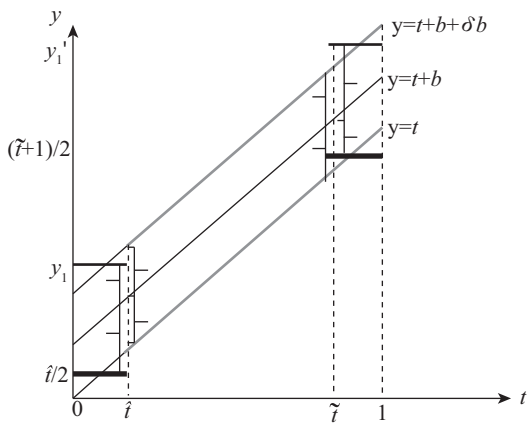

图 6　两侧不连续下的代理机制

接下来，我们将证明总是存在 $\tilde{t}$，当 $t\in[\tilde{t},\ 1]$ 时，中央政府采取完全混同策略：$y_1(t)=y_1'$，$y_2(t)=(1+\tilde{t})/2$；当 $t\in[\hat{t},\ \tilde{t}]$ 时，中央政府采取完全分离策略：$y_1(t)=t+b+\delta b$，$y_2(t)=t$。我们可以证明相比上述定理中的单侧不连续策略，当中央政府在区间右侧也采取这种不连续策略时，可以进一步提高自身效用。

定理 6：总是存在一种激励相容的两侧不连续代理策略——在 $[0,\ \hat{t}]$ 和 $[\tilde{t},\ 1]$ 上采取完全混同策略，在 $[\hat{t},\ \tilde{t}]$ 上采取分离策略——占优于上述单侧不连续代理策略。

证明：和定理 5 的证明类似，在此省略。

对于上述双侧不连续代理策略占优于单侧不连续代理策略，我们有较为直观的解释。在单侧不连续代理策略下，中央政府仅在左侧低类型信息区间上采取两期的完全混同策略，但在右侧的高类型信息区间上，仍然存在进一步的改进空间。在右侧的高类型信息区间上，为了让地方政府揭露其真实信息，第一期的决策水平过于偏离其最优水平，由此导致了过多的效用损失。而在双侧不连续代理策略下，中央政府在右侧信息区间内采取两期的完全混同策略，虽然这将导致第二期的效用损失，但也将大量减少第一期的效用损失，最终仍然能提升中央政府的效用。

在上述证明的左右两侧临界点之间，我们也可以怀疑存在另外的连续区间，在该区间内，完全混同策略占优于完全分离策略。但是不同于之前对 $\hat{t}$ 和 $\tilde{t}$ 的证明，此时我们需要证明在这一连续区间内，这一区间的起点和终点都满足激励相容约束条件，接下来的定理证明了这种连续区间并不存在。

定理 7：最多存在一个连续区间 $[t_1,\ t_2]$，在该区间内完全分离策略是最优的，其中 $t_1>0$，$t_2<1$，并且，从中央政府的角度来说，该区间内决策水平 y 是关于 t 的非减函数。

证明：我们可以先用之前已证明过的代理策略作为基础模型。从 $\hat{t}$ 开始，我们希望得到满足以下条件的另一个区间：在该区间内中央政府在两期均采取完全混同策略，在该区间之后中央政府采取完全分离策略。将这个区间的长度定义为 y_2，因此中央政府在该区间内的第二期决策为 $\hat{t}+y_2/2$。该区间内的第一期决策为 $\hat{t}+b+\delta b+x$，x 由以下激励相容的约束条件决定：$(\delta b)^2+\delta b^2=\delta\ (b-y_2/2)^2+(\delta b+x)^2=\delta\ (b+y_2/2)^2+(\delta b-y_2+x)^2$，其中第一个等式是在该区间的起点 $\hat{t}$ 上的激励相容约束条件，第二个等式是在该区间的终点上的激励相容约束条件。从这两个等式中，我们可以得到 $x=y_2=0$，也就是说满足条件的区间是不存在的。

如果我们只满足起点上的激励相容约束条件，也就是只满足第一个等式时，可以得到 $(\delta b)^2+\delta b^2=\delta(b-y_2/2)^2+(\delta b+x)^2<\delta(b+y_2/2)^2+(\delta b-y_2+x)^2$，显然无法兼容第二个激励相容约束条件。如果我们只满足终点上的激励相容约束条件，也就是只满足第二个等式时，可以得到 $(\delta b)^2+\delta b^2=\delta(b+y_2/2)^2+(\delta b-y_2+x)^2>\delta(b-y_2/2)^2+(\delta b+x)^2$，也无法兼容第一个激励相容约束条件。由此我们证明了最多存在一个连续区间，也就是上述定理给出的 $[\hat{t},\ \tilde{t}]$，只有在这个区间内，完全分离策略是最优的。我们之前只求出了完全分离策略 $s-\delta b=0$ 时的特殊解，并未考虑 $s-\delta b\neq0$ 的情况，但是我们能确定的是，在两种情况下，都有决策水平 y 是关于 t 的非减函数。

此时，我们可以大致描述一下最优的代理策略。当中央政府与地方政府的收益偏差 b 足够小时，最优代理策略中包含以下两种形式：一种是完全混同代理策略，在左右两侧的区间上，中央政府都采用完全混同策略；另一种是完全分离策略，且存在完全分离策略的区间是连通的。接下来，我们将给出在单侧不连续代理策略中满足条件的最优的 $\hat{t}$ 和 y_1。

定理 8：在不连续代理策略中，存在最优的 $\hat{t}$ 和 y_1，两者同时满足以下条件：

$$\frac{(b+\delta b)^2-\frac{\delta\hat{t}^2}{4}-(\hat{t}^2+y_1^2-2\hat{t}y_1)}{\hat{t}^2-2\hat{t}y_1}=\frac{2(\hat{t}+b-y_1)+\delta\left(\frac{\hat{t}}{2}+b\right)}{-2(\hat{t}+b-y_1)},$$

$(\hat{t}+b-y_1)^2+\delta\left(\frac{\hat{t}}{2}+b\right)^2=(\delta b)^2+\delta b^2$。

证明：这种不连续模型策略为：存在一个临界值 $\hat{t}$，当 t 低于 $\hat{t}$ 时，中央政府两期均选择混同策略，即 $y_1(t)=y_1$，$y_2(\hat{t})=\hat{t}/2$；当 t 高于 $\hat{t}$ 时，中央政府将选择完全分离均衡，即 $y_1(t)=t+b+\delta b$，$y_2(t)=t$。

此时中央政府的期望效应为：

$$U_D=-\int_0^{\hat{t}}(y_1-t)^2\mathrm{d}t-\delta\int_0^{\hat{t}}\left(\frac{\hat{t}}{2}-t\right)^2\mathrm{d}t-\int_{\hat{t}}^1(t+b+\delta b-t)^2\mathrm{d}t$$

因此中央政府的目标为：

$$\mathrm{Max}_{\hat{t},y_1}\left[-\int_0^{\hat{t}}(y_1-t)^2\mathrm{d}t-\delta\int_0^{\hat{t}}\left(\frac{\hat{t}}{2}-t\right)^2\mathrm{d}t-\int_{\hat{t}}^1(b+\delta b)^2\mathrm{d}t\right]$$

$$\text{s. t.}\quad(\hat{t}+b-y_1)^2+\delta\left(\frac{\hat{t}}{2}+b\right)^2=(\delta b)^2+\delta b^2$$

上式可简化为：

$$\mathrm{Max}_{\hat{t},y_1}\left[-\frac{\hat{t}^3+3\hat{t}y_1-3\hat{t}^2y_1}{3}-\frac{\delta\hat{t}^3}{12}+(b+\delta b)^2\hat{t}\right]$$

$$\text{s. t.}\quad(\hat{t}+b-y_1)^2+\delta\left(\frac{\hat{t}}{2}+b\right)^2=(\delta b)^2+\delta b^2$$

拉格朗日方程为：

$$L(\hat{t},y_1,\lambda)=-\frac{\hat{t}^3+3\hat{t}y_1-3\hat{t}^2y_1}{3}-\frac{\delta\hat{t}^3}{12}+(b+\delta b)^2\hat{t}+\lambda\left[(\hat{t}+b-y_1)^2+\delta\left(\frac{\hat{t}}{2}+b\right)^2-(\delta b)^2-\delta b^2\right]$$

由此可得：

$$\frac{\partial L}{\partial\hat{t}}=(b+\delta b)^2-\frac{\delta\hat{t}^2}{4}-(\hat{t}^2+y_1^2-2\hat{t}y_1)+\lambda\left[2(\hat{t}+b-y_1)+\delta\left(\frac{\hat{t}}{2}+b\right)\right]=0$$

$$\frac{\partial L}{\partial y_1}=(\hat{t}^2-2\hat{t}y_1)+\lambda\cdot2(y_1-\hat{t}-b)=0$$

通过整理并消除 λ，我们可以得到 $(\hat{t}+b-y_1)^2+\delta(\hat{t}/2+b)^2=(\delta b)^2+\delta b^2$ 和 $[(b+\delta b)^2-\delta\hat{t}^2/4-(\hat{t}^2+y_1^2-2\hat{t}y_1)]/(\hat{t}^2-2\hat{t}y_1)=[2(\hat{t}+b-y_1)+\delta(\hat{t}/2+b)]/[-2(\hat{t}+b-y_1)]$，由此得到最优的 $\hat{t}$ 和 y_1。

以上我们给出的是单侧不连续代理策略中满足条件的最优的 $\hat{t}$ 和 y_1，但是我们无法给出 $\hat{t}$ 和 y_1 的显示解。虽然无法求出两者的显示解，但是我们可以通过计算可知，最优的 $\hat{t}$ 和 y_1 总是存在的。

由之前的定理我们可知总是存在激励相容的不连续代理策略——在临界值 $\hat{t}$ 的左侧混同，在临界值 $\hat{t}$ 的右侧分离——占优于完全分离策略。在这种单侧不连续代理策略下，我们能找到最优的 $\hat{t}$ 和 y_1，使得中央政府能最大化自身效用，只要 $\hat{t}$ 和 y_1 同时满足上述两个条件时，中央政府就能在单侧不连续策略下最大化自身效用。同时由之前的证明可知，在信息区间的右侧，还存在 $\tilde{t}$ 和 y_1' 能进一步优化代理模型，但由于对 $\hat{t}$ 和 y_1 的计算太过复杂，本文在此将不再展开篇幅进行讨论。

四、结论

本文从信息传递的角度出发，一方面构建了多个不同信息传递机制下的理论模型，包括中央集权、廉价磋商和财政分权，通过对不同模型结果的比较，从而验证了财政分权体制的合理性、优越性和局限性。另一方面探讨了更具有现实意义的两期下的中央政府与地方政府的博弈，构建了中央政府有限承诺下的两期代理模型。首先我们给出了在两期动态博弈下中央政府的连续最优激励机制，该机制使得地方政府有动机在第一期揭露其真实信息，从而使得中央政府可以在第二期采取最优决策水平。之后我们研究了对这一连续激励机制进一步优化的方法，从而发现在信息区间上存在不连续的代理机制，可以在完全分离均衡的基础上进一步提升中央政府的效用。

在所有模型中，我们假设中央政府和地方政府的偏好不同，也就是各自对于当地的最优财政支出水平的偏好不同。对于地方政府而言，当其自身激励与居民福利相契合时，其最优支出水平就是最大化居民效用的公共产品供给水平；但是中央政府要考虑的不仅是某个区域的居民利益，更是全国整体的利益，因此中央政府需要统一协调全国各地的公共产品供给，实现资源在全国范围内的最优配置。正是由于两者的角色不同，中央政府与地方政府在当地财政支出水平上必然存在不同的偏好，因此双方的效用函数存在差异。

在两期代理模型中，中央政府在第一期将决策权下放到地方政府，在第二期，中央政府将根据第一期的信息推断地方真实状态 t 并独立做出决策。我们发现，对于任何 $\delta>0$ 的两期代理来说，静态最优代理是激励不相容的。一个激励相容的代理机制应该满足以下条件：对于地方政府而言，任意偏离真实值的 $\tilde{t}$ 带来的收益都小于传递真实值带来的收益。由此我们

得到了一个在第一期的均衡结果，即 $y_1(t)=t+b+\delta b$，当中央政府设计并执行这一机制时，地方政府有动机在第一期揭露其真实信息。

然而本文证明了完全分离策略并不是最优的，因为在完全分离策略下，中央政府为了让地方政府揭露其真实信息，第一期的决策水平过于偏离其最优水平，由此导致了过多的效用损失。之后本文证明了总是存在不连续代理策略占优于完全分离策略，在不连续代理策略下，中央政府将在某些信息区间内采取两期的完全混同策略，虽然这将导致第二期的效用损失，但也将大量减少第一期的效用损失。

此外，本文也证明了在两侧临界值 $\hat{t}$ 和 $\tilde{t}$ 之间不可能存在另外的连续区间，使得中央政府可以采取完全混同策略，因为这一连续区间的起点值和终点值必须同时满足激励相容的约束条件，而我们无法找到同时满足这两个条件的区间值。最后我们给出了最优的代理策略：当中央政府与地方政府的收益偏差 b 足够小时，最优代理策略中包含以下两种形式：一种是完全混同代理策略，在左右两侧的每个区间上，中央政府都采用完全混同策略；另一种是完全分离策略，且存在完全分离策略的区间是连通的。

中央政府拥有两期的承诺能力是很有利的，但因为种种原因，现实中中央政府很多情况下的承诺能力是有限的，本文的模型给出了政府在不同承诺力下的最优机制，具有重要的理论和现实意义。

本文并未考虑两期代理模型下地方真实状态 t 值在两期不同的情形，同时本文假设中央政府和地方政府对于两者之间的收益偏差 b 值的信息是一致的，后续研究可以从上述这两点进行拓展。

参考文献

[1] Tibeout C. Pure Theory of Clubs [J]. American Economic Review, 1956, (64): 416-424.

[2] Musgrave R A. The Theory of Public Finance [M]. McGraw-Hill, New York, 1959.

[3] Oates W. Fiscal Decentralization [M]. Harcourt, Barce and Jovanovich, 1972.

[4] Qian Y, Roland G. Federalism and the Soft Budget Constraint [J]. American Economic Review, 1998, (88): 1143-1162.

[5] Qian Y, Weingast B. Federalism as a Commitment to Preserving Market Incentives [J]. Journal of Economic Perspectives, 1997, (11): 83-92.

[6] Akai N, Sakata M. Fiscal Decentralization Contributes to Economic Growth: Evidence from State-level Cross-section Data for the United States [J]. Journal of Urban Economics, 2002, 52: 93-108.

[7] 林毅夫，刘志强. 中国的财政分权与经济增长 [J]. 北京大学学报，2000 (4): 5-17.

[8] 张晏，龚六堂. 分税制改革、财政分权与中国经济增长 [J]. 经济学（季刊），2005 (4): 75-108.

[9] 周业安，章泉. 财政分权、经济增长和波动 [J]. 管理世界，2008 (3): 6-15, 186.

[10] Holmstrom B. On Incentives and Control in Organizations [D]. Ph. D. Dissertation, Stanford, 1977.

[11] Melumad N D, Shibano T. Communication in Settings with No Transfers [J]. The RAND Journal of Economics, 1991 (22): 173-198.

[12] Martimort D, Semenov A. Continuity in Mechanism Design Without Transfers [J]. Economics Letters, 2006, 93: 182-189.

[13] Martimort D, Koessler F. Optimal Delegation with Multi-dimensional Decisions [J]. Journal of Economic Theory, 2012, 147: 1850-1881.

[14] Dessein W. Authority and Communication in Organizations [J]. The Review of Economic Studies, 2002, 69: 811-838.

[15] Alonso R, Matouschek N. Optimal Delegation [J]. The Review of Economic Studies, 2008, 75: 259-293.

[16] Ambrus A, Egorov G. Delegation and Nonmonetary Incentives [J]. Journal of EconomicTheory, 2017, 171: 101-135.

[17] Ralph B, Tracy R L. Advocacy and Dynamic Delegation [Z]. Working Paper, University of Miami, Department of Economics, 2011.

[18] Yin X. Dynamic Delegation [D]. Ph. D. Dissertation, Toulouse School of Economics, 2012.

[19] Crawford V P, Sobel J. Strategic Ivcformation Transmission [J]. Econometrica, 1982, 50: 1431-1451.

[20] Laffont J, Tirole J. The Dynamics of Incentive Conctracts [J]. Econometrica, 1988, 56 (5): 1153-1175.

