

# 财政运行综合绩效评价的理论、指标与展望

## ——基于财政预算“行为-功能”的分析

Theory, Indicators and Prospects of Comprehensive Performance Evaluation of Fiscal Operation: An Analysis Based on the “Behavior-Function”

Framework of Fiscal Budget

马海涛 曹堂哲 彭珮文

MA Hai-tao CAO Tang-zhe PENG Pei-wen

**[摘要]** 2018年以来,全国各地积极开展财政运行综合绩效评价试点,结合本地的实际情况探索出了各有特点的评价指标体系,虽然实践先行,但尚未就评价的理论、逻辑和指标框架达成共识,一定程度上制约了财政运行综合绩效评价工作的推进,影响到了评价结果运用和质量效益。笔者通过系统梳理财政运行综合绩效评价的相关文献,提出了财政运行综合绩效评价的“财政运行行为-政府预算功能”分析框架,在此框架下阐明了财政运行综合绩效评价的基本框架,明确了关键评价指标,分析了当前实践中面临的突出问题,展望了财政运行综合绩效评价的优化路径,即推动财政运行综合绩效评价全覆盖、健全指标和标准体系、开展专项监测与评价、完善评价机制、加强评价结果长效运用、强化激励约束等。上述建议为凝聚共识,全面开展政府财政运行综合绩效评价提供了可行的方案和路径。

**[关键词]** 财政运行综合绩效评价 财政运行行为 预算功能

**[中图分类号]** F812.3 **[文献标识码]** A **[文章编号]** 1000-1549 (2022) 11-0003-14

**Abstract:** Since 2018, local governments at all levels have actively carried out pilot projects for comprehensive performance evaluation of fiscal operations, through “crossing the river by feeling the stones”, and have explored indicator systems with their own characteristics. However, there are still not much in common on the theoretical basis, logical framework and common indicators of evaluation, which have restricted the promotion of comprehensive performance evaluation of fiscal operations and affected the application of the results. We summarized the origin, history and practical explorations related to comprehensive performance evaluation of fiscal operation in China by systematically reviewing the relevant literature. Then, we proposed a “fiscal behavior-budget function” framework for the performance evaluation. Under this framework, we clarified the basic framework of the comprehensive performance evaluation of fiscal operation, specified the key evaluation indicators, and analyzed the prominent problems. In the end, we proposed methods to optimize the comprehensive performance evaluation of fiscal operation: First, each public sector, every level of government, as well as every fiscal expenditure should be fully covered by the comprehensive performance evaluation of fiscal operation. Second, the index and standard system should be improved and developed. Third, special monitoring and evaluation ought to be carried out. Forth, evaluation mechanism should be well-designed and reformed. Fifth, the system should be established to guarantee the appropriate use of the evaluation results. Last but not least, incentives and punishments related to the feedback should be put well in place. The work mentioned above would provide feasible solutions to build a consensus for comprehensive performance evaluation of government fiscal operations.

**Key words:** Comprehensive performance evaluation of financial operation Fiscal operation behavior Budget function

**[收稿日期]** 2022-07-14

**[作者简介]** 马海涛,男,1966年5月生,中央财经大学中国财政发展协同创新中心教授,博士生导师,研究方向为财税理论与政策、财税管理;曹堂哲,男,1979年4月生,中央财经大学政府管理学院教授,博士生导师,研究方向为预算绩效管理、政府治理;彭珮文,女,2000年1月生,北京大学政府管理学院硕士研究生。本文通讯作者为曹堂哲,联系方式为 caotangzhe@163.com。

**[基金项目]** 国家社科基金重大项目“实质性减税降费与经济高质量发展”(项目编号:19ZDA070);国家社科基金一般项目“我国政府部门预算参与和管理绩效作用机理的模型建构与实证检验研究”(项目编号:20BZZ060);教育部人文社科青年项目“基于序参量识别的政府投资基金绩效评价指标体系研究”(项目编号:19YJC630005);世界银行贷款“现代财政制度与国家治理项目”支援项目——全面实施预算绩效管理背景下的预算绩效评价创新研究子项目(二)“‘十四五’预算绩效评价发展规划”研究。

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、导论

财政运行综合绩效评价涵盖政府全部预算收入和支出,涉及财政运行各个方面,反映财政工作目标达成的结果及其有效实现结果的过程,是政府层面预算绩效管理的核心。自2018年实行财政运行综合绩效评价试点以来,无论是财政运行绩效评价的实践探索,还是关于财政运行评价的相关理论研究,并未就财政运行综合绩效是什么、评什么、怎么评、怎么用等问题达成共识,亟需在理论上厘清财政运行综合绩效的本质和逻辑,建构关键绩效指标框架,为深入全面推进财政运行综合绩效评价工作,提高评价质量和结果运用成效提供理论支撑、擘画实施路径。

从实践来看,2018年《中共中央 国务院关于全面实施预算绩效管理的意见》要求“实施政府预算绩效管理”,“对下级政府财政运行情况实施综合绩效评价”。该意见印发以来,北京、山东、浙江、辽宁等地开始探索省以下政府财政运行综合绩效评价指标体系,目前财政运行综合绩效评价实践及其理论总结还在不断探索和创新中,对指标框架的逻辑尚未达成共识,共性框架并未形成。比较有代表性的指标体系主要有:一是以北京市为代表的“五性”维度体系,该体系突出问题导向,旨在监测和评价各区“钱是否够用”“成本是否节约”“管理是否规范”“效果是否显著”“组收是否努力”等问题(吴素芳,2020<sup>[1]</sup>;王文庚和曹堂哲,2021<sup>[2]</sup>),从财力保障的充分性、成本节约的经济性、运行成效的有效性、财政管理的规范性和财源建设的积极性五个维度构建了指标体系。二是以山东省为代表的整合性指标体系,该体系从收入绩效、支出绩效、管理效率、运行成效、可持续性等方面构建指标体系。比如:韩志毅等(2021)<sup>[3]</sup>;谢可川等(2022)<sup>[4]</sup>基于山东省济南市、济宁市实施的市县级政府财政运行综合绩效评价,临沂、德州部分县(市、区)乡镇(街道)政府财政运行综合绩效评价,日照市市、县、镇街三级试点(自评)试点提出了涵盖收入绩效、支出绩效、规范管理、保障水平、运行成效等维度的指标体系。赵勇和徐晶(2022)<sup>[5]</sup>从收入、支出、管理效率、运行成效、可持续性等方面介绍了廊坊市开展试点县财政运行综合绩效评价和管理的情况。项前和王兆玉(2021)<sup>[6]</sup>基于江苏省南京市雨花台区街道试点提出包括财政运行管理、财政收支情况、财政运行绩效、

财政可持续在内的指标体系。孙逊(2021)<sup>[7]</sup>从财政收入与支出、重大政策落实情况、预算编制与执行、财政综合管理等维度建构了指标体系。三是以浙江、江西等地为代表的“财政管理+政府履职”体系,该类指标更加关注财政预算收支质量、执行状况、财政改革要求(张喆,2020<sup>[8]</sup>;王宏斌等,2020<sup>[9]</sup>),从财政管理、政府履职成效和可持续性等方面建构指标体系。刘国永等(2022)<sup>[10]</sup>基于江西省赣州市财政运行评价试点提出了包括财政管理、财政运行绩效、财政可持续性的指标体系。

从已有的理论研究相关文献来看,财政运行综合绩效评价涉及财政运行的各个方面,诸如财政能力评估、债务风险评估、收入绩效评估、财政运行质量评价、PEFA政府财政管理能力评价等,这些研究仅侧重财政运行的某一具体方面,并未形成逻辑自洽的综合性评价框架。

基于上述实践总结和理论梳理,本文在预算绩效管理视角下,分析财政运行综合绩效问题,将预算编制、执行、决算、公开等预算管理活动作为贯穿财政运行的一条“红线”,串接起财政运行行为和预算功能。首先在既有文献评论的基础上,阐述和论证了“财政运行行为-政府预算功能”的二维分析框架及其指标体系;然后在梳理财政运行综合绩效评价来龙去脉的基础上,分析了现有实践中具有代表性的指标体系的共性、特点和存在的局限;最后提出了进一步深化财政运行绩效综合评价的实施路径。本文提出的“财政运行行为-政府预算功能”二维分析框架吸收了现有实践的成果,在以下几个方面有所突破和创新:一是相较于现有的指标体系对财政运行行为的关注各有侧重的问题,“框架”更为全面地反映了“十四五”时期我国财政运行行为的主要方面,提炼了财政运行行为的关键绩效指标;二是相较于现有指标体系判断财政运行绩效好坏的标准不够完整的问题,“框架”分析了财政运行行为与预算功能的对应关系,给出了判断财政运行各项关键绩效指标好坏的标准;三是相较于现有指标主要反映一般公共预算运行的问题,“框架”展望了“统分结合”的指标体系设计思路,即在分析“四本”预算总体行为和功能的基础上,针对“四本”预算各自的行为和功能特点,设置个性化的关键绩效指标。

## 二、文献综述、理论分析与指标框架

与财政运行综合绩效评价工作的起步探索相一

致, 现有的研究仅明确了财政运行综合绩效评价主体和对象, 即《中共中央 国务院关于全面实施预算绩效管理的意见》要求的“各级财政部门要对下级政府财政运行情况实施综合绩效评价”。但是对于评价什么, 如何评价等问题, 尚未给出一致公认的界定。对此, 本文从绩效和绩效评价的本质要素出发, 演绎推论出财政运行综合绩效评价的“财政运行行为-政府预算功能”二维分析框架, 并在此框架下提炼和归集财政运行综合绩效评价的关键绩效指标。

### (一) 文献综述

与财政运行综合绩效评价相关的研究主要包括财政能力评估、地方政府债务和风险评估、财政收入绩效评价和财政运行质量评价等。这些研究仅着重于财政运行的某个具体环节或者某个层面, 尚未形成综合性的评价框架和指标体系。

1. 财政收入绩效评价。财政收入是财政运行的构成部分, 对地方财政收入绩效评价的研究尚处于起步阶段。比如, 李明等(2022)<sup>[11]</sup>从税负调控成效、收入质量水平、债务风险水平三个维度构建了预算收入绩效评价指标。于树一(2021)<sup>[12]</sup>从地方财政收入的决策、执行和绩效三个维度建构了地方财政收入绩效评价体系。北京市财政局从稳固存量、培育增量、提高质量和贡献等维度构建了财源建设评估指标体系(北京市财政局办公室, 2022<sup>[13]</sup>)。

2. 财政能力评估。财政能力评估与财政运行综合绩效评价在内容上有许多重合之处。财政能力反映一地政府通过其公共权力汲取财政资源, 并有效分配公共资源、提供公共产品和服务的能力(杨娟等, 2021<sup>[14]</sup>)。财政能力评估最早是指可以通过征税以获取财政资源的能力(邱婧, 2017<sup>[15]</sup>)。其衡量标准主要为财政收入的总量、结构和支出效率。财政发挥作用的能力主要包括发挥宏观调控职能、维持市场和社会正常秩序、维持政治秩序、防化社会公共风险等(陈龙, 2020<sup>[16]</sup>)。目前对政府财政能力评估的研究主要采用经济学上的定量分析方法。比如: 辛冲冲和杨文珂(2020)<sup>[17]</sup>使用熵权法, 建构了对31个省级行政区2007—2013年的财政汲取能力指数, 并分析了地区差异。杨志安和邹国庆(2016)<sup>[18]</sup>使用因子分析法, 提取了核心、基础、发展和保障四个主要能力因子。总之, 财政能力评估通过构建多元指标体系, 合成指数进行对比和评价, 这一方法对财政运行综合绩效评价具有启发意义。

3. 地方政府债务和风险评估。地方政府债务和风险评估指标体系是构成财政运行综合绩效评价指标的重要方面。在国家层面和地方政府层面, 债务和风险评估制度和指标体系不断完善。2016年国务院办公厅印发《地方政府性债务风险应急处置预案》, 财政部于同年印发了《地方政府性债务风险分类处置指南》。同年青海省政府出台了《地方政府债务风险评估和预警暂行办法》, 提出了新增债务率、逾期债务率、债务率和偿债率等评价指标(中央政府门户网站, 2016<sup>[19]</sup>)。各地开展债务风险评估的指标与此基本一致。2021年, 国务院印发的《关于进一步深化预算管理制度改革的意见》<sup>[20]</sup>中进一步提出完善政府债务风险评估的指标体系以及对政府偿债能力进行综合评估。

4. 财政运行质量评价。财政运行质量即财政运行的优劣程度, 包括财政收入、财政支出、公债运行等方面, 反映了财政部门组收和管理的质量与效益, 包括现有财政收支规模多大程度上能够保障政府职能充分发挥, 现有收支结构是否合理, 财政收支如何组织管理, 财政预算编制是否按时按质执行, 以及当下财税制度与经济社会运行是否协调(侯燕明, 2020<sup>[21]</sup>)。影响财政运行质量的因素主要包括财源质量和财政管理(赵玮, 2019<sup>[22]</sup>)。高质量的财政运行需要符合效率性、公平性、稳定性和规范性四大特征(朱云飞和赵宁, 2019<sup>[23]</sup>)。上述研究为财政运行综合绩效评价指标和标准的设计提供了参考。

5. 国际 PEFA 政府财政管理能力评价体系。公共支出和财政问责(Public Expenditure and Financial Accountability, PEFA)是由世界银行、国际货币基金组织、欧盟委员会等组织赞助支持的一项多机构合作计划, PEFA 评价体系已被英国、法国、新西兰、巴西等国广泛应用。作为 PEFA 体系的一种工具, 财政管理和绩效测量框架(Public Financial Management Performance Measure Framework, PFM Performance Measure Framework)提出了衡量公共财政管理绩效的指标体系, 包括预算信用, 全面性和透明度, 预算制定与政策目标的符合程度, 预算执行的可行性和控制, 会计、记录和报告, 外部监督和审计六个维度。上述六个维度的28个指标可以归结为系统的结果、关键的横向特征、预算循环三类(PEFA Secretariat, 2011<sup>[24]</sup>)。解洪涛和陈志勇(2016)<sup>[25]</sup>等探讨了 PEFA 政府财政管理能力评价体系与我国实施的《地

方财政管理绩效综合评价方案》之间的关系。2022年湖南省财政厅首次引进 PEFA 系统,用以提升地方政府财政管理(周极宏,2022<sup>[26]</sup>)。PEFA 提出的预算循环思想,是本文将预算过程作为贯穿财政运行的红线,提出财政运行综合绩效评价“二维模型”的重要理论来源。

## (二) 分析框架

考虑到“绩效是一个多维建构,观察和测量的因素不同,其结果也会不同。”(Bates 和 Holton III, 1995<sup>[27]</sup>) 本文在综合行为绩效观和结果绩效观的基础上,将绩效定义为“绩效是目标达成的结果及有效实现结果的过程。”该定义有两个要点,即行为和实现的结果。抓住这两个关键要素,本文从预算管理的视角分析财政运行,对应提出“财政运行行为-政府预算功能”的二维分析框架,财政运行行为对应一级政府财政收入、支出、管理、调节和风险防范

控等方面工作的状态和发展变化情况。政府预算功能则是预算管理在总额控制、配置效率、运营效率等方面发挥的作用和实现的目标。因此,政府财政运行综合绩效可以定义为:一级政府在收入、支出、管理、调节和风险控制等方面,实现总额控制、配置效率、运行效率等方面目标的程度。

绩效评价包含行为测量、目标导向和判断标准三个要素,在上述二维框架中,预算功能维度阐明了财政运行行为达到的目标导向和判断标准。据此可将财政运行综合绩效评价定义为上级财政部门对下级政府财政“收、支、管、调、防”等工作,达成总额控制、配置效率、运营效率等方面目标情况的测量、分析和评判,主要包括收入的合理性、保障的充分性、成本的节约性、管理的规范性、产出的有效性、发展的可持续性等方面。上述演绎推论的逻辑如图 1 所示:

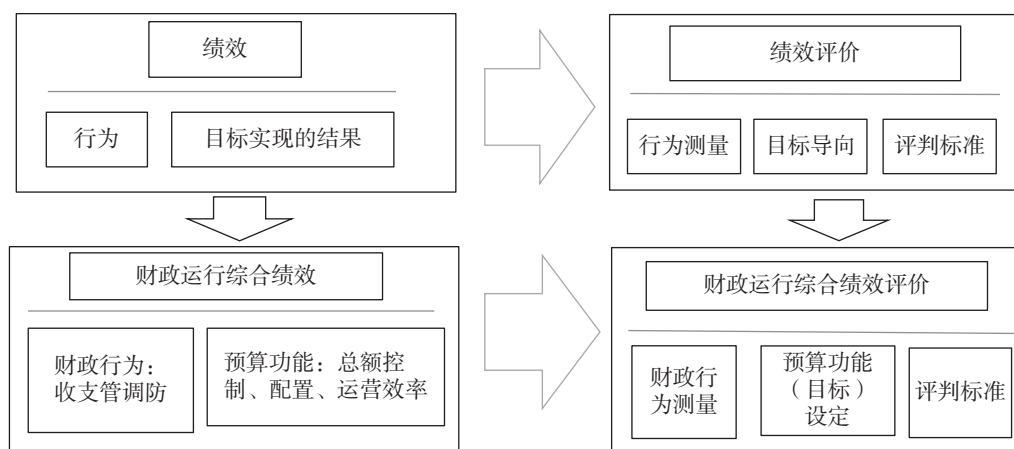


图 1 本文的演绎推论逻辑

1. 二维框架下的财政运行行为。在行为主义经济学中,信息的结构和参与者的特征会对决策和结果产生系统性的影响(Baker 和 Nofsinger, 2010<sup>[28]</sup>)。决策者经常面临一系列巨大不确定性,以及结果可量化的有限性的选择,而选择的建构方式会对选择产生强烈的影响,即客观事实不变,提出问题的方式不同,选择也会不同(Ackert 和 Deaves, 2009<sup>[29]</sup>)。

财政运行行为是财政运行的具体行动选择,财政运行行为随着制度的变化而产生改变,不同时期财政运行问题不同,行为选择就不同。传统上财政运行围绕“收、支、管、平”四个方面开展工作(贾康,2020)<sup>[30]</sup>。随着新《预算法》和《预算法实施条例》的颁布,“讲求绩效”已经成为预算管理的一项基本原则,财政运行拓展为“收、支、管、平、效”五

个方面。“十四五”时期,财政工作坚持以政领财、以财辅政,围绕“收、支、管、调、防”五个方面下功夫(新华社,2021<sup>[31]</sup>)。从财政行为的维度出发,分析其在社会经济环境、制度和有限理性约束下的行为动机和行为结果(目标实现程度),就构成财政运行综合绩效评价的基础内容。

为了做好财政运行综合绩效评价,需要提出若干关键绩效指标对财政运行行为进行测量。一是收入方面。反映政府财政收入政策、制度和管理等方面的合理性与执行情况,以及财力对政府基本职能履行和政府运行的保障情况,主要从财力保障、财源建设、预算平衡等方面评价,涉及的关键绩效指标包括可用财力、转移支付依赖度、减税降费目标完成率、收入质量、财源建设情况、财政自给率等指标。二是支出方

面。反映政府财政支出结构的合理性，主要关注支出的配置效率。配置效率反映各级政府通过支出结构优化，在保障国家重大战略任务、促进经济社会发展、实现财政资源与政策目标匹配、民生保障等方面的情况，涉及国家重大战略任务财政资金到位率、共同事权地方财政资金到位率、基本民生支出标准达标率、“三保”保障率等关键绩效指标。三是管理方面。反映政府财政管理的成本节约和管理规范程度，主要从成本控制和财政管理等方面进行评价，包括财政供养比、政府行政管理支出占比、一般公共预算收入年初预算到位率等关键绩效指标。四是调节方面。反映充分发挥财政宏观调控和收入分配职能，推进经济社会健康发展的结果和成效，包括重大战略任务完成情况、基本公共服务均等化实现情况、基本公共服务投入产出比等关键绩效指标。五是风险防控方面。反映政府财政风险防控、统筹发展和安全，增强财政可持续性方面的情况，包括政府债务率、负债率、偿债率、社保支付能力、政府债务规模、地方政府债务限额、偿债能力评估、民生政策财政可承受能力评估等关键绩效指标。

2. 二维框架下的预算功能。对财政运行行为进行评价，需要引入预算功能理论，以确立评价的目标导向和评价标准。财政预算作为一种公共治理的工具，能够将战略、政策和资源进行整合，并对公共部门的运转产生影响（江月，2020<sup>[32]</sup>）。“过去200年的现代政府的预算历史表明，公共预算在三个层次上影响着公共部门的运作：在宏观层次上公共预算的规则或者程序对于决策者来说将形成一种总财力约束；在中观层次上公共预算的规则或程序将影响以战略重点为基础的资源配置和利用；在微观层次上公共预算的规则或者程序将影响生产和供给公共服务的效率”（Schick, 1998<sup>[33]</sup>；希克，2000<sup>[34]</sup>）。与此相对应，财政预算有三个层次的基本职能和基本目标：财政总额控制、资源配置效率、运作效率（马骏，2011<sup>[35]</sup>）。财政总额控制反映在有限的财政资源限制下，财政总量的变化和财政收支关系，是保障财政可持续性，实现既定支出目标的重要手段（孙克竞和王祯阳，2020<sup>[36]</sup>）。“总额控制问题涉及这样一些问题：在某个预算时期（比如一个预算年度）根据什么确定了可供分配的财力？换言之，在某一个预算时期，根据什么做出了收入决策，即从社会中汲取多少的资源来进行

预算分配。作为一个基本目标的总额控制是指，对于政治家决定在某一时期准备开展的活动，可以通过预先确定财政收支总额来进行约束，进而将公共支出、赤字、债务的比率控制在一个可以承受的水平，最终使得财政是可持续的”（Schick, 1998<sup>[33]</sup>；希克，2000<sup>[34]</sup>；马骏，2011<sup>[35](23)</sup>）。