（责任编辑：李　晟　张安平）

内部控制影响长期并购绩效的中介效应研究
——董事持股的异质性情境

Research on the Intermediary Effect of Internal Control on Long-term Merger Performance: The Heterogeneous Situation of Directors Holding Shares

曾江洪　曾琪姗　黄向荣

ZENG Jiang-hong　ZENG Qi-shan　HUANG Xiang-rong

[摘　要] 上市公司内部控制制度对并购绩效有积极影响，但内部控制影响并购绩效的中介路径与适用情境还有待深入研究。笔者以我国A股市场2009—2017年发生并购事件的上市公司为研究对象，采用固定效应OLS模型，实证考察内部控制能否对高管的真实盈余管理行为发挥治理效应以及这一效应能否正向影响长期并购绩效。在此基础上，笔者进一步检验董事持股对内部控制与长期并购绩效的调节作用。研究发现，真实盈余管理在内部控制与长期并购绩效之间起到中介作用；董事持股比例越高，内部控制对长期并购绩效的正向影响越强。研究还发现，内部控制对真实盈余管理的治理效应在并购完成的次年依然存在，并能继续改善长期并购绩效的表现；内部控制对上市公司长期股票投资收益存在正向影响。本研究及其结论丰富了企业治理效应理论方面的现有文献，也能为上市公司改善长期并购绩效表现提供参考。

[关键词] 并购重组　内部控制　真实盈余管理　董事特征　长期并购绩效

[中图分类号] F275.1　[文献标识码] A　[文章编号] 1000-1549（2020）10-00102-13

Abstract: The internal control system of listed companies has a positive impact on M&A performance, but the intermediary path and applicable situation in which internal control affects M&A performance need further study. The authos use the listed companies that initiated M&A events in China's A-share market from 2009 to 2017 as the research object, and uses the fixed-effect OLS model to empirically investigate whether internal control can exert a governance effect on the true earnings management behavior of executives, and whether this effect can positively affect long-term M&A performance. On this basis, the author further examine the role of directors' shareholding in regulating internal control and long-term M&A performance. The study found that real earnings management acts as an intermediary between internal control and long-term M&A performance. The higher the shareholding ratio of directors, the stronger the positive impact of internal control on long-term M&A performance. The study also found that the governance effect of internal control on true earnings management still exists the year after the completion of M&A, and can further promote the improvement of long-term M&A performance. Internal control has a positive impact on the long-term stock investment returns of listed companies. This study and its conclusions enrich the existing literature on the theory of corporate governance effects and can also provide a reference for listed companies to improve their long-term M&A performance.

Key words: Mergers and acquisitions　Internal control　Real earnings management　Director characteristics　Long-term M&A performance

[收稿日期]　2020-06-18

[作者简介]　曾江洪（通讯作者），男，1968年4月生，中南大学公司治理与资本运营研究中心主任、商学院教授，管理学博士，博士生导师，研究方向为公司治理、企业并购等；曾琪姗，女，1997年4月生，中南大学商学院硕士研究生，研究方向为企业并购、商誉；黄向荣，女，1990年6月生，湖南财政经济学院讲师，管理学博士，研究方向为创业融资、中小企业市场营销和公司治理。

[基金项目]　国家社会科学基金项目“一带一路建设中我国机械制造过剩产能需求侧与供给侧化解机制研究”（项目编号：16BJY079）。感谢匿名评审人提出的修改建议，笔者已做了相应修改，本文文责自负。

一、引言

面对竞争压力，很多上市公司选择通过并购实现外延式扩张。但居高不下的并购失败率，表明部分上市公司并未获得预期收益。因此，如何帮助企业获得理想的并购绩效是学术界长期关注的重要议题之一。在公司治理领域，已有研究证实了内部控制对并购绩效有显著的正向影响（崔永梅和余璇，2011[1]；赵息和张西栓，2013[2]；Harp和Barnes，2018[3]）。但研究内部控制如何影响并购绩效，还需要先厘清内部控制影响并购绩效的中介路径与适用情境。

内部控制可以有效抑制高管的机会主义行为。企业高管是并购活动的主导者，高管有充分动机与能力实施盈余管理行为（张自巧和葛伟杰，2013[4]）。面对日益严格的监管政策，高管的盈余管理行为多由应计盈余管理转变为更为隐蔽的真实盈余管理（Cohen等，2008[5]）。而盈余管理行为会对企业业绩造成负向影响（王福胜等，2014[6]）。内部控制能够有效抑制高管的应计盈余管理行为（Ashbaugh-Skaife等，2008[7]；董望和陈汉文，2011[8]），但内部控制对真实盈余管理的治理效应尚有争论（范经华等，2013[9]；胡明霞，2018[10]；王嘉鑫和王永海，2019[11]）。基于此，笔者将真实盈余管理引入内部控制对长期并购绩效影响的研究中，一方面，拟检验内部控制是否对并购后高管的真实盈余管理行为发挥治理效应；另一方面，拟检验真实盈余管理是否是内部控制影响长期并购绩效的中介路径。

《企业内部控制基本规范》文件指出，董事会对内部控制负有最终责任。董事参与公司治理的程度会影响内控有效性（吕景胜和赵玉梅，2016[12]）与并购绩效（江涛等，2019[13]）。董事持股能有效激励董事参与公司治理。已有研究并未关注到董事持股这一异质性情境对内部控制与长期并购绩效关系的影响。因此，笔者拟将董事持股作为异质性情境，检验董事持股是否在内部控制与长期并购绩效中起到调节作用。

本文基于2009—2017年我国上市公司并购事件的面板数据，实证考察内部控制对并购后真实盈余管理的治理效应，考察这一治理效应是否在内部控制与长期并购绩效间起到中介作用，以及考察董事持股对内部控制与长期并购绩效关系的调节作用。

二、文献综述与研究假设

（一）内部控制、真实盈余管理与长期并购绩效

已有研究用不同研究方法对内部控制与并购绩效的关系进行了分析，得出了较为一致的结论，即内部控制对并购绩效有显著的正向影响。少数学者关注到了内部控制影响并购绩效的中介路径。杨道广等（2014）[14]研究指出，有效的内部控制可以提高并购整合能力、对并购绩效有正向影响，并认为整合能力是内部控制影响并购绩效的中介路径，但未对该中介路径进行实证检验。此外，已有研究并未讨论内部控制与并购绩效的其他中介路径。笔者引入真实盈余管理，拟在已有研究的基础上丰富内部控制影响长期并购绩效的中介路径研究。

真实盈余管理是通过操纵销售、生产成本与酌量性费用等真实活动调整盈余的行为，通常偏离企业正常经营活动决策。在我国并购市场现金支付占比高达80%的背景下（李井林等，2014[15]），并购后上市公司为并购交易承担了较大的现金流支出压力，同时面临较大的整合风险，这不仅给上市公司的财务表现造成压力，也促使高管在并购后进行正向真实盈余管理（张自巧和葛伟杰，2013[4]）。在面临财务压力最大的第三、四季度，这一现象更加明显（孙梦男和吴迪，2017[16]）。因此，并购后高管可能采取增加产品折扣、变更信用政策、分摊固定成本、减少研发费用（朱湘忆，2020[17]）等手段操纵利润。高管的上述行为必然对应企业真实销售、成本计算、研发等行为的变动。一方面，真实生产经营活动的变化涉及人员与流程的变更，这会给企业增加不必要的调整成本。另一方面，暂时提高的盈余不具备可持续性，虽然真实盈余管理行为在短期内可以提高企业利润，但在一年及以上的会计周期中，反而会给企业造成负担，影响绩效表现（王福胜等，2014[6]）。

内部控制的基本职能是保证会计信息质量（Kinney和Mcdaniel，1989[18]；Bell，2000[19]；阎达五和杨有红，2001[20]）。上市公司内部控制制度更有效，上市公司的会计信息质量更高，则并购双方面临的信息不对称程度更低（刘启亮，2013[21]；王晶等，2015[22]）。这将有利于并购双方了解彼此的真实情况，在整合环节促成有效沟通，进一步改善并购表现。此外，内部控制还能通过抑制高管的真实盈余管理行为改善并购表现。一方面，内部控制能有效提升企业的会计稳健性（王宗润和陈艳，2014[23]），而会

计稳健性能有效抑制高管真实盈余管理行为（蒋勇和王晓亮，2019[24]），因此内部控制可以间接影响高管的真实盈余管理行为。另一方面，内部控制制度可以直接约束管理者自利行为（韩岚岚和马元驹，2017[25]）与机会主义行为（Feng 等，2015[26]）。高管通过实施真实盈余管理行为操纵利润，需要公司内部多部门甚至外部其他公司的配合。而贯穿企业生产经营各环节的内部控制制度，可以有效监督与约束公司内外部各关联方的行为。当公司内部控制制度更有效时，内部控制制度会发挥更好的制衡作用，增加真实盈余管理行为的实施成本，进而减少高管实施真实盈余管理行为的空间。因此，公司内部控制制度越有效，高管实施的真实盈余管理行为越少，进而能够减少真实盈余管理行为对并购绩效的毁损。基于上述分析，我们提出假设 1 和假设 2。

H1：内部控制有效性正向影响长期并购绩效。

H2：真实盈余管理在内部控制与长期并购绩效之间起部分中介作用。

（二）内部控制、董事持股与长期并购绩效

Hauser 和 Roie（2018）[27]研究指出，如果一家企业的 CEO 接受了其他公司的董事会任命，那么会影响所担任 CEO 公司的业绩表现。董事的时间、经历与资源均是有限的，然而无论独立董事还是非独立董事，都可以在无竞业限制的多家企业“兼职”，这种“兼职”会影响董事在一家企业的投入。董事“兼职”的可能，以及部分董事与公司利益不一致的现状，均决定了部分董事在公司治理中发挥的作用有限。但董事会作为内部控制制度建设的责任承担方，董事参与内部控制建设的投入程度直接决定了内部控制制度发挥治理效应的有效程度。郑丽和陈志军（2018）[28]研究指出，董事持股越多，与股东利益越一致；董事持股可以激发董事工作的积极性，并有效促进企业采取有一定风险的战略决策。在内部控制建设方面，已有研究证明，董事持股越多，企业的内部控制制度越有效（Kobelsky 等，2013[29]）。并购后的整合中，并购双方要面临生产、销售、财务等多领域业务活动的融合问题。生产经营活动与对应流程的变动意味着内部控制制度需要进行重构，重构需要付出一定的调整成本。董事持股越多，董事工作的积极性就越高，董事更倾向于主动承担建设内部控制并推动企业支付调整成本的责任，进而推动并购后内部控制制度的重构，使得内部控制制度能更有效地应对特殊的并购风险并改善并购绩效的表现。基于此，我们提出假设 3。

H3：董事持股在内部控制与长期并购绩效这一关系中起正向调节作用。

三、研究设计

（一）样本选择与数据来源

《企业内部控制基本规范》要求上市公司自 2009 年 7 月 1 日起对内部控制的有效性进行自我评价，并披露年度自我评价报告。笔者认为 2009 年前后企业对于内部控制所披露的信息量存在差距。考虑到并购绩效的考察通常会持续并购公告日后一年及以上的时间，因此本文拟选取 2009—2017 年上市公司并购事件作为初步研究样本，并进行以下筛选：剔除其中行业代码为 J 开头的金融行业样本数据；剔除未成功并购的样本数据；因为被解释变量为长期并购绩效，对一年内发生多起并购事件的只保留第一起；剔除存在缺失值的样本数据；剔除 ST 公司和 *ST 公司样本数据。最终得到 4 033 个样本。公司治理、财务信息等基本数据来自国泰安数据库，内部控制指数来自“迪博·内部控制与风险管理数据库”，并对所有连续变量进行 1% 的缩尾处理。本文使用 Excel 进行初步数据处理，Stata14.0 进行数据分析。

（二）模型设定与变量说明

本文经过 F 检验、LM 检验与豪斯曼检验后，结果显示选择固定效应。因为异方差的存在，本文回归时用 robust 对异方差进行控制。本文用模型 1 检验主效应，引入中心化后的交乘项，用模型 2 检验调节效应，用模型 1 与模型 3、模型 4 检验中介效应。

$$\Delta ROA/BHAR=\alpha_0+\alpha_1 DIB+\alpha_2 Control+\alpha_3\sum Year+\xi_1 \quad (1)$$

$$\Delta ROA/BHAR=\alpha_0+\alpha_1 DIB+\alpha_4 ShareD+\alpha_5 DIB\times ShareD+\alpha_2 Control+\alpha_3\sum Year+\xi_2 \quad (2)$$

$$DA=\alpha_0+\alpha_1 DIB+\alpha_2 Control+\alpha_3\sum Year+\xi_3 \quad (3)$$

$$\Delta ROA/BHAR=\alpha_0+\alpha_1 DIB+\alpha_6 DA+\alpha_2 Control+\alpha_3\sum Year+\xi_4 \quad (4)$$

本文的解释变量是内部控制，参考前人研究，选择迪博的内部控制质量指数，并对其进行取对数处理。该指数从合法合规、信息披露、经营回报、资产安全、战略执行五个维度衡量企业内部控制的质量，并根据内控缺陷进行修正。

本文的被解释变量是长期并购绩效。关于并购绩效的衡量通常分短期与长期两种。短期绩效通常基于事件研究法，用公告日前后的累积超额收益率来衡量绩效，本质体现的是股票波动；而长期绩效多采用财务指标法与事件研究法两种方法，前者衡量财务绩效，后者衡量市场绩效。本文立足于公司内部治理的角度讨论内部控制对长期并购绩效的影响，所研究的中介路径与适用情境均发生在企业内部，因此所讨论的并购绩效聚焦于财务绩效。财务指标法通过计算并购前后公司主要财务指标的变化来衡量并购绩效。本文参考前人研究，使用并购公告日后一年与前一年总资产报酬率的差额 *ROA*（-1，+1）衡量长期并购绩效，并用净资产报酬率的差额 *ROE*（-1，+1）作为稳健性检验指标。但为与已有研究保持一贯性，本文拟在进一步研究中引入上市公司长期股票投资收益指标代指市场绩效，讨论内部控制对上市公司长期股票投资收益（即市场绩效）的影响。具体指标为上市公司并购公告日后 12 个月的 *BHAR*（购买并持有超额收益率）指标，并以 24 个月的 *BHAR* 指标作为稳健性检验指标。计算公式如下：

$$BHAR_{i,T}=\prod_{t=0}^{T}(1+R_{i,t})-\prod_{t=0}^{T}(1+R_{p,t}) \tag{5}$$

等式左边为企业 i 在首次做出并购公告 t 月后连续 T 个月的买入并持有超额收益，本文中选取 $T=12$ 个月与 24 个月，对应 250 个工作日与 500 个工作日。$R_{i,t}$ 与 $R_{p,t}$ 是考虑现金红利再投资的股票 i 个股日回报率与对应投资组合的综合市场日回报率，将数据代入后根据公式进行连乘作差得出 BHAR。

本文的中介变量是真实盈余管理。参考李增福等（2011）[30]、Roychowdhury（2006）[31] 的衡量方式，从销售、生产、酌量性费用三个方面计算真实盈余管理。本文先通过公式（6）回归估计出企业正常的经营现金净流量，通过公式（7）~公式（9）回归估计出企业正常的生产成本（包括销售产品成本与存货变化额），通过公式（10）回归估计出企业正常酌量性费用，再通过现实值减去估计出的正常值得到异常值，最后通过异常生产成本减去异常酌量性费用与异常现金净流量得到真实盈余管理指标。具体公式如下：

$$\frac{CFO_t}{TA_{t-1}}=\alpha_0+\alpha_1\frac{1}{TA_{t-1}}+\beta_0\frac{SALES_t}{TA_{t-1}}+\varepsilon_t \tag{6}$$

$$\frac{COGS_t}{TA_{t-1}}=\alpha_0+\alpha_1\frac{1}{TA_{t-1}}+\beta_0\frac{SALES_t}{TA_{t-1}}+\varepsilon_t \tag{7}$$

$$\frac{\Delta INV_t}{TA_{t-1}}=\alpha_0+\alpha_1\frac{1}{TA_{t-1}}+\beta_0\frac{SALES_t}{TA_{t-1}}+\beta_1\frac{\Delta SALES_{t-1}}{TA_{t-1}}+\varepsilon_t \tag{8}$$

$$\frac{PROD_t}{TA_{t-1}}=\alpha_0+\alpha_1\frac{1}{TA_{t-1}}+\beta_0\frac{SALES_t}{TA_{t-1}}+\beta_1\frac{\Delta SALES_t}{TA_{t-1}}+\beta_2\frac{\Delta SALES_{t-1}}{TA_{t-1}}+\varepsilon_t \tag{9}$$

$$\frac{DISEXP_t}{TA_{t-1}}=\alpha_0+\alpha_1\frac{1}{TA_{t-1}}+\beta_0\frac{SALES_{t-1}}{TA_{t-1}}+\varepsilon_t \tag{10}$$

其中：*TA* 是总资产；*SALES* 是销售收入；*COGS* 是销售产品成本；*INV* 是存货；*PROD* 是产品成本；*CFO* 是公司经营活动现金流量；*DISEXP* 是酌量性费用，包括营业费用与管理费用。

本文的调节变量是董事持股，即董事会全体成员的持股数量，考虑到不同企业股数的差异，采用并购年度董事持股数/公司总股数，即董事持股占比，以衡量董事持股水平。

本文的控制变量主要从企业层面与并购的交易层面来选取。企业层面选取企业规模、偿债能力、自由现金流、成长性、盈利能力、Herfindahl 指数、是否两职合一、独立董事占比、高管过度自信、股票高估。并购交易层面选取交易规模、支付方式、是否关联交易、控制年份与行业。表 1 为所有涉及的变量的具体含义与衡量方式。