由于在市场经济条件下，存在市场失灵问题，面临公共物品供给、外部性、自然垄断、不完全竞争、信息不对称等问题导致的资源配置低下，需要政府运用公共预算对资源进行合理配置、调节分配、稳定经济社会基础（涂漫漫，2015<sup>[37]</sup>）。“配置效率要求政府建立目标并区分其优先次序，而且它要求政府评估公共支出对于这些目标的实际或者预期的贡献”（Schick, 1998<sup>[33]</sup>；希克，2000<sup>[34]</sup>；马骏，2011<sup>[35](4)</sup>）。

运作效率又称作运营效率，是指政府能够有效率地组织和管理财政交易（马骏，2011<sup>[35](25)</sup>），是政府机构运作的效率，即政府机构所花费的资源与这些资源形成的产出之间的比率（希克，2000<sup>[34]</sup>；马骏，2011<sup>[35](26)</sup>）。重点关注财政资金使用效益，包括财政资金使用经济性（节约性）、效率性（投入产出比）和有效性（即通过合理成本达成预期目标）（王雍君，2016<sup>[38]</sup>）。

我国财政预算由“四本”预算组成，总体上除了需要符合财政预算基本职能和基本目标的共性要求以外，每本预算的职能和目标还有更为丰富的内容，比如：政府性基金预算收支行为的专门特定性，国有资本经营预算反映国家以所有者身份对国有资本的调节作用，社会保险基金预算与特定社会保险制度的高度关联性等。推动财政运行综合绩效评价资金覆盖范围从一般公共预算为主到覆盖“四本”预算，需要在预算功能的共性要求下，进一步完善和丰富不同预算的个性化功能，为评价设定“目标导向”提供基础。

3. 二维框架下的绩效评价标准。在预算管理全过程中，从预算功能的视角评判财政运行“收、支、管、调、防”各项工作中，可形成以下评价标准：一是总额控制。重点关注财政保障方面，即关注财政运行能否做好财政收入、支出及其平衡，保障党委政府履职和财政职能的发挥，其评判标准为财力保障充足性、财源建设积极性、发展可持续性等。二是配置效率。重点关注财政运行能否通过支出结构优化，有效保障国家重大战略任务、促进经济社会发展、实现

财政资源与政策目标匹配，包括战略保障和民生保障等方面。三是运营效率。包括过程性的管理效率和结果性的运行成效两个方面。管理效率侧重成本控制和规范性；运行成效侧重财政支出政策目标完成、基本公共服务均等化、投入产出比等方面。四是财政运行的安全性。这一标准贯穿在预算管理各个环节，反映财政统筹经济发展与财政可持续性、防控风险、保障财政安全等方面。

将财政运行行为和预算功能绘制成二维表格，如表1所示，其中预算功能反映预算编制、执行和决算等环节实现总额控制、配置效率和运营效率的情况；财政运行行为反映财政“收、支、管、调、防”等各项工作。财政运行行为和预算功能之间存在对应匹配关系，比如：财政运行行为的管理方面主要对应预算功能的过程性运营效率，绩效标准是成本控制的合理性和财政管理的规范性。

表1 财政运行综合绩效的“行为-功能”分析框架

维度	评价内容			
预算功能目标	预算功能	总额控制、配置效率、运营效率		
	预算过程	预算编制		预算执行
	绩效目标	总额控制（财力保障）	配置效率	运营效率（过程：管理效率）
	绩效标准	1. 财力保障充足性 2. 财源建设积极性 3. 发展可持续性	1. 战略保障 2. 民生保障	1. 成本控制 2. 管理规范
财政运行行为	收入	充分运用收入政策，不断完善现代税收制度	—	—
	支出	—	保持合理支出强度，有保有压、突出重点	—
	管理	—	—	持续深化预算管理制度改革，提升财政资源配置效率和资金使用效益
	调节	—	—	—
	防控	统筹经济发展与财政可持续性，保障财政安全		

资料来源：作者整理。

（三）关键绩效指标框架设计

结合文献和实践探索，按照财政运行综合绩效的“行为-功能”分析框架，将财政运行行为和预算管理功能匹配对应，本文提出了主要针对一般公共预算的财政运行综合绩效评价的关键绩效指标框架，将这一框架拓展到“四本”预算还需要进一步专文研究，对指标进行丰富和深化。一级指标按照财政运行行为设置，包括收入、支出、管理、调节、风险防控五个一级指标。二级指标对行为进行

细化并反映相应的预算功能，设置财力保障、财源建设、预算平衡、战略保障等19个二级指标。三级指标对二级指标进行细化、量化，三级指标名称反映财政运行的关键要素和特征，指标评分标准和指标值反映财政运行在实现总额控制、配置效率、运营效率等预算功能方面的良好状态和实际状态，设置可用财力、转移支付依赖度、减税降费目标完成率、收入质量、财源建设情况等49个三级指标。关键绩效指标框架如表2所示。

表2 财政运行综合绩效评价关键绩效指标框架

一级指标	二级指标	三级指标
收入方面	财力保障 (财力保障充足性)	可用财力
		转移支付依赖度
	财源建设 (财源建设积极性)	减税降费目标完成率
		收入质量
		财源建设情况
预算平衡 (发展可持续性)	财政自给率	

续前表

一级指标	二级指标	三级指标	
支出方面 (配置效率)	战略保障 (配置效率)	国家重大战略任务财政资金到位率	
		共同事权地方财政资金到位率	
	民生保障 (配置效率)	基本民生支出标准达标率	
		“三保”压力	
		“三保”保障率	
		民生支出占比	
		辖区财政均等化情况	
管理方面 (管理效率)	成本控制 (管理效率)	财政供养比	
		政府行政管理支出占比	
		行政成本占财政支出的比例	
		财政供养人员规模	
		“三公”经费控制率	
		公用经费控制率	
	资金统筹与政府间关系 (管理效率)	“四本预算”统筹完成率	
		资产管理的有效性	
		一般转移支付占转移支付的比例	
		年度间资金统筹完成率	
		财政体制管理的有效性	
		直达资金管理规范性	
	预算执行 (管理效率)	一般公共预算收入年初预算到位率	
		一般公共预算支出年初预算到位率	
		一般公共预算支出进度	
		政府性基金预算支出进度	
		结转结余情况及年度间变化情况	
		国库库款 (管理效率)	库款保障水平
		国库集中支付结余消化进度	
		新增专项债券资金使用进度	
		政府采购执行率	
		预算绩效 (管理效率)	预算绩效管理考核得分
		预算公开 (管理效率)	财政透明度评分
	信息基础 (管理效率)	预算管理一体化系统建设进度	
调节方面 (运行成效)	重大战略任务完成 (运行成效)	重大战略任务完成情况	
	基本公共服务均等化 (运行成效)	基本公共服务均等化实现情况	
		基本公共服务投入产出比	
风险防控 (安全性)	债务风险 (安全性)	政府债务率	
		政府负债率	
		政府偿债率	
		财政暂付款项	
		社保支付能力	

续前表

一级指标	二级指标	三级指标
风险防控（安全性）	风险管控（安全性）	政府债务规模
		政府债务限额
		偿债能力（评估结果）
		民生政策方面财政可承受能力（评估结果）

资料来源：作者整理。

### 三、财政运行综合绩效评价的实践及其评析

我国开展财政运行综合绩效评价可以追溯到2013年开始的县级财政管理绩效综合评价和2014年开始的地方财政管理工作考核。2018年以来，按照《中共中央 国务院关于全面实施预算绩效管理的意见》要求，财政管理绩效综合评价拓展为财政运行综合绩效评价，各地积极探索，形成了北京、山东、

浙江等地为代表的不同指标框架体系。2021年财政部各地监管局实施的地方财政运行分析评估工作与财政运行综合绩效评价也存在紧密的联系。上述实践探索，它们评价的主体、对象、内容不同，反映了财政运行的一些局部、侧面和角度，有必要整合纳入“行为-功能”框架，以便找到各自位置和共通语言。具体如表3所示。

表3 我国财政运行综合绩效评价的实践和探索

序号	名称	主体	对象	内容	评析
1	县级财政管理绩效综合评价	财政部 预算司	县级财政	财政管理：预算编制、预算执行、收入质量、支出结构、财政可持续性以及预决算公开等	财政管理是财政运行的重要组成部分，评价指标是财政运行综合绩效评价指标的重要来源
2	地方财政管理工作考核	财政部 预算司	省级财政	财政管理：财政预算执行管理工作、盘活财政存量资金管理工作、国库库款管理工作、推进财政资金统筹使用、预算绩效管理、预算公开和其他财政管理工作情况等	
3	地方财政运行分析评估	财政部 监管局	驻地财政	财政经济形势、财政政策、财税体制、财政收支情况、财政运行风险、财政基础工作等	关注的的关键问题、分析评估指标是财政运行综合绩效评价指标设计的参考
4	以北京为代表的实践	北京市 财政局	十六区	财政保障、财源建设、成本控制、管理规范、运行成效	突出特色是对成本控制、财源建设的关注。提炼的关键绩效指标可以纳入“行为-功能”分析框架
5	以山东为代表的实践	省、市、 县财政	市、县、 乡镇	收入绩效、保障能力、支出绩效、管理规范性和财政可持续性等方面	提炼的关键绩效指标可以拆分整合纳入“行为-功能”分析框架
6	以浙江等地为代表的实践	省、市、 县财政	市、县、 乡镇	财政管理、政府履职、可持续性	财政管理和政府履职之间相对独立，可在“行为-功能”分析框架下实现财政管理与政府履职的匹配对应

资料来源：作者自制。

#### （一）财政管理综合绩效评价

财政管理综合绩效评价主要包括县级财政管理绩效综合评价和地方财政管理工作考核。它们随着《预算法》和预算管理制度改革的深化不断调整和完善。财政管理综合绩效评价与财政运行综合绩效评价相比，缺少运行成效、财源建设等方面的评价。财政管理综合绩效评价经过多年实践，形成的指标及其工作体系相对完善，可以将结果整体纳入财政运行综合绩效评价体系。

1. 县级财政管理绩效综合评价。2012年9月财政部印发的《预算绩效管理工作规划（2012—2015年）》（财预〔2012〕396号）明确要求“中央和省级财政部门都开展县级财政支出管理绩效综合评价试点”（财政部，2012<sup>[39]</sup>）。2013年，财政部针对2012年县级财政支出管理进行了绩效综合评价（中央政府门户网站，2013<sup>[40]</sup>）。2013年至今，县级财政支出管理评价内容随着预算管理制度改革的深化不断完善。比如：2015年财政部印发的《2015年县级财政



管理绩效综合评价方案》(财预〔2015〕139号)评价内容包括规范预算编制、优化收支结构、盘活存量资金、加强债务管理、控制供养人员五个方面(财政部,2015<sup>[41]</sup>)。2018年财政部印发的《2018年县级财政管理绩效综合评价方案及结果的通知》(财预〔2018〕100号)包括预算编制、预算执行、收入质量、支出结构、财政可持续性以及预决算公开六个方面(财政部,2018<sup>[42]</sup>)。