表 1　　变量定义

	变量代码	变量名	变量说明与衡量方式
解释变量	*DIB*	内部控制	Ln（1+迪博内部控制指数）
被解释变量	*BHAR*	长期股票投资收益	并购公告日后 12 个月（250 个工作日）的买入并持有超额收益率
	BHAR2	长期股票投资收益	并购公告日后 24 个月（500 个工作日）的买入并持有超额收益率
	Δ*ROA*	长期并购绩效	并购事件发生年度（-1，+1）的总资产收益率变动值
	Δ*ROE*	长期并购绩效	并购事件发生年度（-1，+1）的净资产收益率变动值

续前表

	变量代码	变量名	变量说明与衡量方式
调节变量	*ShareD*	董事持股	董事持股数/公司总股数
中介变量	*DA*	真实盈余管理	根据公式（6）~公式（10）计算得出的，企业通过操纵真实活动（销售、生产、酌量性费用）进行的盈余管理
控制变量	Ln*Size*	企业规模	并购前一年总资产的自然对数
	Ln*Pay*	交易规模	并购事件支付金额的自然对数
	DIV	偿债能力	并购前一年年末资产负债率
	FCF	自由现金流	并购前一年年末自由现金流/总资产
	Growth	成长性	并购前一年年末托宾 Q 值
	ROA	盈利能力	并购前一年年末 *ROA*
	*H*1	Herfindahl 指数	并购当年第一大股东持股比例的平方和
	Same	是否两职合一	董事长与总经理是否是同一人，不确定取 0，是取 1，否取 2
	IndependentD	独立董事占比	独立董事人数/董事会总人数
	Overconfidence	高管过度自信	前三名高管薪酬/董监高总薪酬
	Dev	股票高估	上市公司股票价格是否存在高估，参考胡凡和李科（2019）[32] 计算得出
	Rela	关联交易	并购交易是否是关联交易，是取 1，否取 0
	Paytype	支付方式	并购交易是否采用现金支付，是取 1，否取 0
	Industry	行业	控制
	Year	年份	控制

四、实证分析

（一）描述性统计分析

描述性统计结果如表 2 所示。样本企业并购后一年与前一年总资产收益率差额的均值为-0.7%，标准差为 0.055，最小值为-23%，最大值为 18%，净资产收益率差额普遍比总资产收益率差额大 0.3%，且波动更大。可见样本企业的并购事件，在财务绩效层面呈现的是“价值毁损”的特征，且企业间有较大的差异。

样本企业并购事件公告日后 12 个月与 24 个月的 *BHAR* 均值均为正值，分别为 12.3%与 13.8%，说明资本市场普遍对上市公司的并购事件持看好态度，并在股价上给出正向反馈。此外，二者最小值分别是-72.7%与-104%，最大值分别是 264.7%与 378%，说明不同并购事件中，上市公司长期股票投资收益也有较大差异。

样本企业并购当年的内部控制质量指数均值为 6.5，标准差为 0.108，最大值为 6.782，最小值为 6.072，可见样本企业的内部控制基本达到了及格水准，但仍有很大的提升空间。样本企业并购当年的真实盈余管理均值为-0.165，标准差为 0.298，最大值与最小值分别是 0.819 与-1.203，说明上市公司既存在向上的真实盈余管理，也存在向下的真实盈余管理。

样本企业并购当年的董事持股均值为 13.6%，标准差为 0.189，最小值为 0，最大值为 65%，可见样本企业中董事持股占比普遍不高，而且公司之间差别较大，为本文选题的研究提供了一个可行的背景。由其他数据的描述性统计信息可以看出，上市公司在并购事件中更倾向于选择现金支付的支付方式；并购当年，上市公司普遍存在股价高估与管理者过度自信的现象。

表 2 描述性统计

变量	观测数	均值	标准差	最小值	最大值
DIB	4 033	6. 501	0. 108	6. 072	6. 782
BHAR	4 033	0. 123	0. 572	−0. 727	2. 647
*BHAR*2	4 033	0. 138	0. 794	−1. 040	3. 780
Δ*ROA*	4 033	−0. 007	0. 055	−0. 237	0. 181
Δ*ROE*	4 033	−0. 010	0. 112	−0. 502	0. 406
ShareD	4 033	0. 136	0. 189	0. 000	0. 650
DA	4 033	−0. 165	0. 298	−1. 203	0. 819
Ln*Size*	4 033	21. 923	1. 165	19. 748	25. 487
Ln*Pay*	4 033	18. 122	2. 014	11. 918	22. 507
DIV	4 033	0. 416	0. 208	0. 044	0. 863
FCF	4 033	11. 055	9. 210	0. 000	22. 209
Growth	4 033	2. 407	1. 560	0. 974	9. 481
ROA	4 033	0. 052	0. 049	−0. 096	0. 211
*H*1	4 033	0. 142	0. 116	0. 008	0. 539
Same	4 033	1. 706	0. 475	0. 000	2. 000
IndependentD	4 033	0. 373	0. 053	0. 333	0. 571
Overconfidence	4 033	0. 363	0. 134	0. 000	0. 711
Dev	4 033	0. 065	0. 518	−1. 031	1. 467
Rela	4 033	0. 341	0. 474	0. 000	1. 000
Paytype	4 033	0. 847	0. 360	0. 000	1. 000
Industry	4 033	—	—	1. 000	71. 000
Year	4 033	—	—	2009	2017

（二）相关性检验分析

表 3 列示了主要变量的相关性分析信息，主要变量之间均有较强的相关性，且多在 1% 的水平上显著；相关系数较高（>0. 5）的两组变量均为替代变量，分别是 *BHAR* 与 *BHAR*2、*ROA* 与 *ROE*，本身计算逻辑一致且不会出现在同一模型，不影响回归效度。篇幅原因，没有列示控制变量的相关性分析信息。其结果显示控制变量与控制变量之间、控制变量与主要变量之间，大部分的相关系数均低于 0. 5，仅有 *DIV* 与 Ln*Size* 的相关系数为 0. 518，但变量间 VIF 值均小于 3，说明模型不存在多重共线性问题。

表 3 相关性分析

变量	*DIB*	*BHAR*	*BHAR*2	Δ*ROA*	Δ*ROE*	*ShareD*	*DA*
DIB	1. 000						
BHAR	0. 043 ***	1. 000					
*BHAR*2	0. 049 ***	0. 695 ***	1. 000				

续前表

变量	*DIB*	*BHAR*	*BHAR2*	Δ*ROA*	Δ*ROE*	*ShareD*	*DA*
Δ*ROA*	0.060***	0.298***	0.262***	1.000			
Δ*ROE*	0.056***	0.297***	0.260***	0.902***	1.000		
ShareD	−0.033**	0.098***	0.101***	−0.063***	−0.030*		
DA	−0.066***	0.012	0.045***	−0.057***	−0.075***	−0.061***	1.000

注：*、**、*** 分别表示在 10%、5%、1%的水平上显著。下同。

（三）回归结果分析

表 4 列示了以并购绩效 Δ*ROA*、真实盈余管理 *DA* 为因变量的回归结果。模型 1 与模型 2 检验假设 1，模型 2、3、4 检验假设 2，模型 2、5、6 检验假设 3。通过对比各模型可以发现，相同因变量条件下，各模型整体与控制变量的系数与显著性水平均较为稳定，模型 *F* 统计量对应的 *P* 值均为 0.000，说明模型有效且稳定。

模型 1 放入了所有的控制变量，企业规模、成长性、盈利能力、第一大股东持股比例、过度自信、股票高估与支付方式均对长期并购绩效有显著影响。模型 2 在模型 1 的基础上加入了解释变量内部控制，回归系数为 0.069，*p* 值为 0.000，说明内部控制在 1%的水平上显著地正向影响长期并购绩效。假设 1 得以验证。

模型 3 以真实盈余管理作为被解释变量，以内部控制作为解释变量，控制变量不变。因为模型的控制变量多参照并购领域研究选取，大部分对真实盈余管理并无显著影响，导致模型 3 有效性下降。但 *F* 统计量对应的模型 *p* 值仍为 0.000，因此本文认为模型可信。此时内部控制的回归系数为−0.121，*p* 值为 0.039，内部控制在 5%的水平上显著地负向影响真实盈余管理，说明上市公司内部控制制度能有效抑制高管的真实盈余管理行为。模型 4 在模型 2 的基础上加入了真实盈余管理作为解释变量，内部控制的回归系数为 0.068，*p* 值为 0.000，真实盈余管理的回归系数为−0.015，*p* 值为 0.000，均在 1%的水平上显著，结合模型 2 与 3，说明内部控制通过抑制高管的真实盈余管理行为，减少真实盈余管理对并购绩效的毁损。假设 2 得以验证。

模型 5 在模型 1 的基础上加入了董事持股作为解释变量，回归系数为 0.067，*p* 值为 0.001，说明董事会全体董事持股在 1%的水平上显著地正向影响长期并购绩效。模型 6 在模型 5 的基础上加入了内部控制，以及经过中心化处理后的内部控制与董事持股的交乘项，此时内部控制的回归系数为 0.073，*p* 值为 0.000，董事持股的回归系数为 0.066，*p* 值为 0.000，均在 1%的水平上显著，交乘项的回归系数为 0.123，*p* 值为 0.033，在 5%水平上显著，说明董事持股正向调节内部控制与长期并购绩效的关系。假设 3 得以验证。

表 4　　**内部控制影响长期并购绩效的回归结果**

变量	Model1−H1	Model2−H1	Model3−H2	Model4−H2	Model5−H3	Model6−H3
	Δ*ROA*	Δ*ROA*	*DA*	Δ*ROA*	Δ*ROA*	Δ*ROA*
DIB		0.069*** (5.73)	−0.121** (−2.06)	0.068*** (5.53)		0.073*** (6.19)
DA				−0.015*** (−3.92)		
ShareD					0.067*** (3.45)	0.066*** (3.56)
ShareD×*DIB*						0.123** (2.13)
Ln*Size*	−0.023*** (−7.28)	−0.023*** (−7.53)	0.044** (2.48)	−0.022*** (−7.27)	−0.023*** (−7.23)	−0.022*** (−7.40)

续前表

变量	Model1-H1	Model2-H1	Model3-H2	Model4-H2	Model5-H3	Model6-H3
	ΔROA	ΔROA	DA	ΔROA	ΔROA	ΔROA
LnPay	0.000 (-0.41)	0.000 (-0.30)	-0.001 (-0.41)	0.000 (-0.34)	0.000 (-0.39)	0.000 (-0.35)
DIV	0.017 (1.36)	0.019 (1.48)	-0.082 (-1.18)	0.018 (1.38)	0.020 (1.59)	0.023* (1.77)
FCF	0.000 (-0.97)	0.000 (-1.00)	0.000 (0.30)	0.000 (-0.97)	0.000 (-0.92)	0.000 (-0.99)
Growth	-0.002* (-1.89)	-0.002** (-1.96)	-0.016** (-2.17)	-0.002** (-2.20)	-0.002 (-1.60)	-0.002 (-1.62)
ROA	-0.955*** (-27.14)	-0.961*** (-28.00)	-0.091 (-0.48)	-0.962*** (-28.43)	-0.954*** (-27.69)	-0.958*** (-28.47)
H1	0.082*** (3.70)	0.070*** (3.27)	0.147 (1.18)	0.072*** (3.36)	0.079*** (3.64)	0.068*** (3.24)
Same	-0.004 (-1.29)	-0.005 (-1.49)	-0.003 (-0.20)	-0.005 (-1.51)	-0.004 (-1.41)	-0.005 (-1.57)
IndependentD	-0.007 (-0.24)	-0.005 (-0.19)	0.084 (0.48)	-0.004 (-0.15)	-0.011 (-0.39)	-0.009 (-0.35)
Overconfidence	0.021* (1.68)	0.020* (1.71)	-0.179*** (-2.77)	0.018 (1.50)	0.020* (1.67)	0.020* (1.74)
Dev	0.015*** (4.43)	0.015*** (4.61)	-0.053*** (-2.72)	0.014*** (4.40)	0.016*** (4.69)	0.016*** (4.88)
Rela	-0.002 (-0.93)	-0.001 (-0.68)	-0.002 (-0.14)	-0.001 (-0.69)	-0.001 (-0.69)	-0.001 (-0.39)
Paytype	-0.006* (-1.94)	-0.005* (-1.80)	0.006 (0.33)	-0.005* (-1.79)	-0.006** (-2.02)	-0.006* (-1.91)
_cons	0.544*** (8.06)	0.089 (0.86)	-0.339 (-0.60)	0.084 (0.81)	0.527*** (7.86)	0.043 (0.42)
Industry	控制					
Year	控制					
F-statistic	56.29	59.00	6.09	60.01	57.47	57.95
Prob>F	0.000	0.000	0.000	0.000	0.000	0.000
Adjusted_R^2	47.45%	48.76%	8.6%	49.16%	48.05%	49.48%
R^2	47.72%	49.04%	9.1%	49.45%	48.33%	49.78%

五、进一步研究与稳健性检验

（一）进一步研究

1. 内部控制与上市公司长期股票投资收益。在资本市场中，上市公司需要对内部控制信息进行披露，并聘请会计事务所出具内部控制审计报告。因此，内部控制越有效的企业，其内部控制质量指数越高，则能在市场中传递出公司治理越良好的信号。投资者愿意相信拥有高质量内部控制制度企业的并购，比存在内控缺陷企业的并购更加严谨并有希望获

得成功。投资者的积极态度将进一步反映在股票的超额收益上。因此内部控制有效性会正向影响上市公司的长期股票投资收益。但内部控制基于信号传递效应对上市公司长期股票投资收益产生影响的具体机理很明显不同于内部控制基于企业治理效应对财务绩效产生影响的机理。本文将被解释变量替换为并购公告日后 12 个月的 *BHAR*，并以 24 个月的 *BHAR* 作为稳健性检验。

表 5 列示了以 *BHAR* 为因变量的回归结果，篇幅原因没有列示控制变量的回归系数。模型 7 与模型 8 检验假设 1，模型 8、3、9 检验假设 2，模型 8、10、11 检验假设 3。

模型 7 放入了与模型 1 一致的控制变量，*F* 统计量显示模型有效；模型 8 在模型 7 的基础上加入了解释变量内部控制，回归系数为 0.241，*p* 值为 0.028，说明内部控制在 5%的水平上显著地正向影响上市公司长期股票投资收益。

模型 3 与表 4 的模型 3 一致，说明内部控制在 5%的水平上显著地抑制真实盈余管理；模型 9 在模型 8 的基础上加入了中介变量真实盈余管理，此时内部控制的回归系数为 0.247，*p* 值为 0.026，在 5%的水平显著，真实盈余管理的回归系数为 0.052，*p* 值为 0.312，并不显著，说明内部控制并不是通过抑制真实盈余管理行为影响上市公司长期股票投资收益的。

模型 10 在模型 7 的基础上加入了董事持股作为解释变量，董事持股回归系数为 0.133，*p* 值为 0.435，说明董事持股并不会影响上市公司长期股票投资收益；模型 11 在模型 10 的基础上加入了解释变量内部控制以及经过中心化后内部控制与董事持股的交乘项，此时内部控制的回归系数为 0.263，*p* 值为 0.020，在 5%的水平显著，董事持股的回归系数为 0.130，*p* 值为 0.446，交乘项的回归系数为 0.879，*p* 值为 0.178，均不显著，说明董事持股并不能加强内部控制对上市公司长期股票投资收益的正向影响。

通过对 *BHAR* 的回归可以发现，内部控制对上市公司长期股票投资收益也有显著的正向影响，但这一影响并不是通过抑制真实盈余管理行为得以实现，也无法通过董事持股得以加强。该结果说明，在内部控制与并购绩效的研究中，应当对财务绩效与市场绩效加以区分，内部控制对二者的影响具备不同的作用路径与适用情境。

表 5　　**内部控制影响长期股票投资收益的回归结果**

变量	Model7-H1	Model8-H1	Model3-H2	Model9-H2	Model10-H3	Model11-H3
	BHAR	*BHAR*	*DA*	*BHAR*	*BHAR*	*BHAR*
DIB		0.241** (2.19)	−0.121** (−2.06)	0.247** (2.24)		0.263** (2.34)
DA				0.052 (1.01)		
ShareD					0.133 (0.78)	0.130 (0.76)
ShareD×DIB						0.879 (1.35)
Industry	控制					
Year	控制					
F-statistic	26.56	25.78	6.09	24.63	25.52	24.02
Prob>F	0.000	0.000	0.000	0.000	0.000	0.000
Adjusted_R²	27.21%	27.34%	8.6%	27.37%	27.21%	27.40%
R^2	27.59%	27.74%	9.1%	27.78%	27.61%	27.83%

2. 并购次年的内部控制、真实盈余管理与长期并购绩效。

前文已证实，上市公司在并购当年，通过内部控制抑制高管的真实盈余管理行为，影响企业长期并购绩效。但高管并购后的真实盈余管理行为并不仅存在于并购当年。并购整合的长期性与高失败率决定了上市公司在并购后会面临长期的业绩压力。考虑到长期并购绩效的衡量覆盖了并购后一年上市公司的财务业绩，本文将进一步验证并购后一年上市公司的内部控制是否仍能通过抑制后一年的高管真实盈余管理行为影响长期并购绩效。