2. 地方财政管理工作考核。地方财政管理工作考核是财政部对各省级政府财政管理工作完成情况及其成效的考核。2014年财政部印发了《地方财政管理绩效综合评价方案》(财预〔2014〕45号),开启了财政部对全国36个省(直辖市、自治区、计划单列市)上一年度地方财政管理绩效的综合评价,评价结果作为转移支付的重要参考(中国政府网,2014<sup>[43]</sup>)。地方财政管理工作考核指标体系随着预算管理制度改革的深化不断完善,比如:2018年财政部印发的《财政管理工作绩效考核与激励办法》(财预〔2018〕4号)对财政预算执行管理工作、收入质量管理工作、盘活财政存量资金管理工作、国库库款管理工作、预算公开管理工作、推进财政资金统筹使用管理工作和其他财政管理工作进行考核(财政部,2018<sup>[44]</sup>)。2018年12月29日财政部印发了关于修订《财政管理工作绩效考核与激励办法》的通知(财预〔2018〕222号),对国库库款管理工作的考核指标内容、评审程序先进县的数目等方面进行了调整(财政部,2018<sup>[45]</sup>)。2020年财政部印发了《地方财政管理工作考核与激励办法》(财预〔2020〕3号),考核内容包括财政预算执行管理工作(一般公共预算和政府性基金预算的支出进度)、盘活财政存量资金管理工作(各省财政存量资金规模)、国库库款管理工作、推进财政资金统筹使用(地方转移支付的结构)、预算公开和其他财政管理工作情况(财政部,2020<sup>[46]</sup>)。为了适应深化预算管理制度改革的要求,2022年财政部印发了最新的《地方财政管理绩效考核办法》(财预〔2022〕15号)。

## (二) 地方财政运行分析评估

2021年以来,财政部各地监管局按照《财政部各地监管局实施地方财政运行分析评估工作暂行办法》(财预〔2021〕42号),结合驻地实际,重点分析属地财政经济形势、财政政策、财税体制、财政收支情况、财政运行风险、财政基础工作等的内容。各

地监管局形成季度报告,向财政部报送地方财政运行中存在的问题和风险,为财税政策制定、预算安排和调整提供参考依据。

地方财政运行分析评估与财政运行综合绩效评价在实施主体、评价方式方法、评价结果等方面存在差异。地方财政运行分析评估是财政部驻各地监管局开展的工作,侧重财政运行形势的分析和监管;财政运行综合绩效评价是上级财政部门加强下级财政管理的重要手段。两者虽然都是对财政运行进行评估评价,各自的侧重、角度和目的不同;但对财政运行核心指标的关注存在一定的共性,比如都会对债务率等财政风险指标进行分析和评价,从客观刻画财政运行状况的角度,两者完全可能形成一张“体检表”。

## (三) 财政运行综合绩效评价的实践探索

2018年以来,各地开始探索构建财政运行综合绩效评价指标体系,大体形成了北京、山东、浙江等地为代表的三类指标体系。

1. 以北京市为代表的“五性”维度。2019年年底,北京市启动十六区财政运行综合绩效评价,评价更加突出问题导向,重点对各区保障能力(钱是否够用)、财源建设(组收是否努力)、成本节约(支出是否节约)、规范管理(管理是否规范)、运行成效(效果是否显著)等方面进行分析和评价,形成了由财政保障、财源建设、成本控制、管理规范、运行成效5个一级指标,24个二级指标,37个三级指标构成的指标体系。2021年北京市持续开展财政运行专项和综合绩效监测评价,上半年开展了成本节约专项监测与评价,主要包括行政运行、城市运行、重大项目成本管控和民生支出标准等方面(王文庚和曹堂哲,2021<sup>[2]</sup>)。2020年以来,北京市东城、顺义、大兴、昌平等区开展了区对街乡镇财政运行综合绩效评价试点。2021年西城区开展了区对街道成本节约专项监测评价。

2. 以山东为代表的整合体系。2020年山东省财政厅选取济南、济宁开展县级政府财政运行综合绩效评价试点,济宁市积极探索市县乡三级同步实施政府财政运行综合绩效评价,在全国尚属首例(财政部,2021<sup>[47]</sup>)。同年财政厅选取了临沂市罗庄作为乡镇政府财政运行综合绩效评价的试点。通过各级试点,山东省形成了收入绩效、保障能力、支出绩效、管理规范性和财政可持续性五个方面的指标体系(韩志毅等,2021<sup>[3]</sup>;谢可川等,2022<sup>[4]</sup>)。河北廊坊市聚焦

财政运行“收、支、管、调、防”，从收入、支出、管理效率、运行成效、可持续性开展试点县财政运行综合绩效评价和管理（赵勇和徐晶，2022<sup>[51]</sup>）。

3. 浙江等地为代表的财政管理+政府履职体系。浙江义乌、江苏常州、辽宁省、江西省等探索形成了“财政管理+政府履职”的指标体系。比如2019年义乌市财政局联合市人大财经工委、中介机构对后宅街道首次开展2018年财政运行综合绩效评价工作，评价指标包括预算管理、预算执行管理、预算监督、政府债务管理、效益管理等方面，并将政府履职状况进一步细分成64项可量化评价指标四级指标（俞晓丽和王婷，2019<sup>[48]</sup>；弘文，2019<sup>[49]</sup>）。再如2020年江苏常州国家高新区（新北区）开展全市首个镇财政运行综合绩效评价试点工作，评价围绕财政管理、运行绩效、综合效果等方面共设置了12项二级指标和22项三级指标。2020年辽宁启动对下级政府财政运行综合绩效评价试点，评价指标围绕预算编制、收支管理、债务风险和财政可持续等方面，包括9项一级指标、21项二级指标和50多项三级指标（苗宇，2020<sup>[50]</sup>）。

在“行为-功能”框架下审视，上述三种指标体系具有共性，也各有其特点和局限性：

一是从行为维度审视。三类指标都紧扣财政运行的关键环节设置指标，总体上都扣住了财政运行“收、支、管、调、防”的相关方面。北京市的指标注重问题导向，成本节约、财源建设和基本公共服务投入产出效率指标是其特色，反映了财政工作的重点和创新。山东省为代表的指标体系充分借鉴了各地实践做法，结合山东省实际，反映了财政运行收入、支出、管理、成效、风险防控等方面的综合情况。浙江等地将财政管理和政府履职进行组合，政府履职与财政收入、支出之间的匹配对应关系考虑较少。但是三类指标基本上都是以一般公共预算作为主要的评价对象，兼顾其他几本预算，在反映“四本”预算整体运行行为和各本预算行为特点方面还有待进一步深化。比如三类指标都很少涉及预算统筹、国有资本经营预算的调节、政府性基金预算的收支行为等方面。

二是从框架的功能维度审视。三类评价指标都或多或少体现了财政运行应该达到的目标。北京市突出强调收入保障的充分性、成本的节约性。山东省和浙江等地更加突出管理规范性和运行结果的有效性。总体上，三类指标尚未全面地反映各本预算的所有功能

目标，比如政府性基金预算的征收标准、使用效果、对专项债务的支撑能力等；国有资本经营预算的资金战略方向、收益上缴、支出结构、使用效果等；社会保险基金预算中各类社会保险基金收支政策效果、基金管理、精算平衡、地区结构、运行风险等。

三是从评价标准和评分维度审视。三类评价细化了各项指标的评分标准。对于那些缺少标准值的指标如何评价，北京市采用指数评价法进行标准化评分，将财政建设和财政管理考核结果直接纳入评价结果；山东省引入功效系数法进行评分，采用了高质量发展考核、政府绩效考核、财政管理考核等方面的考核指标和结果；浙江等地则更多采用达标考核的方法进行评分，选取了规划指标、政府绩效考核等指标进行评分。上述三种方法都是在有限的条件下，确定评分标准的相对优化的选择，只有在有全国数据和多年数据的基础上，才能找到更加科学的评价基线值。

#### 四、财政运行综合绩效评价存在的问题诊断

与项目支出和部门整体支出绩效评价相比，财政运行综合绩效评价范围包含一级政府预算的全部收支，涉及财政运行的各个环节，系统性、综合性、整体性更强。在“财政行为-政府预算功能”二维框架下审视现有的财政运行实践，存在以下几个突出问题：

##### （一）指标体系不够科学

财政运行综合绩效评价实践在各地都已经取得较好进展，正如本文上段分析，不同地方开展财政运行综合绩效评价的基本维度（一级指标）尚未完全达成共识，不同政府层级的指标的差异性也不够突出，指标权重和评分方法也各不相同。

目前，或多或少存在不同层次政府财政运行综合绩效评价指标框架“俄罗斯套娃”现象。比如：某些地方省对市、市对县、县对街乡镇的评价指标体系基本相同，未能体现不同层级政府财政运行的特点。事实上，我国的地区经济发展水平存在较大差异，不同地区的财政运行水平在财政的收支结构、债务和社会保障可持续性以及基本公共服务均等化方面也存在较大差异（马光荣和吕冰洋，2022<sup>[51]</sup>）。因此亟需建立符合省、市、县、乡镇特点的、多层次的财政运行综合绩效评价指标体系。

目前已有的一些指标体系对指标权重和评分标准确立尚未达成共识，比如北京运用指数评价法、数据

包络分析、分析与打分指标相结合的方法进行评价；山东按照静态和动态评价结合、正向和负向激励结合的原则进行标准和权重确立（济南市财政局，2021<sup>[52]</sup>）。各地的评价方法和权重设计方法都有较大的差异，亟需确立一致的标准。

### （二）绩效标准不够健全

绩效标准为衡量财政运行的好坏提供了一把标尺，包括规范和实证两层含义。

一是实证层面的标准，即绩效标杆或目标，通过实际值与标杆值或目标值进行对照，评判绩效好坏。目前财政运行各项指标的标准体系还不完善，有些指标有较为明确的标杆值，比如债务率指标；有些指标还没有公认的标准值，比如收入质量；有些指标的标准值各地也不完全统一，比如运行成效指标中的普惠性幼儿园覆盖率等；有些历史标准本身的合理性也难以判断，比如人均行政成本等。绩效标准体系的不完善，制约了财政运行综合绩效评价的推广。

二是规范层面的标准，即判断财政运行好坏的价值准则和价值尺度。按照预算功能理论，财政运行综合绩效应该以总额控制、配置效率、运营效率为基础，结合财政运行各项工作，形成系列判断标准，比如财力保障的充分性、财源建设的积极性、财政发展的可持续性、成本的节约性等。这些标准相较于项目支出绩效评价的4E标准（经济、效率、效益和公平）更为丰富，尤其是配置效率、保障充分性、运行安全性等标准是项目支出绩效评价不具有的。在实际评价中，一些评价简单套用4E标准，无法全面地评价财政运行综合绩效。

上述关于指标标准值的问题普遍存在各地财政运行综合绩效评价实践中，很多地方刚刚试点开展财政运行综合绩效评价，一般选择为数不多的地方开展试点，各项指标很难找到行业标准、政策标准等标杆值进行评分，采用历史标准进行评分有时候显得科学性不足。

### （三）运行成效边界不够清晰

财政运行综合绩效评价不仅要关注财政管理及风险，还应重点关注运行成效，这一点已经成为共识。但是对于财政运行成效到底是什么并没有达成一致。实践中存在以下几种做法：一是将“经济社会高质量发展指标”作为衡量财政运行成效的指标，这种

做法一定程度上“夸大”了财政运行成效的范围，超越了财政收支调节经济社会发展的作用域。比如S省和J省对市级财政运行综合绩效评价指标中就借用了诸如“投资结构优化、制造业增加值占比提高幅度及规模以上制造业增加值增速”等指标，<sup>①</sup>这些指标反映的是整个社会、多方投入的结果，与财政运行成效之间的相关关系很难简单地做出判定。

二是将财政运行成效和一级政府的履职绩效或政府绩效考核相等同，政府绩效考核一般涉及履职效能、管理效率、社会效应和创新创优等方面，包括政府部门的各项工作。比如J省对X市和W市对X县的财政运行成效指标，将政府绩效考核的经济、社会、文化、生态等方面的指标直接用作财政运行综合绩效指标，<sup>②</sup>这种“财政管理+政府履职”设计指标体系的做法，无法将履职绩效与财政支出关联起来，无法反映重大财政政策支撑政府履职对应的成效，“以政领财、以财辅政”的匹配对应关系被模糊化了，难以循证归因，评价结果对于财政支出政策调整和支出结构优化的支撑作用难以发挥。

上述问题都是因为没有厘清财政运行成效和政府履职绩效的边界。合理的做法是聚焦一级政府的财政支出政策，按照政策投入和产出对应的原则，聚焦财政核心职能，选取基本公共服务均等化、财政支撑重大战略落实方面投入的重点领域，合理设置财政运行成效指标，避免财政支出“小马”拉动全社会发展成效“大车”的情况。

### （四）成本控制意识不够强

在当前财政紧平衡的形势之下，财政运行必须牢固树立成本意识，强调成本效益和投入产出，通过成本评价和成本管控树立正确的政绩观，避免“不计成本、铺张浪费和好大喜功”等倾向。目前除了北京市将成本节约作为财政运行综合绩效评价的一级指标之外，大部分地区对财政运行成本的测量、分析和评判主要通过财政供养比等指标进行衡量，对行政成本、城市运行成本、民生保障标准、重大项目成本管控等方面尚未设置有效指标进行评价，政府层面成本绩效评价的短板亟待完善。

### （五）评价方法不够完备

定性和定量结合、动态和静态结合、专家评判法、比较分析法、因素法、标杆法等是开展财政运行

①② 根据论文研究过程中的访谈和评价实践。

综合绩效评价的基础方法。考虑到财政运行有些指标缺少标准值的特点，目前大多数财政运行综合绩效评价会综合运用达标考核、功效系数、指数评价等方法开展，财政支出和实际产出各项指标分别进行独立评价，未进行投入产出分析和评价。事实上，财政运行成效不仅要看投入的经济性、产出的有效性，还要看投入产出的效率性，忽略投入产出评价的指标体系可能将财政运行引导到不计成本的“内卷”竞争。

#### （六）评价机制不顺畅

预算绩效管理的经验表明，项目支出绩效评价和部门整体支出绩效评价都走了一条“以评促管、以评促建”的发展路径，通过绩效评价，引导支出责任人开展全过程预算绩效管理。目前财政运行综合绩效尚处于“摸着石头过河”“先干起来”的阶段，与项目和部门整体绩效评价结果运行的广度和深度相比，评价结果对政策调整、预算安排、管理改善的影响还很不够，通过评价引领政府预算绩效管理体系建立的“指挥棒”作用发挥还不明显，亟需以评价为核心，往前端延伸加强绩效目标管理，往后端拓展加强结果运用和激励约束。

### 五、未来展望与路径

基于理论、实践和政策综合考量，财政运行综合绩效评价需要在以下方面不断深化，为深化预算管理制度提供有效支撑。

#### （一）逐步实现财政运行综合绩效评价全覆盖

目前财政运行综合绩效评价尚处于选点试点阶段，大部分省市还没有做到上级财政对全部下级政府的财政运行进行综合绩效评价，按照全面实施预算绩效管理的要求，应力争逐步实现财政运行综合绩效评价全面覆盖各级政府和各级政府全部预算收支。设计评价指标时，上级财政部门应将财政运行综合绩效评价和“十四五”财政工作重点任务结合起来，充分发挥财政在国家治理中的基础和支柱作用，立足“收、支、管、调、防”，从财力保障、预算配置、成本控制、运行成效、管理效率、风险防控、债务管理和财政可持续等方面综合衡量下级政府财政运行综合绩效，同时逐步推动总预算和“四本”预算“统分结合”的财政运行综合绩效评价体系，充分反映“四本”预算的共性要求和个性特征，推动政府预算管理制度改革的深化，提高公共产品和服务质量，助推经济社会高质量发展。

#### （二）健全财政运行综合绩效评价指标和标准体系

针对当前财政运行综合绩效指标和标准体系不够健全的问题，应由各级财政部门牵头，结合本地财政运行的实际和特点，从财政运行行为和预算功能两个维度搭建财政运行综合绩效评价指标框架，统筹财政管理绩效综合评价、地方财政管理工作考核等各类财政督查检查考核，综合协调地方政府债务评估等各类评估评价，搭建整体性、系统性、综合性的地方政府财政运行综合绩效评价指标和标准体系。财政部门和各主管部门根据各自职责分工，研究制定各项指标的标准值和合理区间，比如财政收入、财政管理方面的绩效指标由财政部门牵头制定并确定合理的指标区间；运行成效方面的指标由财政部门会同各主管部门共同研究确定指标值和合理运行区间；成本节约方面的指标由财政部门会同各主管部门明确指标含义、口径和数据来源与收集方式等。

#### （三）开展专项监测与评价

考虑到财政运行综合绩效评价涉及财政运行各方工作，各方工作虽然紧密联系，但有一定的相对独立性，比如财源建设和成本节约在管理主体、工作内容、评价标准等方面具有相对独立性。为了避免对财政运行各项工作评价“蜻蜓点水”不够深入的问题，应完善模块式、积木式组合评价机制，适时开展财力保障、成本节约、财政管理、分行业分领域运行成效、基本公共服务均等化、产业政策落实等方面的专项监测与评价。比如开展成本节约专项监测评价引导财政运行厉行节约；开展运行成效评价以财政职能为基础，以支出政策为导向，以支出预算为主线，按照政府功能分类科目为基本单元，评价财政投入、产出以及投入产出比。

#### （四）完善财政运行综合绩效评价方式方法

财政运行综合绩效评价与已经开展的项目支出绩效评价、部门整体支出绩效评价相比，样本数量有限，能够比对的绩效目标和标杆有限，对此可以考虑综合运用指数评价、功效系数、数据包络分析、定性定量结合等方法优化财政运行综合绩效评价方法。各级财政部门要完善绩效自评和上级财政评价相结合的评价机制。上级财政部门制定下发年度评价方案。下级政府组织开展综合绩效自评工作，撰写自评报告。在自评工作基础上，上级财政部门组织评价工作组，对各地上报数据信息、佐证资料、绩效自评情况

进行审核,同时获取与评价相关的统一口径数据,采用科学合理的方法,对下级政府财政综合绩效情况进行综合评定,形成财政运行评价结果。财政运行评价工作必要时可委托第三方中介机构或相关领域专家参与,提高评价的客观性和公正性。通过财政运行综合绩效评价“以评促管、以评促建”推动各级政府全过程预算绩效管理,实现政府预算与财政运行综合绩效管理一体化。

#### (五) 加强绩效评价结果运用

针对目前财政运行综合绩效评价结果运用于预算安排、政策调整、管理改善方面较为薄弱的问题,各级财政部门要加强评价结果的运用。对各评价对象财政运行状况进行分析,选取关键性指标进行标准化处理,生成各评价对象财政保障能力指数、财源建设指数、成本节约指数、规范管理指数和运行成效指数等,合成综合绩效总指数,客观反映财政运行状况和相对位置。建立一套财政运行综合绩效分析的数据采集模型,实现财政运行综合绩效动态监控,形成绩效评价长效机制。强化财政运行综合绩效评价结果运

用,将评价结果运用于财政体制改革,与管理改善、政策调整和转移支付分配挂钩。比如将财源建设评价结果与特定功能一般性转移支付分配挂钩,将运行成效评价结果作为基本公共服务均等化预算安排的调节系数等。

#### (六) 强化激励约束机制

针对财政运行综合绩效评价激励约束作用不强等问题,各级政府应将财政运行综合绩效评价结果纳入各级政府年度考核,比如北京市已将各级政府成本控制情况纳入政府绩效考核。通过财政运行综合绩效评价,将预算绩效理念传导到各级政府、财政和各个部门,牢固树立“花钱必问效,无效必问责”的理念,从政府财政运行向部门整体和项目层层传导,推动各部门、各单位形成以成本、质量、效果为核心内容的预算绩效考核体系,保障各方责任有效落实。目前预决算绩效信息公开主要针对项目和部门整体绩效,各级政府应探索将地方财政运行综合绩效评价结果随同政府决算公开,将地方财政运行综合绩效纳入政府财务报告。

## 参考文献

- [1] 吴素芳. 创新全成本预算绩效管理“北京模式”[J]. 中国财政, 2020(21): 36-39.
- [2] 王文庚, 曹堂哲. 北京市区级财政运行综合绩效评价指标概述[N]. 中国财经报, 2021-03-20.
- [3] 韩志毅, 吴国昭, 唐韶龙. 地方财政运行综合绩效评价的实践与思考[J]. 财政监督, 2021(24): 53-56.
- [4] 谢可川, 逢锦丽, 安佰欣. 下级政府财政运行综合绩效评价新思路——以山东省日照市五莲县三个乡镇试点为例[J]. 财政监督, 2022(6): 29-33.
- [5] 赵勇 徐晶. 河北廊坊试点开展县级政府财政运行综合绩效管理[J]. 中国财经报“绩效新时代”专刊, 2022-07-30.
- [6] 项前, 王兆玉. 街道财政运行综合绩效评价探索——以江苏省南京市雨花台区街道试点为例[J]. 财政监督, 2021(18): 57-61.
- [7] 孙逊. 县域财政运行综合绩效评价指标体系研究[J]. 财政监督, 2021(24): 40-48.
- [8] 张喆. 政府财政运行综合绩效评价研究[J]. 市场周刊, 2020(9): 120-123, 166.
- [9] 王宏斌, 刘敏, 程文. 关于政府财政运行情况综合绩效评价的思考[J]. 时代经贸, 2020(7): 85-87.
- [10] 刘国永, 夏和飞, 俞顺良. 地方政府财政运行综合绩效评价体系探索与实践[J]. 财政监督, 2022(5): 37-41.
- [11] 李明, 姚学辉, 王帅. 中国地方财政收入预算绩效评价: 依据、体系与结果[J]. 学习与探索, 2022(4): 90-100, 186.
- [12] 于树一. 以财政收入绩效评价深化预算管理制度改革——基于地方层面的探索[J]. 财政监督, 2021(11): 5-10.
- [13] 北京市财政局办公室. 北京市财政局聚焦九方面重点工作 2022 年全市财政工作这样干 [EB/OL]. (2022-02-07) [2022-07-03]. [http://www.mof.gov.cn/zhengwuxinxi/xinwenlianbo/beijingcaizhengxinilianbo/202201/t20220130\\_3786013.htm](http://www.mof.gov.cn/zhengwuxinxi/xinwenlianbo/beijingcaizhengxinilianbo/202201/t20220130_3786013.htm).
- [14] 杨娟, 刘继东, 杨卓文. 地方政府预算能力评估: 基于省级层面数据的研究[J]. 财政研究, 2021(2): 44-58.
- [15] 邱婧. 湖南省省本级政府财政能力评价与改进研究[D]. 湘潭大学, 2017.
- [16] 陈龙. 国家治理“3+1”架构下的财政能力集——基于公共风险视角的分析[J]. 财政研究, 2020(11): 21-32.
- [17] 辛冲冲, 杨文珂. 地方政府财政汲取能力: 分布动态、地区差异及空间收敛性[J]. 财经论丛, 2020(7): 23-32.
- [18] 杨志安, 邱国庆. 地方政府财政能力非均衡性分析: 来自辽宁的例证[J]. 地方财政研究, 2016(5): 64-70.
- [19] 中央政府门户网站. 青海省出台地方政府债务风险评估和预警暂行办法 [EB/OL]. (2016-06-28) [2022-07-03]. [http://www.gov.cn/xinwen/2016-06/28/content\\_5086271.htm](http://www.gov.cn/xinwen/2016-06/28/content_5086271.htm).
- [20] 国务院. 国务院关于进一步深化预算管理制度改革的意见: 国发〔2021〕5号 [A/OL]. (2021-3-7) [2022-07-03]. [http://www.gov.cn/zhengce/content/2021-04/13/content\\_5599346.htm](http://www.gov.cn/zhengce/content/2021-04/13/content_5599346.htm).