去除部分缺失值后，样本量变更为 3 955。表 6 中模型 12 放入了与主效应一致的控制变量，显示模型有效；模型 13 加入了并购次年上市公司的内部控制质量指数，回归系数为 0.117，p 值为 0.000，说明并购次年，内部控制仍在 1%水平上显著地正向影响长期并购绩效。模型 14 用并购次年的内部控制对并购次年的真实盈余管理进行回归，内部控制指数的系数为-0.135，p 值为-0.001，说明并购次年内部控制仍在 1%水平显著地负向影响真实盈余管理。模型 15 将并购次年的内部控制指数与真实盈余管理均放入回归模型，内部控制的回归系数为 0.113，p 值为 0.000，真实盈余管理的回归系数为-0.029，p 值为 0.000，说明并购次年，内部控制仍在 1%的水平上通过抑制真实盈余管理行为影响长期并购绩效。模型 16 与模型 17 分别将被解释变量替换为 ΔROE，以检验该部分结论的稳健性。可以发现关键变量的系数符号与显著性水平均与模型 13、15 一致，该部分的研究结论稳健。

表 6　　并购次年内部控制影响长期并购绩效的回归结果

变量	Model12	Model13	Model14	Model15	Model16	Model17
	ΔROA	ΔROA	DA_{t+1}	ΔROA	ΔROE	ΔROE
DIB_{t+1}		0.117*** (10.26)	-0.135*** (-3.23)	0.113*** (10.01)	0.252*** (10.75)	0.246*** (10.49)
DA_{t+1}				-0.029*** (-6.42)		-0.051*** (-5.37)
Industry	控制					
Year	控制					
F-statistic	59.79	72.31	4.97	73.95	54.92	57.48
Prob>F	0.000	0.000	0.000	0.000	0.000	0.000
*Adjusted_*R^2	49.40%	54.79%	7.83%	56.11%	47.57%	48.58%
R^2	49.67%	55.04%	8.34%	56.36%	47.86%	48.88%

（二）稳健性检验

表 7 列示了替换被解释变量后，前文研究的稳健性检验结果。模型 18 与模型 19 检验假设 1，模型 19 与模型 20 检验假设 2，模型 19、21、22 检验假设 3，模型 23 与模型 24 检验进一步研究第一部分，进一步检验的第二部分已在表 6 模型 16 与模型 17 列示。

模型 18 放入了全部控制变量，F 统计量显示模型有效；模型 19 在模型 18 的基础上加入解释变量内部控制，内部控制回归系数为 0.148，p 值为 0.000，说明假设 1 结论稳健，内部控制显著地正向影响长期并购绩效。

模型 20 在模型 19 的基础上加入了中介变量真实盈余管理，此时内部控制的回归系数与 p 值没有变化，真实盈余管理的回归系数为-0.035，p 值为 0.000，说明假设 2 结论稳健，内部控制通过抑制高管的真实盈余管理行为影响长期并购绩效。

模型 21 在模型 18 的基础上加入了董事持股作为解释变量，董事持股的回归系数为 0.129，p 值为 0.002。模型 22 在模型 21 的基础上加入了内部控制以及中心化后内部控制与董事持股的交乘项，此时内部控制的回归系数为 0.156，p 值为 0.000，董事持股的回归系数为 0.127，p 值为 0.002，交乘项的回归系数为 0.315，p 值为 0.034，说明假设 3 结论稳健，董事持股对内部控制与长期并购绩效有正向调节作用。

模型 23 在模型 19 的基础上，将被解释变量替换

为并购公告日后 24 个月的 *BHAR*，加入解释变量内部控制，此时内部控制的回归系数为 0.135，*p* 值为 0.326，并不显著，这可能是由于内部控制指数衡量的是并购当年的内控质量，而并购公告日后 24 个月的 *BHAR* 跨度太长，当年的内控难以产生如此长远的影响所致。因此，本文选取了并购后一年内部控制指数作为自变量，在模型 24 中对并购公告日后 24 个月的 *BHAR* 进行回归。并购次年内部控制指数对上市公司长期股票投资收益的回归系数为 0.373，*p* 值为 0.000，说明第一部分的进一步研究结论稳健，内部控制显著地正向影响上市公司长期股票投资收益。

表 7　　**基于变量替换的稳健性检验结果**

变量	Model18	Model19	Model20	Model21	Model22	Model23	Model24
	Δ*ROE*	Δ*ROE*	Δ*ROE*	Δ*ROE*	Δ*ROE*	*BHAR*2	*BHAR*2
DIB		0.148*** (5.19)	0.148*** (5.00)		0.156*** (5.60)	0.135 (0.98)	0.373*** (3.68)
DIB_{t+1}							
DA			−0.035*** (−3.90)				
ShareD				0.129*** (3.06)	0.127*** (3.13)		
ShareD×*DIB*					0.315** (2.12)		
Industry	控制						
Year	控制						
F-statistic	38.30	39.15	38.85	38.99	38.07	29.39	24.90
Prob>*F*	0.000	0.000	0.000	0.000	0.000	0.000	0.000
*Adjusted*_R^2	38.17%	39.61%	40.15%	38.71%	40.36%	32.92%	26.99%
R^2	38.50%	39.94%	40.49%	39.04%	40.71%	33.29%	27.40%

此外，考虑到内部控制更好的企业可能本身并购绩效更优，本文还进行了内生性检验，确保不存在样本选择偏差问题。本文参考叶康涛等（2015）[33]的研究，选取公司规模、负债率、成长性、前十大股东持股占比、独立董事比例、行业作为内控质量的解释变量，采用 Heckman 两阶段模型进行回归。计算出的逆米尔斯比率 lambda 的 *p* 值为 0.209，因此认为本文研究不存在样本选择偏差的内生性问题。

六、研究结论与展望

（一）研究结论

本文以我国 A 股市场 2009—2017 年发生并购事件的上市公司为研究对象，构建固定效应 OLS 模型，验证内部控制对并购后真实盈余管理的治理效应，以及内部控制与长期并购绩效之间真实盈余管理的中介作用与董事持股的调节作用，得出以下几点主要结论。

第一，有效的内部控制制度能抑制并购后高管的真实盈余管理行为，并进一步正向影响企业长期并购绩效；这一治理效应与中介关系在并购当年与并购次年均存在。该结论从高管行为角度回答了内部控制如何影响长期并购绩效的问题。结论说明我国上市公司的内部控制制度对企业高管的机会主义行为起到了监督与约束的作用，从侧面证明了高管在并购后确实存在真实盈余管理的倾向，且高管的真实盈余管理行为确实会对企业的长期并购绩效造成损失，而内部控制的治理效应可以有效减少该损失。

第二，董事持股在内部控制与长期并购绩效关系中起到正向调节作用。该结论说明，虽然我国上市公司的内部控制制度在并购后能发挥有效的治理效应，但这一治理效应仍存在提升的空间。并购后，上市公司面临特殊的并购风险，该风险对内部控制制度提出

了新的要求。上市公司激励董事参与内部控制制度建设，是提升内部控制制度治理效应的有效手段。

第三，内部控制对上市公司长期股票投资收益有显著的正向影响，但不存在真实盈余管理的中介路径、不受董事持股的异质性情境调节。该结论说明，内部控制对财务绩效与市场绩效的影响机理与适用情境并不相同：内部控制通过约束高管行为、提高整合能力等路径影响财务绩效，这一过程主要发生在企业内部；而内部控制基于信号传递效应影响市场绩效，这一过程主要发生在企业外部。

（二）管理启示

我们从以上研究结论得出如下管理启示。

第一，上市公司在并购后应当高度关注高管的真实盈余管理倾向，利用内部控制制度监督高管的机会主义行为。并购当年与并购后，上市公司应当做好“不相容职务分离”“授权批准”等内部控制的机制设计，将一项企业活动的授权、执行与监督等责任分配到不同岗位，并落实好各生产经营环节中内部控制制度的执行工作，以有效监督与约束高管的机会主义行为。

第二，上市公司应当充分利用董事持股的激励作用，促进并购后内部控制制度的重构。上市公司应当针对特殊的并购风险，进一步完善与之匹配的内部控制制度。在此过程中，上市公司应当充分重视董事的作用，采用股权激励等方案，鼓励董事关注并购后内部控制制度的修正与重构，促使内部控制制度发挥更有效的治理效应。

（三）局限与展望

本文从高管机会主义行为角度实证了内部控制影响长期并购绩效的中介路径问题，并讨论了董事持股的异质性情境状况，其局限及相关的研究展望主要归结如下：第一，本文验证了内部控制对上市公司长期股票投资收益有显著的正向影响，并指出这一影响中不存在真实盈余管理的中介路径，二者关系不受董事持股的异质性情境调节，但并未深入研究内部控制影响上市公司长期股票投资收益的内在机理与适用情境，未来的研究可以基于信号传递理论，进一步探究内部控制影响上市公司长期股票投资收益的具体机理与适用情境问题。第二，本文从高管行为角度验证了真实盈余管理在内部控制与长期并购绩效之间发挥的中介作用，但并未验证其他可能的中介路径，从并购整合等视角分析内部控制影响并购绩效的其他中介路径，是有待学术界未来深入探讨的一个重要课题。

参考文献

[1] 崔永梅，余璇. 基于流程的战略性并购内部控制评价研究［J］. 会计研究，2011（6）：57-62.

[2] 赵息，张西栓. 内部控制、高管权力与并购绩效——来自中国证券市场的经验证据［J］. 南开管理评论，2013（2）：75-81.

[3] Harp N L，Barnes B G. Internal Control Weaknesses and Acquisition Performance［J］. The Accounting Review，2018，93（1）：235-258.

[4] 张自巧，葛伟杰. 股份支付并购中存在不同的盈余管理吗？——来自中国上市公司的经验证据［J］. 证券市场导报，2013（1）：25-30.

[5] Cohen D A，Dey A，Lys D T Z. Real and Accrual-Based Earnings Management in the Pre-and Post-Sarbanes-Oxley Periods［J］. Accounting Review，2008，83（3）：757-787.

[6] 王福胜，吉姗姗，程富. 盈余管理对上市公司未来经营业绩的影响研究——基于应计盈余管理与真实盈余管理比较视角［J］. 南开管理评论，2014（2）：95-106.

[7] Ashbaugh-Skaife H，Collins D W，Kinney W R，LaFond R. The Effect of SOX Internal Control Deficiencies and Their Remediation on Accrual Quality［J］. Accounting Review，2008，83（1）：217-250.

[8] 董望，陈汉文. 内部控制、应计质量与盈余反应——基于中国2009年A股上市公司的经验证据［J］. 审计研究，2011（4）：68-78.

[9] 范经华，张雅曼，刘启亮. 内部控制、审计师行业专长、应计与真实盈余管理［J］. 会计研究，2013（4）：81-88，96.

[10] 胡明霞. 管理层权力、内部控制质量与盈余管理［J］. 重庆大学学报（社会科学版），2018（2）：71-81.

[11] 王嘉鑫，王永海. 强制性内部控制制度影响真实盈余管理吗？——基于中国上市公司的准自然实验研究［J］. 证券市场导报，2019（9）：20-30.

[12] 吕景胜，赵玉梅. 董事会特征对内部控制失效的影响研究——基于中国上市公司的数据分析［J］. 中国软科学，2016（5）：93-106.

[13] 江涛，陈富永，汤思禹. 基于“关系型”社会情境的董事网络对并购绩效影响研究［J］. 中国软科学，2019（11）：183-192.

[14] 杨道广，张传财，陈汉文. 内部控制、并购整合能力与并购业绩——来自我国上市公司的经验证据［J］. 审计研究，2014（3）：43-50.

[15] 李井林，刘淑莲，韩雪. 融资约束、支付方式与并购绩效［J］. 山西财经大学学报，2014（8）：114-124.

［16］孙梦男，吴迪. 中国上市公司并购中盈余管理的择机行为研究［J］. 云南财经大学学报，2017（6）：132-139.
［17］朱湘忆. 真实盈余管理、产权性质与创新绩效［J］. 中央财经大学学报，2020（5）：53-64.
［18］Kinney W R，Mcdaniel L S. Characteristics of Firms Correcting Previously Reported Quarterly Earnings［J］. Journal of Accounting & Economics，1989，11（1）：0-93.
［19］Bell T B，Carcello J V. A Decision Aid for Assessing the Likelihood of Fraudulent Financial Reporting［J］. Auditing A Journal of Practice & Theory，2000，19（1）：169-184.
［20］阎达五，杨有红. 内部控制框架的构建［J］. 会计研究，2001（2）：2-9.
［21］刘启亮，罗乐，张雅曼，等. 高管集权、内部控制与会计信息质量［J］. 南开管理评论，2013（1）：15-23.
［22］王晶，彭博，熊焰韧，张萍，张娟. 内部控制有效性与会计信息质量——西方内部控制研究文献导读及中国制度背景下的展望（一）［J］. 会计研究，2015（6）：89-97，99.
［23］王宗润，陈艳. 基于熵模型的内部控制质量与会计稳健性研究——来自我国金融行业上市公司的面板数据［J］. 中央财经大学学报，2014（1）：55-63.
［24］蒋勇，王晓亮. 会计稳健性、管理层防御与真实盈余管理［J］. 中央财经大学学报，2019（6）：58-69.
［25］韩岚岚，马元驹 . 内部控制对费用粘性影响机理研究——基于管理者自利行为的中介效应［J］. 经济与管理研究，2017（1）：131-144.
［26］Feng M，Li C，Mcvay S，Skaife H. Does Ineffective Internal Control over Financial Reporting Affect a Firm's Operations? Evidence from Firms' Inventory Management［J］. Accounting Review，2015，90（2）：529-557.
［27］Hauser，Roie. Busy Directors and Firm Performance：Evidence from Mergers［J］. Journal of Financial Economics，2018（1）：1-46.
［28］郑丽，陈志军. 企业扩张行为是否与董事会结构特征、激励方式有关——来自中国上市公司数据的实证检验［J］. 现代财经（天津财经大学学报），2018（9）：48-60.
［29］Kobelsky K，Lim J H，Jha R. The Impact of Performance-based CEO and CFO Compensation on Internal Control Quality［J］. Social Science Electronic Publishing，2013，29（3）：913-934.
［30］李增福，董志强，连玉君. 应计项目盈余管理还是真实活动盈余管理？——基于我国 2007 年所得税改革的研究［J］. 管理世界，2011（1）：129-142.
［31］Roychowdhury S. Earnings Management Through Real Activities Manipulation［J］. Journal of Accounting and Economics，2006，42（3）：0-370.
［32］胡凡，李科. 股价高估与商誉减值风险［J］. 财经研究，2019（6）：71-85.
［33］叶康涛，曹丰，王化成. 内部控制信息披露能够降低股价崩盘风险吗？［J］. 金融研究，2015（2）：192-206.

（责任编辑：郜　霖　张安平）

（上接第 69 页）

［17］Bebchuk L A，Fried J M，Walker D I. Managerial Power and Rent Extraction in the Design of Executive Compensation［J］. University of Chicago Law Review，2002，69（3）：751-846.
［18］权小锋，吴世农，文芳. 管理层权力、私有收益与薪酬操纵［J］. 经济研究，2010（11）：73-87.
［19］刘星，徐光伟. 政府管制、管理层权力与国企高管薪酬刚性［J］. 经济科学，2012（1）：86-102.
［20］陈修德. 区域市场化、管理层权力与高管薪酬契约［D］. 广州：华南理工大学，2012.
［21］王克敏，王志超. 高管控制权、报酬与盈余管理——基于中国上市公司的实证研究［J］. 管理世界，2007（7）：111-119.
［22］Jayaraman S，Milbourn T T. The Role of Stock Liquidity in Executive Compensation［J］. The Accounting Review，2012，87（2）：537-563.
［23］马晨，程茂勇，张俊瑞. 事务所变更能帮助公司提升审计质量么？——来自二次财务重述的经验证据［J］. 中国软科学，2014（10）：109-120.
［24］Lobo G J，Zhao Y. Relation between Audit Effort and Financial Report Misstatements：Evidence from Quarterly and Annual Restatements［J］. The Accounting Review，2013，88（4）：1385-1412.
［25］权小锋，吴世农. CEO 权力强度、信息披露质量与公司业绩的波动性——基于深交所上市公司的实证研究［J］. 南开管理评论，2010（4）：142-153.
［26］Ma C，Zhang J，Yang B. Financial Restatement and Auditor Dismissal［J］. China Journal of Accounting Studies，2015，3（3）：209-229.
［27］马晨，张俊瑞. 管理层持股，领导权结构与财务重述［J］. 南开管理评论，2012，15（2）：143-160.

（学术顾问：吴　溪，责任编辑：王克方　张安平）

污染就近转移的驱动力：环境规制抑或经济动机？

The Driving Forces of Pollution Transfer Nearby: Environmental Regulation or Economical Motivation?

钟 娟 魏彦杰

ZHONG Juan WEI Yan-jie

［摘 要］本文使用中国城市数据作为样本，借鉴地理重心的测度方法，探讨省域内污染就近转移的动态演化及其驱动因素。研究发现污染源企业正朝着省域内环境规制以及经济水平更高的地区聚集，中国省域内并未出现明显的污染就近转移，反而更多表现出转移粘性。进一步的探讨发现：造成这一现象的原因在于阻滞污染源企业迁移的力量大于推动其迁移的力量；产业集聚因素阻滞污染转移的作用非常强烈，这是造成中心城市难以摆脱污染困扰的主因；而环境规制和产业升级因素却表现出推动和阻滞污染转移的双向作用，其促进污染减排的效力并未得到有效和充分的发挥。因此，未来的环境治理工作应以中心城市为重点，依据城市类型安排好污染减排的层次性；同时，有必要通过加强环境规制的政策协调与配套、推动“规制+升级”的深层次去污模式，充分释放环境规制和产业升级促进污染减排的效力。

［关键词］污染转移 环境规制 产业集聚 产业升级

［中图分类号］F062.9 F062.2 ［文献标识码］A ［文章编号］1000-1549（2020）10-00115-14

Abstract: Based on the data of Chinese cities and the measurement method of geographical gravity centers, this paper studies the dynamic evolution and driving forces of pollution transfer nearby within the provinces. The results show that, the pollution source enterprises are converging to areas with stricter environmental regulations and higher economic level within the provinces. In China, there is no obvious pollution transfer nearby, but more transfer viscosity. Further studies show that, the reasons for these phenomena are that the power to block the transfer of pollution source enterprises is greater than the power to promote it. Industrial agglomeration plays a very strong role in blocking the transfer of pollution, which is the main reason that makes it difficult for central cities to solve the pollution problems. However, environmental regulation and industrial upgrading have two-way effects on promoting and blocking pollution transfer, and their effects on promoting pollution reduction have not brought into full and effective play. Therefore, for the pollution reduction, we should focus on the central cities and implement differentiation strategy according to types of cities. At the same time, it is necessary to give full play to the role of environmental regulation and industrial upgrading to promote pollution reduction.

Key words: Pollution transfer Environmental regulation Industrial agglomeration Industrial upgrading

［收稿日期］ 2019-12-30

［作者简介］ 钟娟（通讯作者），女，1980年6月生，安徽财经大学国际经济贸易学院副教授，经济学博士，研究方向为产业集聚与环境污染；魏彦杰，男，1973年11月生，安徽财经大学国际经济贸易学院副教授，经济学博士，研究方向为环境经济学。

［基金项目］ 教育部人文社会科学研究基金青年项目“基于产业结构和产业关联的产业集聚环境效应及其协调政策研究”（项目编号：15YJC790155）；安徽财经大学校级科研重点研究项目“产业内贸易的环境效应研究——基于异质性企业贸易理论”（项目编号：ACKYB19009）。

感谢匿名评审人提出的修改建议，笔者已做了相应修改，本文文责自负。

一、引言

环境规制是否引发污染转移存在争议。“污染避难所假说”认为，严格的环境规制将驱使污染产业向规制较为宽松的地区转移（Copeland 和 Taylor，2004[1]）。赞同者发现，这一现象在中国各区域间已有所显现，中西部地区正成为东部地区产业转移的“污染避难所”（林伯强和邹楚沅，2014[2]；彭文斌等，2014[3]）。但反对者认为，由于中国经济发展中产业集聚的极化效应以及围绕经济和开放中心的本地市场效应，污染产业跨区域转移会受到较强阻滞，污染源企业反而向环境规制严格的东部地区聚集，即“污染转移粘性”（陈秀山和徐瑛，2008[4]；耿文才，2015[5]）。事实上，环境规制能否在大尺度空间上引发“污染避难所”效应确实值得怀疑，因为跨区域的远距离迁移更多地受到迁出与迁入区域特征的影响（Weterings 和 Knoben，2013[6]），那些具有良好经济环境、更多可用外部资源以及雅各布斯（Jacobs）和马歇尔（Marshall）效应更明显的地区会对企业产生空间锁定（Knoben 和 Oerlemans，2008[7]；Knoben，2011[8]）。另外，Lee（2008）[9]、Conroy 等（2016）[10]针对美国企业国内迁移的实证研究也发现制造业企业很少跨州搬迁。

在此背景下，研究者开始将视角转向更小的空间尺度，探究污染产业是否在区域内的临近城市间就近转移。因为区域内发展水平、交通设施、技术能力与投资环境差异较小，此时决定企业迁移的因素主要来自内部属性（如企业规模、年龄等）而非区位影响（Weterings 和 Knoben，2013[6]；Ravulaparthy 等，2017[11]）。由于企业是有限理性的决策主体，对不确定性的风险厌恶导致其更倾向于短距离搬迁（Van Dijk 和 Pellenbarg，2000[12]）。因此，污染产业即便出于降低环境治理成本动机而重新选址，也更可能在较小的空间尺度上展开（沈坤荣等，2017[13]）。

不过，由于缺乏经验研究的支持，有关污染就近转移问题仍有许多待澄清之处。

一是“有、无”问题，即污染就近转移现象在中国是否存在？事实上，从“污染避难所假说”问世以来，污染转移问题一直存在争议。在国家间层面，有证据显示污染产业并未向规制水平更弱的国家大规模移动（Xing 和 Kolstad，2002[14]）；在中国区域间层面，研究者也发现污染产业在区域间的大规模转移并不明显，转移粘性现象更为突出（耿文才，2015[5]），尤其当研究者使用更为精确的省份投入产出表时，污染产业不断在东部地区集聚的现象也非常明显（王奇等，2017[15]）。2019 年 5 月，生态环境部发言人就否认了北京这样的大城市空气质量改善源于污染源就近或跨区域转移，因为京津冀及周边地区“2+26”城市、汾渭平原 11 城市①的空气质量存在相同变化趋势，并且由于这些城市空气优良天数、重度以上污染天数比例均有所增加，说明空气污染仍集中于中心城市，所谓污染转移的说法并不准确。

二是“驱动力”问题。无论区域内表现为污染就近转移还是转移粘性，企业搬迁决策必然受到了一系列因素的影响。例如“污染避难所假说”认为，严格的环境规制将驱使污染产业转移到环境规制较为宽松的地区，从而构成污染转移的推动力。但迁移决策也并非简单的成本计算，Eslamipoor 和 Sepehriar（2014）[16]发现，企业决策者最关心的是迁移带来的劣势，而不是环境规制所导致的成本压力或是迁入环境规制宽松地区的成本优势。一般认为，搬离经济集聚中心的迁移活动容易导致企业劣势，因为靠近集聚中心能带来 Jacobs 和 Marshall 式的外部性（Holl，2004a[17]；Holl，2004b[18]；Frenken 等，2007[19]）。因此，企业间由于集聚所导致的地理临近和关系亲近②将抑制企业迁移倾向（Kirat 和 Lung，1999[20]；Torre 和 Rallet，2005[21]；Carrincazeaux 和 Coris，2015[22]），从而构成污染转移的阻滞力。此外，有关空间产品生命周期理论（SPLC）的实证研究发现，产业升级也具有驱使企业迁移的作用：一方面，为了借助本地市场效应与集聚正外部性，处于产品竞争阶段的初创公司以及新兴产业大多集中于中心城市。其后，随着产业过渡到成熟的生命周期阶段，成本竞争将促使企业向外围迁移，以便从低工资、低地价以及宽松的环境规制中获益（Kronenberg，

① 京津冀及周边地区“2+26”城市、汾渭平原 11 城市是《2018—2019 年蓝天保卫战重点区域强化督查方案》（环环监［2018］48 号）重点防控区域，包含了京津冀大气污染传输通道的主要城市。

② 地理临近（geographical proximity）和关系亲近（relational proximity）是法国“接近性经济学派”（School of Proximity Economics）的核心概念，该学派认为地理临近以及随之而来的关系亲近对企业搬迁决策具有锚定效应。

2013[23]；Hong，2014[24]；Capasso 等，2016[25]；Yi，2018[26]）。另一方面，即使在同一产业中，技术进步带来的竞争压力也会导致低生产率企业从中心向外围地区迁移，因为低生产率企业对成本更为敏感（Okubo 和 Tomiura，2012[27]；Forslid 和 Okubo，2015[28]）。因此，如果产业升级导致经济增长逐渐摆脱对污染产业的依赖，或者产业升级导致污染产业内部出现“波特”式的创新竞争，污染产业（或企业）都可能被迫从中心城市向外围地区迁移，从而构成污染转移的另一种推动力。不过，由于缺乏经验证据的支持，这些因素在中国是否影响了污染就近转移？其作用方向、大小如何？目前尚无定论。

三是“经济含义”问题，即污染就近转移或转移粘性对中国可持续发展究竟有利还是有害？一方面，如果污染就近转移归因于“污染避难”，这种迁移活动无疑是有害的（沈坤荣 等，2017[13]）。但污染就近转移的驱动力也可能来自产业升级，Pumain 等（2006）[29]发现，随着法国经济增长的重心从高耗能、高污染产业转向信息技术及其相关部门，1960—2000 年间，化工、纺织、金属制品、纸浆和造纸等行业纷纷从大都市迁移到周边较小城市。同时，英国经验证据也显示，那些依据自身产品生命周期从高成本地区（拥挤、高工资和租金）转向周边低成本地区（规模较小、不那么密集）的企业，其搬迁后的生产率反而更高（Foreman-Peck 和 Nicholls，2015[30]）。因此，这种升级驱动型的迁移活动既是良性经济现象，也是企业面对成本竞争的理性行为。另一方面，即便企业对产业集聚的依赖阻滞了污染就近转移，其后果也未必有利。只要污染产业依旧是经济增长的引擎，伴随着大城市的集聚效益超过环境规制或拥挤效应的成本，更多企业将涌入大城市，并引发极为严重的环境问题，这一点在发展中国家尤其明显（Deichmann 等，2008[31]；Eslamipoor 和 Sepehriar，2014[16]）。当前问题在于：由于中国的污染就近转移可能受到多种驱动力的影响，它究竟体现为何种经济含义？需要何种对策安排？目前也是一个谜。

为探究上述问题，我们作如下处理：第一，借鉴地理重心的测度方法（徐建华和岳文泽，2001[32]；丁焕峰和李佩仪，2009[33]；赵小凤等，2010[34]），将污染就近转移表示为省域内“污染重心”偏离“经济-规制重心”的动态演化过程。这种量化处理的好处是可以直观地度量污染就近转移，不必通过邻近城市环境规制与本地污染排放之间的正相关关系①来间接理解这一问题。第二，将环境污染的关注点集中于污染物浓度而非污染物数量。因为有关环境库兹涅兹曲线（EKC）的研究发现，由于污染物数量隐含了规模因素，即使污染物浓度已开始从拐点下降，污染物排放量仍可能持续增加（Stern 等，1996[35]）。这会低估相关因素影响环境污染的实际效果，进而也可能低估污染转移的真实状况。第三，在前述基础上，围绕污染就近转移的驱动因素，并按照其在污染转移中的推动和阻滞两种作用，系统验证污染就近转移的动因及其隐含的经济意义。

总之，针对区域内污染源迁移方向、迁移原因的研究有助于深刻揭示中国环境污染问题的特性，并说明这种空间布局的改变究竟来自外在的环境规制压力还是企业内在的经济动机，进而引申出截然不同的政策含义。

二、污染就近转移问题：一个奇特的现象

污染转移反映了污染物空间分布的动态演变。现有研究集中于跨区域的污染转移，不过由于污染转移量难以测度，通常使用污染产业转移加以替代，研究方法包括区位熵等份额指标、区域间投入产出表、偏离份额法等。当然，也有少量文献尝试使用偏离份额法来测度区域间污染物排放转移情况（成艾华和赵凡，2018[36]），但其地理空间含义仍不够直观。

为克服上述问题，更准确测度污染就近转移，我们首先引入地理重心的概念，将省域内某一空间现象的差异性分布简化为一个均衡中心点，以便从时间维度上观察该重心点的移动。地理重心的计算方法为：假设某一省域由 n 个城市构成，其中第 i 个城市的中心坐标（即经纬度）为（x_i，y_i），其属性值为 m_i，则该属性的重心坐标可用公式（1）求算：

$$\bar{x}=\frac{\sum_{i=1}^{n} m_i x_i}{\sum_{i=1}^{n} m_i} \qquad \bar{y}=\frac{\sum_{i=1}^{n} m_i y_i}{\sum_{i=1}^{n} m_i} \tag{1}$$

我们选取每个城市 i 的三类属性“工业二氧化硫

① 沈坤荣等（2017）[13]的研究正是基于这种相关关系间接证明了环境规制在城市间引起了污染就近转移。

排放强度（WSO_2）、工业二氧化硫去除率（*SDRR*）、工业总产值（*IOV*）”分别构建地理重心点，并称作“污染重心、环境规制重心、经济重心”。其中：WSO_2 为工业二氧化硫排放量与工业总产值的比值；*SDRR* 为工业二氧化硫去除量与产生量的比值；*IOV* 使用分省份工业品出厂价格指数进行近似性折算，基期为 2000 年，以消除价格因素影响。由此，重心坐标（$\bar{x}$、$\bar{y}$）代表了省域内某一类属性 m 的一个均衡点，均衡点总是朝着省域内该属性不断强化的地区移动。

随后，我们借鉴空间产品生命周期理论（SPLC）的“核心—外围”思想来测度污染就近转移。第一步，我们尝试在省域内选定一个核心和参照点，称作地理空间上的“经济－规制”重心。方法如下：（1）假定省域内的环境规制重心总是靠近其经济重心，并且二者将不断靠拢（乃至重合）①。因为经济活动产生环境污染，随着污染危害的增强，经济活动会受到环境规制越来越强的约束。（2）如果经济重心与环境规制重心不断靠拢并接近重合，则选取二者之间的中心点作为“经济-规制”重心。该重心点总是朝着省域内环境规制以及经济水平更高的地区移动。第二步，我们将污染就近转移问题界定为“经济-规制”重心与污染重心之间相对距离的动态变化。原理如下：（1）由于“经济-规制”重心总是朝着省域内环境规制严格的经济中心区域移动，因此它将表现出推动污染转移（来自环境规制、产业升级）或阻滞污染转移（来自产业集聚）的双重力量，从而对污染重心产生“推”“拉”两种作用。（2）由此，二者之间相对距离的变化代表了不同的迁移模式，包括污染重心不断远离核心（即经济－规制重心）的“核心→外围”迁移，即污染就近转移现象；以及污染重心不断靠近核心的“外围→核心”迁移，即污染转移粘性现象。

围绕“经济-规制”重心的“核心—外围”迁移模式全面反映了多因素驱动下的污染转移问题。（1）环境规制对污染转移的影响。如果“污染避难所”效应真的存在，污染源企业将被迫从环境规制严格的地区向宽松地区迁移，污染重心与“经济-规制”重心的相对距离将增大，并表现出“核心→外围”迁移，即环境规制引发的污染转移，环境规制构成污染转移的推动力。（2）经济动机对污染转移的影响。一方面，如果“空间产品生命周期”效应足够强烈，产业升级会推动污染产业步入成本竞争的成熟生命周期阶段，其引发的创新竞争也会导致低效率污染企业面对更强的成本压力，进而驱使污染源企业逐渐远离经济中心区域，向外围的低成本地区移动。相应地，污染重心与“经济－规制”重心之间将表现为“核心→外围”迁移，即产业升级引发的污染转移，产业升级构成污染转移的另一种推动力。另一方面，如果集聚效应足够强烈，产业集聚的正外部性将吸引污染源企业不断靠近经济中心区域，并表现为污染重心与“经济-规制”重心相对距离不断缩小的“外围→核心”迁移，即产业集聚引发的转移粘性，产业集聚构成污染转移的阻滞力。总之，无论污染转移的驱动力来自环境规制还是经济动机，我们都可以全面观察到这些驱动机制的综合作用结果。

基于此，我们做如下处理：第一，选择《中国城市统计年鉴》的城市范围作为测度基础，并剔除 2005—2016 年间撤并或新增的几个城市。第二，去除城市数据严重缺失的少数省份，如西藏、台湾等。第三，为解决单个直辖市无法测度其地理重心之间相对距离的问题，将北京和天津并入河北、上海并入江苏、重庆并入四川。第四，分年份计算各省份污染重心与“经济-规制”重心的相对距离 *DPER*，并按照重要性原则，剔除 2005—2016 年间 *DPER* 变化很小的省份②，获得 20 个省份最终样本的 $DPER_{jt}$（$j=20$；t 为 2005—2016）。其中 6 个省份为东部沿海地区，14 个省份为其他地区。第五，分年份计算 20 个省份经济重心与环境规制重心的相对距离 DER_{jt}，观察二者是否不断接近（乃至重合），以验证前文的假设。由此，2005—2016 年 DER_{jt}、$DPER_{jt}$ 的年平均值 *DERavg*、*DPERavg* 见图 1。

① 为保证假设的合理性，我们在下文中测度了省域内两类重心之间相对距离的变化。2005—2016 年，按照本文选取的省域样本，环境规制重心与经济重心之间相对距离的平均值从 60.24KM 快速缩小到 26.58KM，二者之间存在靠拢并重合的趋势。

② 剔除方法是计算各省份 *DPER* 变化率绝对值的平均值，并使用 Fisher 最优分割法将其区分为“变化很小、中等变化、变化很大”三种类型，并剔除掉“变化很小”的省份，如新疆、青海、宁夏等。出现这种情况的原因在于这些省份的经济活动过度集中于某一中心城市，导致 *DPER* 在观察时间段内几乎不发生变化，从而失去研究意义。