- [21] 侯燕明. 基于财税大数据构建地方财政运行质量评价体系的探索 [J]. 预算管理与会计, 2020 (1): 18-23.
- [22] 赵玮. 提高财政运行质量 助推县域经济高质量发展 [J]. 预算管理与会计, 2019 (4): 45-47.
- [23] 朱云飞, 赵宁. 我国地方财政质量评估与省际比较 [J]. 地方财政研究, 2019 (6): 45-53.
- [24] PEFA Secretariat. Public Financial Management Performance Measurement Framework [R]. Washington DC: PEFA Secretariat, 2011: 1-6.
- [25] 解洪涛, 陈志勇. PEFA 政府财政管理能力评价体系及在中国地方政府的适用性——兼评《地方财政管理绩效综合评价方案》[J]. 财政研究, 2016, (4): 2-14.
- [26] 周极宏. 引进国际评估系统提高财政治理效能——湖南省和世行合作促进地方政府财政管理提质增效 [M]. 中国财经报, 2022-06-14.
- [27] Bates R A, Holton III E F. Computerised Performance Monitoring: A Review of Human Resource Issues [J]. Human Resource Management Review, 1995, 5 (4): 267-288.
- [28] Baker H, Nofsinger J. Behavioral Finance: Investors, Corporations and Markets [M]. Hoboken: John Wiley & Sons, Inc., 2010: 3-7.
- [29] Ackert L, Deaves R. Behavioral Finance: Psychology, Decision-Making, and Markets [M]. Boston: South-Western Cengage Learning, 2009: 83-120.
- [30] 贾康. 当前财政政策的发力方向 [J]. 财政监督, 2020 (9): 37-43.
- [31] 新华社. 国新办举行贯彻落实“十四五”规划纲要 加快建立现代财税体制发布会 [EB/OL]. (2021-04-07) [2022-07-03]. <https://baijiahao.baidu.com/s?id=1696387750976690796&wfr=spider&for=pc>.
- [32] 江月. 预算僵化: 理论、检验及修复机制 [J]. 地方财政研究, 2020 (6): 51-59.
- [33] Schick A. A Contemporary Approach of Public Expenditure Management [M]. World Bank; Washington, D. C. 1998: 89.
- [34] 艾伦·希克. 当代公共支出管理方法 [M]. 王卫星, 译. 北京: 经济管理出版社, 2000: 112.
- [35] 马俊. 公共预算比较研究 [M]. 北京: 中央编译出版社, 2011: 23.
- [36] 孙克竞, 王祯阳. 跨期总额控制、预算执行规则弱化与辖区公共支出绩效 [J]. 财政研究, 2020 (3): 40-54.
- [37] 涂漫漫. 公共预算的功能价值与制度优化 [J]. 广东财经大学学报, 2015 (2): 105-112.
- [38] 王雍君. 公共支出管理改革: 进展与突破 [J]. 中国财政, 2016 (24): 38-39.
- [39] 财政部. 预算绩效管理规划 (2012—2015 年): 财预 [2012] 396 号 [A/OL]. (2012-10-30) [2022-07-03]. [http://www.gov.cn/zw/gk/2012-10/30/content\\_2253898.htm](http://www.gov.cn/zw/gk/2012-10/30/content_2253898.htm).
- [40] 中央政府门户网站. 财政部开展县级财政支出管理绩效综合评价 [EB/OL]. (2013-08-02) [2022-07-03]. [http://www.gov.cn/govweb/jrzg/2013-08/02/content\\_2460238.htm](http://www.gov.cn/govweb/jrzg/2013-08/02/content_2460238.htm).
- [41] 财政部. 财政部关于印发《2015 年县级财政管理绩效综合评价方案》的通知: 财预 [2015] 139 号 [A/OL]. (2015-07-24) [2022-07-03]. [http://www.mof.gov.cn/gkml/caizhengwengao/wg2015/wg201509/201601/t20160120\\_1652607.htm](http://www.mof.gov.cn/gkml/caizhengwengao/wg2015/wg201509/201601/t20160120_1652607.htm).
- [42] 财政部. 2018 年县级财政管理绩效综合评价方案及结果的通知: 财预 [2018] 100 号 [A]. 2018-07-10.
- [43] 中国政府网. 财政部印发《地方财政管理绩效综合评价方案》: 财预 [2014] 45 号 [A/OL]. (2014-04-17) [2022-07-03]. [http://www.gov.cn/xinwen/2014-04/17/content\\_2661108.htm](http://www.gov.cn/xinwen/2014-04/17/content_2661108.htm).
- [44] 财政部. 财政部关于印发《财政管理工作绩效考核与激励办法》的通知: 财预 [2018] 4 号 [A/OL]. (2018-01-08) [2022-07-03]. [http://yss.mof.gov.cn/zhengceguizhang/201802/t20180223\\_2817079.htm](http://yss.mof.gov.cn/zhengceguizhang/201802/t20180223_2817079.htm).
- [45] 财政部. 财政部关于修订《财政管理工作绩效考核与激励办法》的通知: 财预 [2018] 222 号 [A/OL]. (2018-12-29) [2022-07-03]. [http://yss.mof.gov.cn/zhengceguizhang/201902/t20190228\\_3179954.htm](http://yss.mof.gov.cn/zhengceguizhang/201902/t20190228_3179954.htm).
- [46] 财政部. 关于印发《地方财政管理工作考核与激励办法》的通知: 财预 [2020] 3 号 [A/OL]. (2020-01-03) [2022-07-03]. [http://yss.mof.gov.cn/zhengceguizhang/202001/t20200117\\_3460945.htm](http://yss.mof.gov.cn/zhengceguizhang/202001/t20200117_3460945.htm).
- [47] 财政部. 山东省济宁市财政: 实行市县乡三级联动全面提升财政绩效全流程管理水平 [EB/OL]. (2021-12-03) [2022-07-03]. [http://www.mof.gov.cn/zhengwuxinxi/xinwenlianbo/shandongcaizhengxinxi/xinwenlianbo/202112/t20211203\\_3771695.htm](http://www.mof.gov.cn/zhengwuxinxi/xinwenlianbo/shandongcaizhengxinxi/xinwenlianbo/202112/t20211203_3771695.htm)
- [48] 俞晓丽, 王婷. 义乌首次实施镇街财政运行综合绩效评价 [N/OL]. 浙江日报 (2019-06-24) [2022-07-03]. <https://baijiahao.baidu.com/s?id=1637177518284515755&wfr=spider&for=pc>.
- [49] 弘文. 浙江东阳迈出乡镇政府整体绩效评价全覆盖第一步 [N]. 中国财经报, 2019-01-26.
- [50] 苗宇. 辽宁财政启动对下级政府财政运行综合绩效评价试点 [N]. 中国财经报, 2020-10-24.
- [51] 马光荣, 吕冰洋. 2021 中国各地区财政发展指数报告 [R]. 北京: 中国人民大学财税研究所, 2022.
- [52] 济南市财政局. 济南市县级政府财政运行综合绩效评价体系“上岗”构建全方位预算绩效管理新格局 [EB/OL]. (2021-04-11) [2022-07-03]. [http://jncz.jinan.gov.cn/art/2021/4/11/art\\_7850\\_4764815.html](http://jncz.jinan.gov.cn/art/2021/4/11/art_7850_4764815.html).

(责任编辑: 孙亦军 张安平)

# 财政生态补偿：一个理论逻辑

## Fiscal Ecological Compensation: A Theoretical Logic Perspective

陈少强 覃凤琴

CHEN Shao-qiang QIN Feng-qin

**[摘要]** 生态补偿是我国为推进生态文明建设而实施的一项改革举措，中央和地方政府也出台了一系列生态补偿财政政策，但学术界对生态补偿的财政理论研究明显不足。研究财政生态补偿必须结合中国国情，笔者尝试从国家所有者和公共管理者的“双重身份”视角展开研究。政府是国家所有者的代表，国有资产所有者身份赋予了政府财产权利，政府以自然资源所有者身份向国有企业征收部分红利，以探矿权、采矿权等非税收收入的形式向部分开采企业征收探矿权、采矿权使用金，并通过财政补助和转移支付等方式支持国家生态安全区和重要生态功能区。社会管理者身份赋予了政府公共权力，政府应承担资源配置、公平分配和监督管理等职能。沿着这一逻辑，本文总结了我国财政生态补偿的实践，并提出今后要着力解决的理论和实践问题。

**[关键词]** 国家所有者 公共管理者 财政生态补偿

**[中图分类号]** F812 **[文献标识码]** A **[文章编号]** 1000-1549 (2022) 11-0017-09

**Abstract:** Ecological compensation is a reform measure implemented in China to promote the construction of ecological civilization. So far the central and local governments in China have introduced a series of fiscal ecological compensation policies, but relevant academic research seems to be obviously insufficient. The study of fiscal ecological compensation should be based on China's national conditions. This paper studies from the perspective of dual identities, i. e., state owners and public managers. The government is the representative of the owner of the country, and the ownership of state-owned assets grants the government property rights, levying some dividends from state-owned enterprises as owners of natural resources, and collecting exploration rights and mining rights from some mining companies in the form of non-tax revenue such as exploration rights premium and mining rights premium. Subsidies and transfer payments are used to support national ecological security zones and important ecological function zones. The status of public manager endows the government with the public power to undertake the functions of resource allocation, fair distribution and supervision and management. Following the above logic, this paper summarizes the practice of fiscal ecological compensation in China, and proposes theoretical and practical problems to be solved in the future.