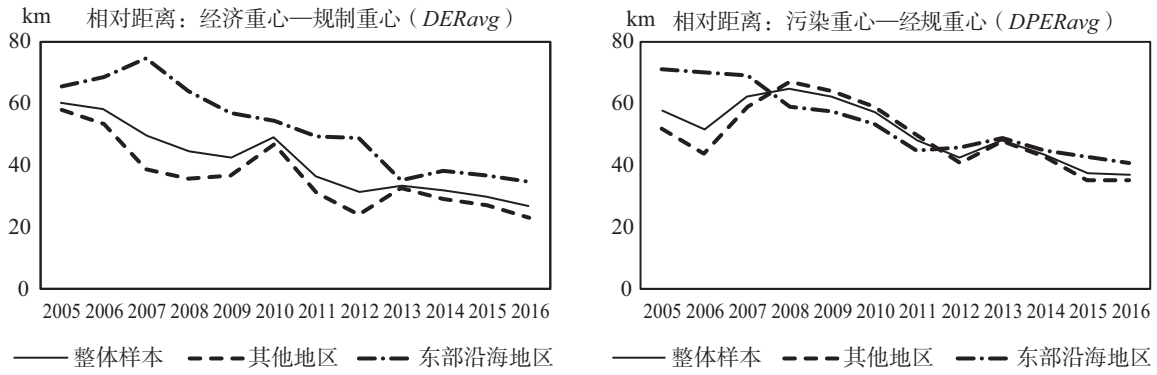

图 1 2005—2016 年 *DERavg*、*DPERavg* 变化情况

注：20 个样本省份被区分为东部沿海地区（包括 6 个省份，河北（含北京、天津）、山东、江苏（含上海）、浙江、福建、广东）和其他地区（包含 14 个省份，黑龙江、吉林、辽宁、河南、安徽、湖北、湖南、江西、广西、山西、陕西、内蒙古、四川（含重庆）、云南）。图中标题的“规制重心”“经规重心”分别是“环境规制重心”“经济-规制重心”的简称。

图 1 左图显示，2005—2016 年，经济重心与环境规制重心相对距离的年平均值 *DERavg* 在整体样本、东部沿海以及其他地区中均快速下降（整体样本的 *DERavg* 从 60. 24km 缩小至 26. 58km，已非常接近）。这与前文假设一致，经济活动受到越来越强的环境规制压力。但与此同时，污染重心与“经济-规制”重心相对距离的年平均值 *DPERavg*（见图 1 右图）也在下降（2005—2016 年，整体样本的 *DPERavg* 从 57. 55km 缩小至 36. 83km），*DPERavg* 并未表现出“核心→外围”迁移的污染转移，反而更多地表现出“外围→核心”迁移的转移粘性，污染源企业正朝着省域内环境规制以及经济水平更高的地区聚集，这一趋势在东部沿海以及其他地区都非常明显。我们推测，这一奇特现象说明“经济-规制”重心对污染重心的“推”“拉”作用发生了意想不到的变化。也就是说，由于集聚效应非常强烈，它超过了“污染避难所”效应和“空间产品生命周期”效应，这使得阻滞污染转移（来自产业集聚）的力量大于推动污染转移（来自环境规制、产业升级）的力量，从而将污染重心从外围地区不断拉近至“经济-规制”重心（即核心）。

为佐证这一点，我们也测度了污染重心与环境规制重心相对距离 DPR_{ji} 的年平均值 *DPRavg*、污染重心与经济重心相对距离 DPE_{ji} 的年平均值 *DPEavg*。如图 2 左图所示，*DPRavg* 表现出明显的下降（整体样本的 *DPRavg* 从 2005 年的 55. 62km 缩小至 2016 年的 43. 34km），污染重心不断靠近环境规制重心而不是远离它，环境规制作为污染转移“推动”力的作用并未得到发挥。如图 2 右图所示，*DPEavg* 也表现出下降趋势（整体样本的 *DPEavg* 从 2005 年的 72. 82km 缩小至 2016 年的 53. 32km），污染重心正朝着经济重心不断靠拢，这一现象说明产业集聚阻滞污染转移的作用非常明显，产业升级推动污染转移的作用则不明显，经济重心的集聚效应显然超过了其“空间产品生命周期”效应。

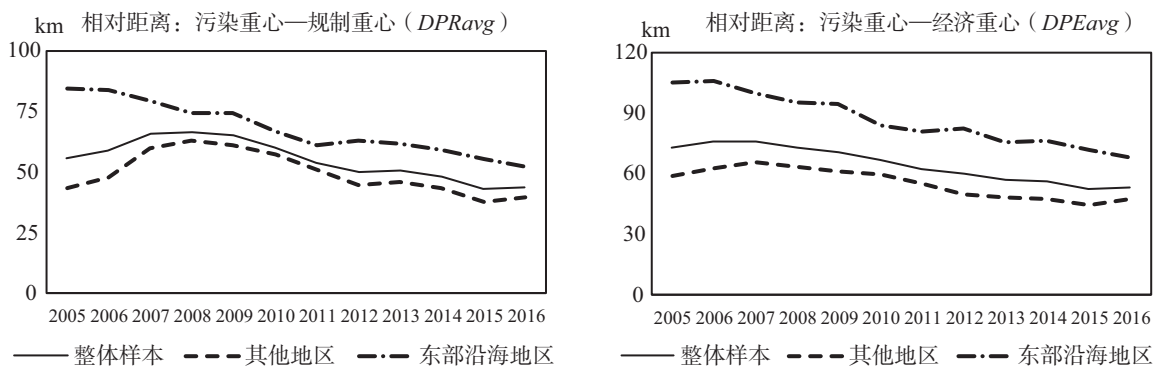

图 2 2005—2016 年 *DPRavg*、*DPEavg* 变化情况

总之，针对 2005—2016 年间中国 20 个省份的研究发现，*DPERavg*、*DPRavg*、*DPEavg* 的变化都反映出同一种现象：就多因素驱动的污染就近转移问题而言，阻滞污染转移的力量明显大于推动污染转移的力量。据此，我们提出以下假设，以实证检验其是否成立。

假设 1：*DPERavg* 之所以表现出“外围→核心”迁移的转移粘性，就在于阻滞污染转移（来自产业集聚）的力量大于推动污染转移（来自环境规制、产业升级）的力量。

假设 2：*DPERavg* 的变化说明：由于“经济-规制”重心的集聚效应更为强烈且稳定，产业集聚因素显著地阻滞了污染转移。

假设 3：虽然严格的环境规制可能引发“污染避难所”效应，但 *DPERavg* 的变化说明：环境规制因素对污染转移的影响可能并非单向作用，其推动力显然受到了某种抑制。

假设 4：虽然产业升级可能引发“空间产品生命周期”效应，但 *DPERavg* 的变化说明：产业升级因素对污染转移的影响可能并非单向作用，其推动力同样受到了某种抑制。

三、计量模型、变量与数据

（一）计量模型的设定

为验证上文假设，我们构建计量模型，实证检验多种驱动因素对污染就近转移的影响。考虑到省域内污染就近转移情况可能具有持续性特征，使用以下动态面板模型加以估计：

$$DPER_{jt}=\alpha_0+\alpha_1 L.DPER_{jt}+\alpha_2 ER_{jt}+\alpha_3 IA_{jt}+\alpha_4 IU_{jt}+\beta Z_{jt}+\varepsilon_t \quad (2)$$

其中：下标 j 和 t 分别代表省份和年份；被解释变量 $DPER_{jt}$ 反映了省域内的污染就近转移情况；解释变量 $L.DPER_{jt}$、ER_{jt}、IA_{jt}、IU_{jt}、Z_{jt} 分别表示被解释变量滞后一期、环境规制因素、产业集聚因素、产业升级因素和控制变量；ε_t 为随机扰动项。

（二）变量的指标选取与度量

（1）污染转移 $DPER_{jt}$。该变量反映污染就近转移情况，测度方法如上文所示，表示为省份 j 内污染重心与“经济-规制”重心的相对距离。当相对距离增大，表示存在污染就近转移现象；当相对距离减小，表示存在污染转移粘性现象。我们在回归时做了取对数处理。

（2）环境规制 *ER*。环境规制构成污染转移的推动力，其机制来自“污染避难所”效应，并迫使污染源企业从“核心→外围”迁移。衡量标准为城市 i 的工业二氧化硫去除率。

为了全面反映省域内环境规制的变化，我们选取两个指标：第一，规制水平差异 ERd_{jt}。假设省份 j 由 n 个城市构成，其中第 i 个城市在年份 t 的工业二氧化硫去除率为 ER_{it}（其中，$i=1, 2, \cdots, n$），则 ER_{it} 的基尼系数代表规制水平差异，即省份 j 内不同城市间环境规制水平的离散程度。由于环境规制水平差异 ERd_{jt} 可能引发污染源企业从规制严格地区向宽松地区迁移，因此预测 ERd_{jt} 与污染转移呈正相关关系。第二，平均规制水平 ERa_{jt}。考虑到环境规制的变化也表现为水平值的变化，因此选取 ER_{it}（$i=1, 2, \cdots, n$）在年份 t 的平均值作为省份 j 的平均规制水平。我们预测，如果平均规制水平 ERa_{jt} 的增加主要来自严格环境规制城市更快地提升其规制水平，那么城市间规制水平差异将增大，“污染避难所”效应变得强烈，ERa_{jt} 与污染转移之间呈正相关关系；当然，平均规制水平 ERa_{jt} 的增加也可能来自环境规制宽松城市更快地提升其规制水平，城市间规制水平差异反而缩小，此时 ERa_{jt} 与污染转移之间的关系将变得不确定。

（3）产业集聚 *IA*。产业集聚构成污染转移的阻滞力，其机制来自集聚效应，并促使污染源企业从“外围→核心”迁移。衡量标准为城市 i 工业总产值（百万元）占全国工业总产值（亿元）的比重。

为了全面反映省域内产业集聚的变化，我们选取两个指标：第一，集聚水平差异 IAd_{jt}。假设省份 j 由 n 个城市构成，其中城市 i 在年份 t 的产业集聚水平为 IA_{it}（$i=1, 2, \cdots, n$），则 IA_{it} 的基尼系数代表集聚水平差异，即省份 j 内不同城市间产业集聚水平的离散程度。随着城市间集聚水平差异 IAd_{jt} 的增大，更多污染源企业将涌入中心城市，因此预测 IAd_{jt} 与污染转移呈负相关关系。第二，平均集聚水平 IAa_{jt}。考虑到产业集聚的变化也表现为水平值的变化，因此选取 IA_{it}（$i=1, 2, \cdots, n$）在年份 t 的平均值作为省份 j 的平均集聚水平。一般而言，如果某个省份的平均集聚水平 IAa_{jt} 更高，则其中心城市的工业发展更充分、工业行业门类也更富多样性，集聚正外部性也更加明显，这类似于一种规模经济，此时 IAa_{jt} 与污

染转移将呈负相关关系；当然，随着平均集聚水平 IAa_{jt} 的增加，中心城市也可能受制于拥挤和高成本，这类似于一种规模不经济，其集聚能力反而边际递减，此时 IAa_{jt} 与污染转移之间的关系将变得不确定。

（4）产业升级 *IU*。产业升级构成污染转移的另一种推动力，其机制来自“空间产品生命周期”效应。按照 SPLC 理论，产业升级推动污染转移表现为两个方面：第一，产业升级将导致工业的细分行业结构发生变化，工业增长对重污染行业的依赖性降低，污染行业被迫从“核心→外围”迁移。因此，我们选取轻度污染行业产值占工业总产值的比重作为衡量指标，简称为行业结构 IS_{jt}，预测 IS_{jt} 与污染转移呈正相关关系。测度方法如下：首先计算 2005—2016 年间全国工业各细分行业二氧化硫排放强度的平均值 θ，并使用 Fisher 最优分割法将其区分为“极重度污染行业（$\theta \geqslant 0.062\ 4$）、重度污染行业（$0.027\ 2 \leqslant \theta \leqslant 0.044\ 8$）、中度污染行业（$0.008\ 3 \leqslant \theta \leqslant 0.021\ 4$）、轻度污染行业（$\theta \leqslant 0.004\ 9$）”四种类型。随后，我们将轻度污染行业中的石油天然气开发、供水、废旧回收、烟草制造 4 个行业剔除，得到 13 个轻度污染行业名单（见表 1）。最后，根据 13 个细分行业名单，使用 20 个样本省份工业细分行业产值数据，计算出分年份、分省份行业结构 IS_{jt}。第二，产业升级也将引发技术创新，这会导致低效率的污染企业面对更强的成本压力，进而被迫从“核心→外围”迁移。因此，我们使用分省份的 R&D 投入强度作为衡量指标，简称为技术创新 RD_{jt}，预测 RD_{jt} 与污染转移之间呈正相关关系。计算方法为 R&D 经费内部支出（万元）与 R&D 人员全时当量（人年）的比值，即单位全时当量的 R&D 经费投入，并取其对数。其中，R&D 活动的范围包括研究与开发机构、企业、高等学校和其他。之所以选择省份 j 在年度 t 的全部 R&D 活动作为计算基础，是因为工业部门的技术创新不仅仅只受到其自身 R&D 活动的影响。

表 1　　轻度污染行业名单

二氧化硫排放强度 θ（单位：吨/十万元）	分类	细分行业名称
$\theta \leqslant 0.004\ 9$	轻度污染行业	橡胶和塑料制品业、金属制品业、皮革毛皮羽毛（绒）及其制品业、通用设备制造业、专用设备制造业、纺织服装鞋帽制造业、仪器仪表及文化办公用机械制造业、交通运输设备制造业、印刷和记录媒介复制业、家具制造业、文教工美体育和娱乐用品制造业、电气机械及器材制造业、通信设备计算机和其他电子设备制造业

注：表 1 中细分行业名单顺序按照污染强度降序排序。其中，剔除掉的 4 个行业分别为“石油和天然气开采业”“水的生产和供应业”“烟草制品业”“废弃资源和废旧材料回收加工业”，这些行业与生产生活服务、禀赋资源开发等特定事项有关，对地区工业发展不具有普遍意义。

在计算 2005—2016 年间的 θ 值时，为保证各细分行业前后一致，我们将 2005—2011 年的“橡胶制品业”“塑料制品业”做合并处理，以便与 2012—2016 年间的“橡胶和塑料制品业”对应；将 2011—2016 年的“汽车制造业”“铁路、船舶、航空、航天和其他运输设备制造业”做合并处理，以便与 2005—2010 年间的“交通运输设备制造业”对应。

（5）控制变量。为保证估计结果稳健，本文选取如下控制变量：第一，出口贸易 EX_{jt}。针对贸易与环境的研究发现，中国更倾向于出口清洁产品（李小平和卢现祥，2010[37]）。由此我们推测，对于那些出口依存度很高的地区，这一贸易模式可能改变其行业结构，导致工业增长更偏重于低污染行业。贸易的这种间接作用类似于 SPLC 理论所谓的产业升级，因此预测 EX_{jt} 与污染转移呈正相关关系。测度方法为省份 j 出口额（亿元）与 GDP（亿元）的比值。其中，出口额按照汇率年平均价折算为人民币；GDP 按平减指数换算为不变价 GDP，基期为 2000 年。第二，外商直接投资 FDI_{jt}。“污染避难所”假说认为，跨国公司会通过 FDI 将高污染产业或生产环节转移至发展中国家，部分中国经验研究也支持这一观点（张宇、蒋殿春，2014[38]；杨子晖和田磊，2017[39]）。因此，如果 FDI 在省域内的选址偏向中心城市，它将导致污染源企业从“外围→核心”聚集，FDI_{jt} 与污染转移呈负相关关系；如果偏向非中心城市，它将导致污染源企业从“核心→外围”扩散，FDI_{jt} 与污染转移呈正相关关系。测度方法为省份 j 内外商投资总额（亿元）与 GDP（亿元）的比值。其中，外商投资总额按照汇率年平均价折算为人民币；GDP 按平减指数换算为不变价 GDP，基期为 2000 年。第三，环境压力差异 ESd_{jt}。尽管经济活动的空间集聚带来正外部性，但它也具有负面影响，例如环境恶化压力（Eslamipoor 和 Sepehriar，2014[16]）。与集聚导致的拥挤效应类似，集聚所引发的环境污染也会降低福利，并产生负外部性，进而迫使污染源企业从“核心→外

围”迁移（Foreman-Peck 和 Nicholls，2015[30]），所以预测 ESd_{jt} 与污染转移呈正相关关系。测度方法为：假设省份 j 由 n 个城市构成，其中城市 i 在年份 t 的环境压力指标 ES_{it}（$i=1, 2, \cdots, n$）等于工业二氧化硫排放量（吨）与总人口（千人）的比值，则 ES_{it} 的标准差代表环境压力差异，并取其对数。更大的 ESd_{jt} 意味着省域内的部分城市承受着远高于其他城市的环境压力。第四，行业景气 IC_{jt}。工业行业景气水平也可能影响污染转移，因为“经济-规制”重心总是朝着省域内环境规制严格的经济中心区域移动，这一区域的环境治理成本、工资和租金更为昂贵，只有当工业行业处于繁荣周期，污染源企业才有可能承受从“外围→核心”迁移所带来的高成本，因此预测 IC_{jt} 与污染转移呈负相关关系。测度方法为各省份工业生产者出厂价格指数的定基指数，基期为 2000 年。

（三）数据来源与变量描述性统计结果

本文数据来源主要包括《中国城市统计年鉴》《中国统计年鉴》《中国环境统计年鉴》《中国工业统计年鉴》《中国科技统计年鉴》、各省和城市统计年鉴等。

面板数据共包括 20 个样本省份，时间从 2005—2016 年，描述性统计结果见表 2。

表 2　　变量的描述性统计结果（$N=20$，$T=12$，$NT=240$）

变量		含义	均值	标准差	最小值	最大值
$DPER_{jt}$		省份内的污染就近转移情况	3.708	0.711	0.909	5.790
$L.DPER_{jt}$		被解释变量 $DPER_{jt}$ 的滞后一期	3.735	0.705	0.909	5.790
ER	ERd_{jt}	省份内不同城市间的环境规制水平差异	0.253	0.137	0.015	0.633
	ERa_{jt}	省份内的平均环境规制水平	0.416	0.190	0.050	0.841
IA	IAd_{jt}	省份内不同城市间的产业集聚水平差异	0.372	0.116	0.147	0.611
	IAa_{jt}	省份内的平均产业集聚水平	0.418	0.341	0.081	1.551
IU	IS_{jt}	该省份的工业细分行业结构	0.329	0.147	0.076	0.655
	RD_{jt}	该省份的 R&D 投入强度	3.120	0.398	1.921	3.869
EX_{jt}		该省份的出口依存度	0.333	0.366	0.039	1.680
FDI_{jt}		该省份的 FDI 依存度	0.648	0.487	0.140	2.992
ESd_{jt}		省份内不同城市间的环境压力差异	2.617	0.480	1.078	3.603
IC_{jt}		该省份的工业行业景气水平	1.283	0.202	1.005	1.959

四、污染就近转移问题：回归分析与稳健性检验

（一）回归结果分析

首先，我们针对动态模型（见公式 2）可能存在的内生性问题作如下处理：由于被解释变量 $DPER_{jt}$ 表示省域 j 内污染重心与“经济-规制”重心的相对距离，这一距离的远近实际上也可以理解为该省份内经济与环境规制水平较高区域所具有的污染排放特征，如果距离较近，则污染程度较重；如果距离较远，则污染程度较轻。而污染物排放作为生产过程的副产品，它并非生产目的，因此 $DPER_{jt}$ 与产业集聚 IA、产业升级 IU 等因素之间仅是单向的投入产出关系。但是，$DPER_{jt}$ 与环境规制因素 ER 之间却可能互为因果，规制会影响污染排放，污染水平也会影响规制。为防范这一问题，我们在模型回归中将环境规制因素 ER 作为内生变量。其次，为防止伪回归现象，使用 LLC 方法对各变量进行单位根检验，P 值显示均在 5% 置信水平上拒绝原假设，所有变量均为平稳。由此，使用系统 GMM 方法对动态模型进行估计，估计结果使用稳健标准误，表 3。

表 3　　动态面板模型估计结果

解释变量	被解释变量：省域内污染就近转移情况 $DPER_{jt}$					
	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
$L.DPER_{jt}$	0.409***	0.495***	0.479***	0.459***	0.405***	0.495***
	（7.10）	（5.16）	（4.16）	（6.05）	（4.12）	（7.25）

续前表

解释变量		被解释变量：省域内污染就近转移情况 $DPER_{jt}$					
		模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
ER	ERa_{jt}	−1. 524 (−1. 43)	−1. 543 * (−1. 74)	−1. 327 * (−1. 75)			
	ERd_{jt}				1. 062 *** (2. 73)	2. 531 ** (2. 41)	1. 323 *** (3. 84)
IA	IAa_{jt}	−0. 541 ** (−2. 55)		−0. 535 *** (−2. 77)	−0. 425 ** (−2. 15)		−0. 325 ** (−2. 18)
	IAd_{jt}		−1. 006 ** (−2. 13)	−0. 891 ** (−2. 16)		−0. 458 * (−1. 66)	−0. 403 * (−1. 66)
IU	IS_{jt}	−1. 381 ** (−2. 03)	−0. 871 (−1. 58)	−0. 735 (−1. 50)	−0. 745 * (−1. 85)	−1. 225 ** (−2. 37)	−0. 604 * (−1. 66)
	RD_{jt}	0. 489 * (1. 79)	0. 313 (1. 29)	0. 479 ** (1. 97)	0. 321 * (1. 87)	0. 440 * (1. 66)	0. 325 ** (2. 06)
EX_{jt}		0. 564 *** (3. 53)	0. 355 ** (2. 16)	0. 498 *** (3. 64)	0. 389 *** (3. 58)	0. 392 ** (2. 27)	0. 383 *** (3. 47)
FDI_{jt}		0. 047 (0. 29)	−0. 030 (−0. 23)	0. 071 (0. 53)	0. 011 (0. 08)	−0. 099 (−0. 89)	−0. 034 (−0. 29)
ESd_{jt}		0. 572 *** (4. 76)	0. 515 *** (4. 62)	0. 565 *** (4. 98)	0. 615 *** (4. 78)	0. 655 *** (4. 93)	0. 611 *** (4. 87)
IC_{jt}		−0. 520 (−1. 31)	−0. 316 (−0. 96)	−0. 367 (−1. 27)	0. 025 (0. 13)	0. 110 (0. 42)	0. 062 (0. 32)
常数项		0. 859 (0. 99)	1. 183 * (1. 75)	0. 426 (0. 59)	−0. 639 (−1. 04)	−1. 127 (−1. 04)	−0. 667 (−1. 13)
AR（1）*P* 值		0. 008	0. 023	0. 029	0. 016	0. 044	0. 019
AR（2）*P* 值		0. 742	0. 981	0. 910	0. 931	0. 952	0. 965
Sargan *P* 值		0. 283	0. 236	0. 234	0. 606	0. 310	0. 242
省份 FE		是	是	是	是	是	是
年份 FE		是	是	是	是	是	是
样本量		240	240	240	240	240	240

注：***、**、*分别表示在 1%、5%、10%的水平上显著。下同。

从表 3 结果看，各类驱动因素对污染就近转移的影响基本符合上文中 4 个假设的预定结论，$DPER_{jt}$的变化正是这些驱动机制综合作用的结果。

（1）同假设 2 的预期一致，产业集聚因素 *IA* 阻滞污染转移的作用非常强烈且稳定。无论平均集聚水平 IAa_{jt}还是集聚水平差异 IAd_{jt}，它们都与污染转移 $DPER_{jt}$呈现显著的负相关关系。这说明集聚效应引发对企业的空间锁定，它是导致污染转移粘性现象的主因。一方面，从模型 2、3 和模型 5、6 的结果看，当城市间集聚水平差异 IAd_{jt}增大，中心城市的吸引力增强，污染源企业从“外围→核心”迁移，$DPER_{jt}$减小。另一方面，从模型 1、3 和模型 4、6 的结果看，随着省域内平均集聚水平 IAa_{jt}的增长，$DPER_{jt}$也会减小。按照 IAa_{jt}变量设定时的讨论，这说明各省份中心城市的集聚效应增长仍处于一种类似规模经济的阶段，省域内更高的平均集聚水平带来中心城市更充分的工业发展和更丰富的工业行业类型，其集聚效应也更强，进而吸引污染源企业从“外围→核心”迁移。

（2）同假设 3 的预期一致，环境规制因素 *ER* 对污染转移的影响存在正反两方面的双向作用，其推动力受到了抑制。可以发现，规制水平差异 ERd_{jt}与污染转移 $DPER_{jt}$呈现显著的正相关关系（见模型 4、5、6），城市间环境规制水平差异的增大引发“污染避难所”效应，污染源企业被迫从“核心→外围”迁移、向环境规制相对宽松的地区移动，$DPER_{jt}$随之变大。但与此同时，平均规制水平 ERa_{jt}与污染转移 $DPER_{jt}$却呈现显著的负相关关系（见模型 1、2、3），省域内平均规制水平的增长反而阻滞了污染转移，并

导致 $DPER_{jt}$ 减小。按照 ERa_{jt} 变量设定时的讨论，这说明观察期内 ERa_{jt} 的增加更多地来自省域内环境规制宽松城市更快的规制水平增长，从而城市间规制水平差异缩小。此时，污染源企业作为理性的决策主体，如果在外围地区也要面对快速增长的环境规制压力，其最优决策必然是向“经济-规制”中心靠拢，以便借助“外围→核心”迁移从集聚的正外部性中受益。

（3）同假设 4 的预期一致，产业升级因素 *IU* 对污染转移的影响也存在双向作用，其推动力受到了抑制。一方面，技术创新 RD_{jt} 与污染转移 $DPER_{jt}$ 呈现显著的正相关关系，技术创新带来竞争压力，这迫使低效率的污染企业从“核心→外围”迁移，以便从外围地区的相对低成本中获益，$DPER_{jt}$ 随之变大。但另一方面，行业结构 IS_{jt} 与污染转移 $DPER_{jt}$ 却呈现显著的负相关关系，工业增长对重污染行业的依赖性降低反而阻滞了污染转移，并导致 $DPER_{jt}$ 减小。这明显有悖于 SPLC 理论。一个可能原因是行业结构与污染转移之间具有类似 EKC 曲线那样的非线性关系：在 IS_{jt} 增长初期，轻度污染行业的兴起导致出现更多细分行业，集聚中心的 Jacobs 效应增大，其对污染源企业的吸引力反而增强，污染产业从“外围→核心”迁移；只有当超越某个拐点后，IS_{jt} 的增长才可能驱使污染产业远离集聚中心，从“核心→外围”迁移。从 2005—2016 年行业结构的数值来看，样本省份的均值确实不高，多数省份（12 个省份）IS_{jt} 值在 25% 至 40% 之间，其余省份则处于高、低两极，广东、江苏、福建、浙江在 50% 左右，黑龙江、山西、内蒙古、云南在 10% 左右。也就是说，由于观察期内中国样本省份行业结构的绝对水平偏低，其与污染转移的关系可能仍未超越拐点，因此二者之间呈负相关关系。当然，探究这一非线性关系超出了本文研究目的，仍有待后续研究加以验证。

（4）总体来看，同假设 1 的预期一致，阻滞污染转移（来自产业集聚）的力量大于推动污染转移（来自环境规制、产业升级）的力量。在产业集聚因素 *IA* 中，无论平均水平还是水平差异，它们都具有阻滞污染转移的显著性作用；但在环境规制因素 *ER* 中，规制水平差异和平均规制水平分别表现出推动和阻滞污染转移的双向作用，环境规制促使污染转移的能力受到抑制；产业升级因素 *IU* 的情况类似，技术创新和行业结构也分别具有推动和阻滞污染转移的双向作用，产业升级促使污染转移的能力同样受到了抑制。

（5）此外，从表 3 结果看，上一期的 $L.DPER_{jt}$ 显著正向影响本期的 $DPER_{jt}$，中国省域内的污染就近转移现象具有较强的动态效应。在控制变量方面，出口贸易 EX_{jt} 显著正向影响污染转移 $DPER_{jt}$，出口依存度的提升能够促进污染就近转移，迫使污染源企业从“核心→外围”迁移；外商直接投资 FDI_{jt} 与污染转移 $DPER_{jt}$ 之间则没有显著性关系；环境压力差异 ESd_{jt} 也显著正向影响污染转移 $DPER_{jt}$，集聚所引发的环境压力也会迫使污染源企业从“核心→外围”迁移；行业景气 IC_{jt} 与污染转移 $DPER_{jt}$ 之间没有显著性关系，工业行业景气水平的提升或下降并不具有阻滞或推动污染就近转移的作用。

表 3 中各类驱动因素的回归结果反映出其独特的经济含义，结合图 1 中 *DPERavg* 表现出的污染转移粘性现象，它们一起揭示出中国环境问题的长期性和复杂性。

（1）产业集聚因素阻滞了污染转移，它是造成中国省域内污染转移粘性现象的主因。集聚正外部性会吸引污染源企业靠近经济集聚中心。虽然产业集聚也可能引发技术进步与结构优化，从而提升区域内的污染减排能力（Copeland 和 Taylor，2004[1]），但集聚也带来规模效应，规模扩大导致的污染增加会削弱产业集聚的有利环境作用。从中国目前的情况看，集聚效应将导致相当程度的环境污染长期滞留于中心城市，并带来严峻的环境问题。

（2）环境规制因素同时表现出推动和阻滞污染转移的双向作用，这导致其经济含义相对模糊。一方面，规制水平差异推动污染转移，这会在一定程度上（但不是根本上）缓解经济集聚中心的环保压力，但人为地增大城市间规制水平差异显然不是环境政策的应有之义。另一方面，平均规制水平则阻滞污染转移，虽然区域内平均规制水平的提升是环保监管的必然要求，但这一趋势也可能导致集聚中心更难以摆脱污染的困扰。就中国目前的发展阶段而言，环境规制因素促进可持续发展的政策效果可能未必如预想的那样乐观。

（3）产业升级因素也具有推动和阻滞污染转移的双向作用，技术创新推动污染转移，行业结构则阻滞污染转移。事实上，Pumain 等（2006）[29] 发现，环境问题的根本改善有赖于行业结构优化发挥其双重“挤出”作用：首先是工业部门中低污染行业逐渐

“挤出”重污染行业，随之而来的是污染源企业被从集聚中心（核心）“挤出”到周边地区。不过，从中国目前情况看，行业结构绝对水平偏低，而且行业结构优化反而引发污染源企业从“外围→核心”聚集。这一有悖于 SPLC 理论的现象说明：产业升级因素促进中国可持续发展的效力并未充分释放，污染问题的真正解决仍是一个长期经济过程，并非短期规制能够一蹴而就。

（二）稳健性检验

为增强研究结论的可靠性，我们对各类驱动因素中实证结果与理论预期差异较大的变量进行稳健性检验。同时，鉴于产业集聚因素 IA 阻滞污染转移的作用非常一致，我们在稳健性检验模型中不再区分其水平差异和平均水平，而是构造一个新指标 $IAad_{jt}$ 来代表 IA，其计算方法为 $IAa_{jt}\times IAd_{jt}$，用以综合反映水平差异和平均水平两种效应。

（1）放宽 IS_{jt} 的行业选择范围。在表 3 中，行业结构显著负向影响污染转移，这明显有悖于 SPLC 理论的预期。为此，我们尝试放宽其行业选择范围，使用轻度与中度污染两类行业产值占工业总产值的比重作为新的衡量指标，称为行业结构 IS^{*}_{jt}。事实上，从表 1 的轻度污染行业名单看，其范围确实更偏重技术含量较高的机电、通信、运输设备、计算机等产品，少数东部沿海省份（广东、江苏、福建、浙江）与其他省份在该行业名单下的 IS_{jt} 值差异较大，同时全部样本省份 2005—2016 年间的 IS_{jt} 平均值也仅为 0.329，其绝对水平明显偏低。因此，相对狭窄的行业名单可能不够贴近中国产业发展的现实状况、低估了多数省份产业升级的努力程度。中度污染行业名单见表 4，其中我们剔除了黑色金属矿采选业、燃气生产和供应业、煤炭开采和洗选业 3 个行业。

表 4　　中度污染行业名单

二氧化硫排放强度 θ（单位：吨/十万元）	分类	细分行业名称
$0.008\ 3\leqslant\theta\leqslant 0.021\ 4$	中度污染行业	饮料制造业、食品制造业、纺织业、木材加工及木竹藤棕草制品业、医药制造业、工艺品及其他制造业、农副食品加工业

注：表 4 中细分行业名单顺序按照污染强度降序排序。其中，剔除掉的 3 个行业为黑色金属矿采选业、燃气生产和供应业、煤炭开采和洗选业，这些行业与生产生活服务、禀赋资源开发等特定事项有关，对地区工业发展不具有普遍意义。

使用 IS^{*}_{jt} 替换 IS_{jt} 后，各变量的回归结果（见表 5 中模型 7、8）与表 3 相比基本一致。不过，虽然 IS^{*}_{jt} 依然负向影响 $DPER_{jt}$，但二者之间不再具有显著性关系。一个可能的解释是，由于 IS^{*}_{jt} 同时考虑了轻度与中度污染两类行业，增大了低污染行业产值在工业总产值中的比重（全部样本省份 2005—2016 年间 IS^{*}_{jt} 的平均值为 0.495），而新加入的中度污染行业与轻度污染行业相比，其受到集聚中心吸引、并从“外围→核心”迁移的特征相对较弱，因此 IS^{*}_{jt} 对 $DPER_{jt}$ 的负向影响不再显著。不过这也说明，虽然 IS^{*}_{jt} 扩大了低污染行业范围，但它对 $DPER_{jt}$ 依然不具有显著正向影响的作用，IS^{*}_{jt} 的水平仍未超越拐点，行业结构要真正发挥其降污减排效力，仍有待于工业增长进一步摆脱对重污染行业的依赖。