**Key words:** State owner Public manager Fiscal ecological compensation

**[收稿日期]** 2022-04-02

**[作者简介]** 陈少强，男，1971年12月生，中国财政科学研究院资源环境和生态文明研究中心研究员，博士生导师，研究方向为资源环境财政政策；覃凤琴，女，1984年11月生，中国财政科学研究院资源环境和生态文明研究中心助理研究员，经济学博士，研究方向为环境公共经济学。本文通讯作者为覃凤琴，联系方式为 qinfengqin@chineseafs.org。

**[基金项目]** 山东省社科规划项目“山东省财政生态补偿的绿色发展效应研究”（项目编号：21DJJJ01）；新安江-千岛湖生态补偿试验区建设的财政支持政策研究。

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、导言

研究生态补偿财政问题，主要基于以下考量：

一是生态补偿财政理论研究明显滞后于生态补偿财政实践发展。生态补偿是改革开放以来特别是党的十八大以来我国在生态环境保护方面的一项重要举措，它作为生态文明体制的“四梁八柱”之一已引起了社会广泛重视，我国也逐步确定了加快健全有效市场和有为政府更好结合、分类补偿与综合补偿统筹兼顾的生态补偿发展目标。财政作为国家治理的基础和重要支柱，在推进生态补偿方面采取了一系列举措，但生态补偿财政方面的理论文献严重不足，理论研究明显滞后于实践发展。笔者在中国知网数据库（CNKI）中以“生态补偿”为主题进行了搜索，发现生态补偿相关文献有2万多篇，其中关于生态补偿理论的研究仅有500多篇，与其相关的财政理论研究只有20篇左右，而关于生态补偿财政职能研究的文献则几乎没有。

二是西方的经济理论和公共财政理论对我国生态补偿财政政策实践解释乏力。传统的庇古税和科斯定理理论难以解释中国中央和地方财政共同应对横向流域生态补偿实践问题，特别是上下游区域政府间的对赌协议很难用西方财政理论来解释，生态补偿财政理论的研究应结合中国社会主义公有制的基本国情，深入推进基于所有者为基础的生态补偿财政理论研究成为必然。

三是由于市场发展阶段等方面的差异，国内外生态补偿实践有较大差别。国外生态系统服务付费是基于买卖双方自愿下的市场行为，中国生态补偿虽然也是以市场化、多元化为重要目标，但是在实践层面，生态补偿对财政政策的依赖度较高，相关生态补偿财政职能边界问题需要引起关注。

综上，笔者认为在借鉴国外生态补偿经验及其财政实践的基础上，生态补偿财政理论研究要更多地立足我国国情，以讲好财政生态补偿的“中国故事”，推动我国生态文明建设财政理论建设向纵深发展，也为创新我国财政理论研究奠定基础。

## 二、文献综述

### （一）关于生态补偿

我国生态补偿的概念源于国外的生态系统服务付

费（Payment for Ecosystem Services, PES）或生态环境服务付费（Payment for Environmental Services, PES）。国外对PES概念表述不尽相同，较为典型的是Wunder（2005）<sup>[1]</sup>提出的观点，即PES是一个自愿的、附条件的交易，至少有一个卖方，一个买方，并且具有明确定义的生态系统服务。这一定义是基于科斯定理的市场私人交易，并且遵循使用者付费原则。在实践中，除了私人参与外，还有比如哥斯达黎加国家生态补偿项目（PSA）为代表的公共项目，以及津巴布韦本土资源公共区域管理项目（CAMPFIRE）为代表的公私混合生态补偿（Ezzine-de-Blas D等，2016<sup>[2]</sup>）。因此，Muradian等人将生态补偿定义为社会参与者之间的资源（包括现金和非现金）转移，这种资源转移可以通过市场或者准市场的方式进行，也可以通过公共政策（如补贴）的方式进行（张晏，2016<sup>[3]</sup>）。

我国学者于20世纪80年代提出生态补偿概念（张诚谦，1987<sup>[4]</sup>）。《国务院关于落实科学发展观加强环境保护的决定》（国发〔2005〕39号）提出尽快建立生态补偿机制和开展生态补偿试点，“生态补偿”首次在政府文件出现，当时这一概念与环境赔偿关系较为紧密<sup>①</sup>。目前，国内对生态补偿的理解尚不完全统一。较为流行的观点认为，生态补偿机制是以保护生态环境、促进人与自然和谐发展为目的，根据生态系统服务价值、生态保护成本、发展机会成本，运用政府和市场手段，调节生态保护利益相关者之间利益关系的制度安排（刘桂环等，2021<sup>[5]</sup>）。这种观点较好地借鉴了国外理论，但对于生态保护利益相关者到底是谁并没有明确说明，结果可能导致在实践中难以操作执行。

### （二）关于生态补偿财政职能

生态补偿概念在20世纪80年代后期才提出并于21世纪初才正式体现在国家政策中，在此期间几乎没有相关财政职能研究文献，学术界关于财政职能的研究和讨论集中在国家分配论和公共财政论方面。

在计划经济时代，主流财政理论国家分配论（邓子基，1962<sup>[6]</sup>；许廷星，1982<sup>[7]</sup>；许毅，1995<sup>[8]</sup>），将财政定义为一种特殊的分配关系，也就是社会主义财政分配理论。国家分配论继承了马克思的劳动价值

① 该《决定》在加强环境监管制度部分提出，“国家加强跨省界环境执法及污染纠纷的协调，上游省份排污对下游省份造成污染事故的，上游省级人民政府应当承担赔付补偿责任，并依法追究相关单位和人员的责任。赔付补偿的具体办法由环保总局会同有关部门拟定。”



论的理论思想,但是对自然资源或者生态环境的价值没有提及。

在建立社会主义市场经济体制过程中,学术界陆续讨论基于市场失灵基础上的公共财政理论,到20世纪90年代关于国家分配论与公共财政论的讨论非常热烈,一大批专家学者都参与了关于财政职能的讨论(贾康,1994<sup>[9]</sup>;郭庆旺等,1999<sup>[10]</sup>;邓子基,2000<sup>[11]</sup>)。伴随着社会主义市场经济理论的逐步确立,学术界也就开始研究社会主义市场经济条件下的财政职能问题,楼继伟(2006)<sup>[12]</sup>、张馨等(2000)<sup>[13]</sup>、高培勇(2012)<sup>[14]</sup>等将财政职能概括为:优化资源配置、调节收入分配和促进经济稳定等。西方公共财政论中的“外部性”在一定程度上解释了财政生态补偿的必要性和存在依据,但对于我国这样一个以公有制为主体的国家而言,财政作为国家资源资产所有者代表如何履行生态补偿职能却并没有引起重视。

当然,在国家分配论和公共财政论之间也有一种折中的理论,就是既承认财政的国家所有者代表的职能,又承认公共财政相关职能,如叶振鹏(1998)<sup>[15]</sup>提出社会主义市场经济条件下的财政分为两种:一是以政权组织的身份为主体,为满足社会公共需要而进行的财政分配,即公共财政;二是以全民所有制生产资料所有者身份为主体,为维护所有者的权益和进行国有资产宏观经营而进行的财政分配,即国有资产经营财政。这种理论观点认为,公共财政和国有资产经营财政都是以国家分配为主体,为实现国家的职能,依据国家的权力进行的政府分配行为,是国家通过财政收支所进行的计划资源配置活动,因而本质上归属于国家分配论。这种理论注意到财政发挥职能作用要以公有制为前提,但其缺陷在于,看到了国有企业和国有资产的经济价值,没有关注自然资源的生态价值;看到了要处理好财政与国家企业的经济分配关系,但没有看到财政与全民所有制自然资源的分配关系,对于财政如何维护、保护和分配自然资源价值,对于财政作为所有者如何对自然资源生态价值进行补偿则更是没有提及。

党的十八届三中全会提出了“财政是国家治理的基础和重要支柱”的重要论断,财政从经济管理范畴演变为国家治理范畴,财政职能也就有了新的表述:优化资源配置、维护市场统一、促进社会公平和实现国家长治久安(高培勇,2018<sup>[16]</sup>;高培勇,2021<sup>[17]</sup>)。我国提出从全面小康到共同富裕、生态文明建设向纵深推进、碳达峰碳中和战略目标和美丽中国建设目标

以来,有关财政职能研究视野也逐步拓展,例如,陈诗一和祁毓(2022)<sup>[18]</sup>研究了应对气候变化的财政政策 and 财政职能,刘尚希(2018)<sup>[19]</sup>指出全球公共风险条件下我国需要发挥大国财政职能。

随着我国生态补偿概念的正式引入,学术界也开始关注相关财政理论问题。孙亦军(2006)<sup>[20]</sup>认为当前学术界都无一例外地把财政分配活动局限在了GDP或国民收入的再分配领域,但对于国家所有者的利益却并没有给予足够的关注。钟大能(2006)<sup>[21]</sup>以西部民族地区为例,提出建立一种兼容政府纵向和省区际间、流域间政府横向的财政补偿机制。

总体看来,我国有关生态补偿财政职能的文献较少,相关文献主要集中在以下三个方面:一是介绍国外生态补偿财政职责立法。徐丽媛(2017)<sup>[22]</sup>提出德国联邦政府负责政策制定以及协调跨界补偿,州政府主要是立法责任(细化和完善相关法律),地方主要履行评估、监督管理以及出资责任。二是讨论流域生态补偿财政职责。郑云辰等(2019)<sup>[23]</sup>提出中央政府负责生态补偿顶层设计、统筹协调以及全国性生态补偿的出资,省级政府负责制定政策以及区域性生态补偿出资,地方政府主要是执行和落实。三是明确资源环境领域财政事权和支出责任的划分。《国务院办公厅关于印发生态环境领域中央与地方财政事权和支出责任划分改革方案的通知》(国办发〔2020〕13号)规定了生态环境规划制度制定、生态环境监测执法、生态环境管理事务与能力建设、环境污染防治、生态环境领域其他事项的事权和支出责任划分。《国务院办公厅关于印发自然资源领域中央与地方财政事权和支出责任划分改革方案的通知》(国办发〔2020〕19号)规定了自然资源产权管理、自然资源有偿使用和权益管理、国土空间规划和用途管制、生态保护修复、自然资源安全、自然资源领域灾害防治以及自然资源领域其他事项的事权和支出责任划分。

生态补偿是我国生态文明建设中的一项改革举措,我国在推进财政生态补偿方面采取了一系列举措,但相关财政职能理论研究明显滞后于财政政策实践,西方财政理论也难以有效解释中国丰富的财政生态补偿政策实践。本文从国家所有者和公共管理者的视角进行研究,为研究我国生态补偿财政职能提供一个新的理论逻辑。

### 三、财政生态补偿的理论逻辑

财政在生态补偿中发挥职能作用的理论逻辑有两

个维度：一是所有者维度，财政作为所有者代表履行出资人职能；二是公共管理者维度，财政作为政府公共管理部门履行公共管理者职能。

### （一）作为国家所有者代表的逻辑

财政作为国家资源资产所有者和出责者代表，在国有资产分配和生态补偿中应承担起应有的作用。马克思在《哥达纲领批判》<sup>①</sup>中批判了拉萨尔所谓的“公平分配不折不扣的劳动所得”之类口号的虚伪性和荒谬性，并提出在生产资料公有制条件下社会总产品分配的“社会扣除理论”，这些用来被扣除的内容包括：第一，用来补偿消费掉的生产资料的部分；第二，用来扩大再生产的追加部分；第三，用来应付不幸事故、自然灾害等的后备基金或保险基金；第四，和生产没有直接关系的一般管理费用；第五，用来满足共同需要的部分，如学校、保健设施等；第六，为丧失劳动能力的人等设立的基金。马克思还指出，学校、保健设施等社会共同需要的部分一开始就会显著增加，并将随着新社会的发展而日益增加（马克思，2012<sup>[24]</sup>）。马列主义经典作家在很早就预测到，公有制国家要对学校、保健设施社会公共需要进行保障，而优美的自然环境显然也是社会公共需要的一个组成部分，也应是马列主义经典作家认为应当扣除的一部分，这为社会主义国家财政支持包括生态补偿在内的社会共同消费提供了理论支撑。