（2）重新衡量环境规制水平。在表 3 的回归结果中，与规制水平差异 ERd_{jt} 的作用不同，平均规制水平 ERa_{jt} 反而阻滞了污染就近转移，这似乎有悖于“污染避难所”假说的直观理解。我们怀疑，ERa_{jt} 的异常结果是否与我们衡量环境规制水平的特定视角有关？为排除这种可能性，我们放弃了之前的单一指标（工业二氧化硫去除率），转而使用综合指标重新衡量环境规制水平。方法如下：第一步，选择城市 i 的污水处理率、工业二氧化硫去除率作为指标体系①，并简单加权为一个综合指标，用以全面衡量城市 i 的环境规制水平。其中，污水处理率的数据来自《城市建设统计年鉴》。第二步，依据新的环境规制衡量标准，重新测度污染转移 $DPER^{*}_{jt}$、规制水平差异 ERd^{*}_{jt} 和平均规制水平 ERa^{*}_{jt}。回归结果见表 5。

首先，当在表 5 的模型 9、10 中替换进 $DPER^{*}_{jt}$、ERd^{*}_{jt} 和 ERa^{*}_{jt} 以后，规制水平差异依然显著正向影响

① 与其他使用综合指数法的文献有所不同：（1）本文没有使用烟尘、粉尘去除率作为衡量指标，因为 2011 年该项目统计口径发生较大变化，在本文时间序列内不具有连续可比性。（2）本文使用污水处理率而非工业废水排放达标率，基于两个原因：一是工业废水排放达标量数据从 2011 年起不再有公开数据；二是因为我国城市污水处理率严重偏低，在 21 世纪初仍有相当数量城市的污水处理率为 0，到 2020 年才力争达到 85%以上（《水污染防治行动计划》（国发［2015］17 号）），该项指标在本文时间序列内不仅变化幅度更高，而且城市间差异比较明显，更有利于区分不同城市间环境规制水平高低。

污染转移，平均规制水平则显著负向影响污染转移，这与表 3 中模型的回归结果一致。同时，与表 3 中的结果相比，其他解释变量的回归结果也比较相似，唯一区别是行业结构 IS_{jt} 虽然回归系数依然为负，但不再显著①。随后，当在表 5 的模型 11、12 中进一步使用 IS_{jt}^{*} 替换 IS_{jt} 后，ERd_{jt}^{*}、ERa_{jt}^{*} 与 $DPER_{jt}^{*}$ 的关系仍没有改变，而行业结构 IS_{jt}^{*} 与污染转移 $DPER_{jt}^{*}$ 负向关系的显著性则有所增强。

总体来看，一个城市更高（或更低）的环境规制水平并不局限于某一类污染物，它也同时表现为多种污染物的减排能力。因此，无论使用单一或综合指标衡量环境规制水平，环境规制因素都同时表现出推动和阻滞污染转移的双向作用，这一结果并没有因为衡量标准的变化而发生改变。

表 5　　**稳健性检验结果**

被解释变量：$DPER_{jt}$			被解释变量：$DPER_{jt}^{*}$			被解释变量：$DPER_{jt}^{*}$		
解释变量	模型 7	模型 8	解释变量	模型 9	模型 10	解释变量	模型 11	模型 12
$L.DPER_{jt}$	0.499*** (5.10)	0.488*** (5.12)	$L.DPER_{jt}^{*}$	0.445*** (5.29)	0.476*** (4.10)	$L.DPER_{jt}^{*}$	0.437*** (6.10)	0.463*** (3.81)
ERa_{jt}	−1.308* (−1.71)		ERa_{jt}^{*}	−0.486* (−1.67)		ERa_{jt}^{*}	−0.565* (−1.91)	
ERd_{jt}		0.992*** (3.49)	ERd_{jt}^{*}		1.411*** (3.03)	ERd_{jt}^{*}		1.363*** (2.98)
$IAad_{jt}$	−1.127*** (−2.69)	−0.848*** (−2.73)	$IAad_{jt}$	−0.767*** (−2.93)	−0.586** (−2.23)	$IAad_{jt}$	−0.731** (−2.44)	−0.573** (−2.02)
IS_{jt}^{*}	−0.809 (−1.60)	−0.399 (−1.04)	IS_{jt}	−0.522 (−1.42)	−0.353 (−0.87)	IS_{jt}^{*}	−0.608** (−2.07)	−0.415 (−1.46)
RD_{jt}	0.439* (1.82)	0.312* (1.71)	RD_{jt}	0.487** (2.47)	0.436** (2.18)	RD_{jt}	0.502** (2.55)	0.427** (2.22)
EX_{jt}	0.456*** (3.94)	0.353*** (3.85)	EX_{jt}	0.436*** (4.38)	0.371*** (2.81)	EX_{jt}	0.442*** (5.07)	0.371*** (3.21)
FDI_{jt}	0.033 (0.23)	−0.016 (−0.13)	FDI_{jt}	−0.037 (−0.36)	−0.072 (−0.72)	FDI_{jt}	−0.062 (−0.59)	−0.082 (−0.80)
ESd_{jt}	0.539*** (4.52)	0.608*** (4.96)	ESd_{jt}	0.497*** (6.03)	0.517*** (5.69)	ESd_{jt}	0.507*** (6.33)	0.528*** (5.28)
IC_{jt}	−0.290 (−0.94)	0.176 (0.89)	IC_{jt}	−0.056 (−0.24)	0.102 (0.48)	IC_{jt}	−0.086 (−0.37)	0.082 (0.41)
常数项	0.340 (0.37)	−0.536 (−0.96)	常数项	−0.121 (−0.17)	−0.847 (−1.00)	常数项	−0.106 (−0.25)	−0.960 (−1.28)
AR(1)P 值	0.026	0.015	AR(1)P 值	0.031	0.019	AR(1)P 值	0.032	0.017
AR(2)P 值	0.914	0.972	AR(2)P 值	0.826	0.901	AR(2)P 值	0.798	0.877
Sargan P 值	0.228	0.386	Sargan P 值	0.219	0.384	Sargan P 值	0.290	0.376
省份 FE	是	是	省份 FE	是	是	省份 FE	是	是
年份 FE	是	是	年份 FE	是	是	年份 FE	是	是
样本量	240	240	样本量	240	240	样本量	240	240

五、结论与建议

本文研究结论如下：中国省域内更多地表现出“外围→核心”迁移的转移粘性，污染源企业正朝着省域内环境规制以及经济水平更高的地区聚集。造成这一现象的原因在于阻滞污染源企业迁移的力量大于推动其迁移的力量。其中，在阻滞力方面，产业集聚因素阻滞污染转移的作用非常强烈且稳定；但在推动

① 虽然行业结构 IS 在此处不显著，但从表 3 的整个回归结果看，它整体上仍表现出阻滞污染转移的明显特征。

力方面，环境规制和产业升级因素对污染转移的影响具有正反双向作用，其推动力受到了抑制。虽然省域内更高的规制水平差异和技术创新具有推动污染转移的作用，但平均规制水平的提升和行业结构优化反而阻滞了污染转移。

从上述结论可以看出，中国所处的特定发展阶段决定了其环境治理的长期性和复杂性。首先，产业集聚作为推动中国发展的重要因素，在可预见的未来都将在中国经济中扮演重要角色，但集聚效应也可能导致相当程度的环境污染长期滞留于中心城市。其次，平均规制水平的提升也可能造成意想不到的结果，如果污染源企业在所有地区均要面对快速强化的环保监管，其最优决策反而是向中心城市靠拢，以便从集聚的正外部性中受益，这将造成中心城市更难摆脱污染的困扰。最后，从目前情况看，行业结构优化对促进污染减排的决定性作用尚未得到发挥，工业部门中轻度污染行业的发展不仅远未达到取代重度污染行业的水平，其发展水平也不具有驱使污染源企业远离中心城市的作用。

上述分析结果对规划中国未来的环境治理具有重要启示意义：（1）污染减排的地区布局应注重层次性。中心城市作为区域的经济集聚中心，不仅具有吸引污染源企业的内在机制，而且这种吸引力也极易受到省域内平均规制水平、行业结构等外在因素的正向影响。因此，污染减排的地区布局应以中心城市为重点，按照重要性高低为不同类型城市制定更具针对性的减排目标，以求找准污染减排的核心关注点、避免“四面出击”。（2）强化环境规制的过程要重视政策协调与配套。更严格的环保监管是污染减排的必由之路，但强化环境规制的过程也可能产生多重政策效果。如果仅仅关注中心城市的环保监管，城市间更大的规制水平差异会推动污染就近转移，不利于区域整体环境质量的改善；如果“一刀切”式的提升平均规制水平，则会强化污染转移粘性，中心城市也更难以摆脱污染的困扰。因此，省域各城市间不仅需要就环境治理的目标达成协同规制的共识，也有必要构建分工明确、定位合理的城市差异化发展战略，以减少污染源企业因为环保压力在城市间的迁移活动。（3）应结合使用污染减排的短期和长期工具。从短期看，中国现实的环境压力急需环境规制的补救，“不痛不痒”的环保监管将累积环境污染、带来不可逆的环境破坏。但从长期看，污染问题的根本解决仍有赖于市场规律而非规制政策，只有借助产业升级，重污染行业才能逐渐被取代，并被迫从集聚中心向外围地区迁移，中心城市乃至整个区域的环境质量才能得到真正改善。因此，未来的污染减排规划应兼顾其短期和长期目标，结合使用环境规制和产业升级两类工具，推动“规制+升级”双轮驱动的深层次去污模式。

当然，本研究也存在一些不足，有待后续研究进一步完善。使用省域内地理重心轨迹来测度污染就近转移虽然提供了明确的地理含义，但也有其局限性。由于污染转移被限定在省域尺度内，忽略了污染在省域间转移的可能性。这可能导致某些发达省份内污染重心向“经济-规制”重心的靠拢隐含了污染源企业从省域边缘迁出到其他就近省份所带来的变化，而这一变化并未体现在省域的“核心—外围”框架内。一个可能解决方法是扩展地理范围，使用城市群的“核心—外围”框架来探讨污染就近转移问题。当然，这一优化方法的科学性仍有待进一步研究加以证实。

参考文献

[1] Copeland B R，Taylor M S. Trade，Growth，and the Environment [J]. Journal of Economic Literature，2004，42 (1)：7-71.

[2] 林伯强，邹楚沅. 发展阶段变迁与中国环境政策选择 [J]. 中国社会科学，2014 (5)：81-95，205-206.

[3] 彭文斌，吴伟平，邝嫦娥. 环境规制对污染产业空间演变的影响研究——基于空间面板杜宾模型 [J]. 世界经济文汇，2014 (6)：99-110.

[4] 陈秀山，徐瑛. 中国制造业空间结构变动及其对区域分工的影响 [J]. 经济研究，2008 (10)：104-116.

[5] 耿文才. 新经济地理学视角下中国纺织业区际转移的粘性分析 [J]. 地理研究，2015 (2)：259-269.

[6] Weterings A，Knoben J. Footloose：An Analysis of the Drivers of Firm Relocations over Different Distances [J]. Papers in Regional Science，2013，92 (4)：91-809.

[7] Knoben J，Oerlemans L A G. Ties That Spatially Bind? A Relational Account of the Causes of Spatial Firm Mobility [J]. Regional Studies，2008，42 (3)：385-400.

[8] Knoben J. The Geographic Distance of Relocation Search：An Extended Resource-based Perspective [J]. Economic Geography，2011，87 (4)：

371-392.
[9] Lee Y. Geographic Redistribution of US Manufacturing and the Role of State Development Policy [J]. Journal of Urban Economics, 2008, 64 (2): 436-450.
[10] Conroy T, Deller S, Tsvetkova A. Regional Business Climate and Interstate Manufacturing Relocation Decisions [J]. Regional Science and Urban Economics, 2016 (60): 155-168.
[11] Ravulaparthy S K, Selby B, Kuppam A. Spatial Firm Demographic Micro-simulator Development and Validation for Phoenix and Tucson, Arizona, Megaregion [J]. Transportation Research Record, 2017 (2671): 59-70.
[12] Van Dijk J, Pellenbarg P H. Firm Relocation Decisions in the Netherlands: An Ordered Logit Approach [J]. Papers in Regional Science, 2000, 79 (2): 191-219.
[13] 沈坤荣, 金刚, 方娴. 环境规制引起了污染就近转移吗? [J]. 经济研究, 2017 (5): 44-59.
[14] Xing Y, Kolstad C D. Do Lax Environmental Regulations Attract Foreign Investment? [J]. Environmental and Resource Economics, 2002, 21 (1): 1-22.
[15] 王奇, 刘巧玲, 李鹏. 我国污染密集型产业的显性转移与隐性转移研究 [J]. 北京大学学报 (自然科学版), 2017 (1): 91-100.
[16] Eslamipoor R, Sepehriar A. Firm Relocation As Potential Solution for Environment Improvement Using SWOT-AHP Hybrid Method [J]. Process Safety and Environmental Protection, 2014, 92 (3): 269-276.
[17] Holl A. Start-Ups and Relocations: Manufacturing Plant Location in Portugal [J]. Papers in Regional Science, 2004a, 83 (4): 649-668.
[18] Holl A. Manufacturing Location and Impacts of Road Transport Infrastructure: Empirical Evidence from Spain [J]. Regional Science and Urban Economics, 2004b, 34 (3): 341-363.
[19] Frenken K, van Oort F, Verburg T. Related Variety, Unrelated Variety and Regional Economic Growth [J]. Regional Studies, 2007, 41 (5): 685-697.
[20] Kirat T, Lung Y. Innovation and Proximity: Territories as Loci of Collective Learning Processes [J]. European Urban and Regional Studies, 1999, 6 (1): 27-38.
[21] Torre A, Rallet A. Proximity and localization [J]. Regional Studies, 2005, 39 (1): 47-59.
[22] Carrincazeaux C, Coris M. Why Do Firms Relocate? Lessons from a Regional Analysis [J]. European Planning Studies, 2015, 23 (9): 1695-1721.
[23] Kronenberg K. Firm Relocations in the Netherlands: Why Do Firms Move, and Where Do They Go? [J]. Papers in Regional Science, 2013, 92 (4): 691-714.
[24] Hong S H. Agglomeration and Relocation: Manufacturing Plant Relocation in Korea [J]. Papers in Regional Science, 2014, 93 (4): 803-818.
[25] Capasso M, Cefis E, Frenken K. Spatial Differentiation in Industrial Dynamics: The Case of the Netherlands (1994—2005) [J]. Tijdschrift Voor Economische En Sociale Geografie, 2016, 107 (3): 316-330.
[26] Yi Y J. Firm Relocation and Age-dependent Reliance on Agglomeration Externalities [J]. Annals of Regional Science, 2018, 61 (2): 439-456.
[27] Okubo T, Tomiura E. Industrial Relocation Policy, Productivity and Heterogeneous Plants: Evidence from Japan [J]. Regional Science and Urban Economics, 2012, 42 (1/2): 230-239.
[28] Forslid R, Okubo T. Which Firms Are Left in the Periphery? Spatial Sorting of Heterogeneous Firms with Scale Economies in Transportation [J]. Journal of Regional Science, 2015, 55 (1): 51-65.
[29] Pumain D, Paulus F, Vacchiani-Marcuzzo C, Lobo J. An Evolutionary Theory for Interpreting Urban Scaling Laws [J]. Cybergeo: European Journal of Geography, 2006, No. 343.
[30] Foreman-Peck J, Nicholls T. Inter-Regional Mobility of Entrepreneurial SMEs [J]. Annals of Regional Science, 2015, 54 (1): 57-87.
[31] Deichmann U, Lall S V, Redding S J, Venables A J. Industrial Location in Developing Countries [J]. World Bank Research Observer, 2008, 23 (2): 219-246.
[32] 徐建华, 岳文泽. 近20年来中国人口重心与经济重心的演变及其对比分析 [J]. 地理科学, 2001 (5): 385-389.
[33] 丁焕峰, 李佩仪. 中国区域污染重心与经济重心的演变对比分析 [J]. 经济地理, 2009 (10): 1629-1633.
[34] 赵小风, 黄贤金, 张兴榆, 陆汝成, 李衡. 区域经济重心及COD、SO2、TSP排放重心演变路径分析——以江苏省为例 [J]. 长江流域资源与环境, 2010 (3): 225-230.
[35] Stern D, Common M, Barbier E. Economic Growth and Environmental Degradation: The Environmental Kuznets Curve and Sustainable Development [J]. World Development, 1996, 24 (7): 1151-1160.
[36] 成艾华, 赵凡. 基于偏离份额分析的中国区域间产业转移与污染转移的定量测度 [J]. 中国人口·资源与环境, 2018 (5): 49-57.
[37] 李小平, 卢现祥. 国际贸易、污染产业转移与中国CO2排放 [J]. 经济研究, 2010 (1): 15-26.
[38] 张宇, 蒋殿春. FDI、政府监管与中国水污染——基于产业结构与技术进步分解指标的实证检验 [J]. 经济学季刊, 2014 (2): 491-514.
[39] 杨子晖, 田磊. "污染天堂"假说与影响因素的中国省际研究 [J]. 世界经济, 2017 (5): 148-172.

(责任编辑: 李　晟　张安平)