在计划经济时代，我国提出将国家分配论作为财政学基本理论的思想，但是自然资源或者生态环境的价值没有被提及，这主要是源于国家作为所有者实施了统收统支财政体制，当时也没有可以参与交换的市场主体，生态产品的交换价值无法得到体现。

随着经济的快速发展，特别是到了20世纪90年代，资源环境的补偿问题逐步显现。孙毅等（1990）<sup>[25]</sup>在《补偿经济论》一书中提出造成自然资源危机和生态环境破坏的主要原因是价值补偿不足，从根源上来说，就是绝大多数资源型初级产品价格在其形成中并没有对资源开采和耗用进行核算，资源等生产要素是游离于社会再生产价值之外的，资源不能以一般商品进入市场，导致该类产品的价值难以计算，这种“忽略”和“游离”造成人们形成了自然资源等要素价值不高甚至没有价值的意识，补偿不足是导致经济

社会发展失衡和生活环境恶化的重要原因之一。当然，形成这样认识的原因很多，一个重要原因是在社会主义经济转轨过程中，国有资产的管理体制尚不健全，履行出资人的职责尚未到位。

党的十八大以来生态文明建设不断推进，自然资源的作用和价值进一步彰显。习近平生态文明思想和习近平经济思想也都多次强调要发挥财政在国家治理中的基础和重要支柱的作用，重视财政在生态补偿领域的职能作用，在财政理论研究中重点关注其在生态补偿领域的财政职能理论的研究，为生态补偿实践提供有力的理论支撑，也为生态文明建设奠定理论基础。

全民资产所有者的身份赋予了政府财产权利。相应地，财政所有者职能作用体现在：一是国家以自然资源所有者身份向国有企业征收部分红利，弥补国有企业在资源开发利用过程中对资源环境造成的损害，并通过国有资本经营预算支出等方式对生态产品价值进行补偿；二是国家以自然资源所有者身份，以探矿权、采矿权等非税收入的形式，向部分开采企业征收探矿权、采矿权使用金，以弥补生态价值；三是国家以自然资源所有者身份，通过财政补助和转移支付等方式支持国家生态安全区和重要生态功能区，以补偿其为国家生态环境做出的贡献。

### （二）作为公共管理者的逻辑

在社会主义公有制国家，政府具有国有资产所有者和社会管理者的双重身份<sup>②</sup>。作为公共管理者，财政在生态补偿中也发挥重要职能作用，特别是在市场经济条件下，财政生态补偿的职能作用更多的是以公共财政的形式显现出来。具体而言，生态补偿公共财政的职能作用主要体现在：

#### 1. 资源配置职能。

资源配置职能即财政配置资源用于生态补偿，具体职能实现方式有：

一是配置财政资源，实现生态补偿的内部循环。政府资源的方式既有计划配置资源，如通过预算拨款程序实现生态补偿，也有通过创造市场化的方式配置政府资源。

二是配置市场资源。政府通过政府采购、财政支

① 《哥达纲领批判》在马克思生前并没有发表过，该文是恩格斯为了反击德国党内日渐抬头的机会主义思潮，根据马克思的《德国工人党纲领批注》和1875年5月5日给威廉·白拉克的信归集而于1891年1月在《新时代》上发表的。

② 严格来讲，政府仅仅是国家和人民的代表。如《民法典》第246条规定：“法律规定属于国家所有的财产，属于国家所有即全民所有。国有财产由国务院代表国家行使所有权。法律另有规定的，依照其规定。”

出、政府补助、税收优惠、政府和社会资本合作 (PPP) 模式等财政政策手段, 支持和引导市场化主体进入生态补偿领域, 实现自然资源和生态环境保护的外溢价值的内部化。中央和地方财政通过调动社会资本和资源共同配置资源, 多元化市场化生态补偿, 进而达到生态补偿的外部循环。

### 2. 公平分配职能。

生态补偿的财政公平分配职能体现为通过财政政策和资金的导向作用, 实现生态补偿的使用者和受益者之间的成本收益公平分担和分配。由于生态补偿的使用者和受益者既可以是地区 (区域), 也可以是政府和企业, 因而生态补偿财政分配职能又与区域均衡协调发展、社会阶层和成员之间的社会公平等问题密切相关。

当然, 代际之间的公平也是生态补偿财政职能需要考虑的重要方面。就生态补偿而言, 客观上存在受益者和提供者可能不在同一时期或者同一时代, 在生态补偿保护方面, 客观上存在前人保护、后人受益的“前人栽树、后人乘凉”问题, 也存在前人破坏、后人遭殃的“吃子孙饭”问题。财政作为管理者身份的代表, 其重要职责是通过预算政策等管理手段实现全生命周期生态补偿的客观公正, 以平滑代际之间的成本受益公平负担和分配。

### 3. 监督管理职能。

作为政府公共管理者, 中央财政和地方各级财政在生态补偿中也发挥其监督管理职能, 这些职能包括但不限于预算编制、预算执行、政府采购、财政监督和绩效管理。

## 四、财政生态补偿的实践探索

### (一) 所有者视角下的财政生态补偿

财政以所有者身份进行生态补偿, 既可以针对国有企业的开采利用而进行生态补偿, 也可以针对特定国家重点生态功能区生态保护而进行补偿; 既可以通过国有资本经营预算而进行补偿, 也可以通过一般公共预算而进行补偿。

1. 通过国有资本经营预算支出保护自然生态环境。

国家以自然资源所有者身份向国有企业征收部分红利, 并通过国有资本经营预算支出方式对生态产品价值进行补偿。例如, 石油石化企业、煤炭企业等资源型企业向国家上缴利润, 并通过国有资本预算支出于油气管道建设和环境保护等, 以弥补国有企业在资源开发利用过程中对资源环境造成的损害。

2. 通过征收资源类权益金保护国家所有的自然资源。

国家以自然资源所有者身份, 以探矿权、采矿权等非税收收入的形式, 向部分开采企业征收探矿权采矿权使用金, 以弥补生态价值。

矿产资源开发利用的税费由来已久, 也是自然资源中较早被作为有偿使用的资源之一, 自 1986 年立法以来, 开采矿山资源的相关税费备受关注, 也在不断地改革, 矿产资源生态补偿经历了征收生态补偿费、缴存保证金到现在生态补偿基金、资源税、环保税的多元补偿阶段 (见表 1)。

表 1 矿产资源生态补偿发展历程

年份	政策文件	主要内容
1986	《中华人民共和国矿产资源法》	开采矿产资源必须按照国家有关规定缴纳资源税和资源补偿费; 给他人生产生活造成损失的除赔偿外, 还要采取必要的补救措施
2003	《矿山地质环境专项资金》	支持地方开展历史遗留和矿业权人灭失矿山的地质环境治理
2006	《关于逐步建立矿山环境治理和生态恢复责任机制的指导意见》(财建〔2006〕215号)	按矿产品销售收入的一定比例, 由矿山企业分年预提矿山环境治理恢复保证金
2009	《矿山地质环境保护规定》	采矿权人应当依照国家有关规定, 缴存矿山地质环境治理恢复保证金
2010	《国土资源部关于将矿产资源专项收入统筹安排使用的通知》	国家设立矿山地质环境专项资金, 支持地方开展历史遗留和矿业权人灭失矿山的地质环境治理

资料来源: 根据官方网站整理。

3. 通过转移支付等体现国家对重点生态功能区的保护。

国家以自然资源所有者身份, 通过财政补助和转移支付等方式支持国家生态安全区和重要生态功能区等, 以保护长江、黄河等关乎中华民族永续发展的生

态资源, 以补偿其为国家生态环境保护所做出的贡献。纵向生态补偿主要是通过一般转移支付和专项转移支付对重要生态系统和重点生态功能区进行补偿。早期阶段, 生态补偿主要是通过专项转移支付对我国与生态建设工程相关的生态保护行动进行补偿。到

2001年后,国家又陆续向森林生态效益、矿产资源、退耕还林还草等项目进行资金支持和物资补偿。之后,生态补偿范围也逐渐扩大到荒漠、海洋、水流、土地等领域。自2008年起,国家在一般转移支付中设立重点生态功能区转移支付,对主体功能区中的禁止开发区与部分限制开发区进行补偿,生态补偿范围已经从当初的221个县扩大到818个县。根据笔者汇总,重点生态功能区生态补偿资金从2008年的60.51亿元提高到2022年的984.02亿元,主要是用于弥补生态系统服务供给地区丧失的发展机会成本,同时也用于生态环境保护和民生的改善。

## (二) 公共管理者视角下的财政生态补偿

### 1. 财政配置各类资源进行生态补偿。

财政生态补偿的资源配置职能体现在以下三个方面:

(1) 预算计划配置。政府通过预算计划的方式配置生态补偿资源,即政府在编制年度预算之后,根据预算标准、支出结构等方式支持生态补偿。这是一种常见的财政生态补偿方式。

(2) 竞争性评审方式配置。这种方式仍然以预算计划配置方式为主,但借鉴了国外新公共管理的部分特点,通过多个地方政府参与中央财政专项评审,地方政府之间作为“准市场主体”参与竞争,中央政府最终让有实力、有意愿、准备充分的地方政府率先获得中央财政资源。竞争性评审方式在资源环境和经济建设等领域得到较为广泛的运用。例如,2016年以来,财政部等国家部委通过竞争性评审方式在全国24个省(自治区、直辖市)共安排25个山水林田湖草生态保护修复工程试点,涉及24个省(直辖市、自治区),覆盖约111万平方公里的国土面积,第一批和第二批每个试点项目中央财政给予基础奖补资金20亿元,第三批每个试点10亿元,从2016年开始合计安排补助资金360亿元,对提升区域生态系统服务功能和生态环境质量发挥了重要作用。

(3) 引导市场配置。财政生态补偿也可以通过

政府引导市场方式配置资源,这种生态补偿属于政府市场共治类型的生态补偿,也是通常所说的多元化市场化生态补偿。这种类型的补偿表现形式多样,包括但不限于:水权交易试点、绿色基金、政府和社会资本合作、产业和人才补偿、碳交易市场等,这类补偿的核心要义是政府和市场形成各类契约关系,双方或者多方共担风险和收益,与国外的PES方式有一定类似。

### 2. 财政生态补偿体现社会公平分配。

财政在生态补偿中也体现公平分配职能,主要表现在:

(1) 横向公平分配。财政生态补偿体现财政的公共分配职能,即促进区域协调均衡发展,在实现生态补偿的同时缓解和解决区域之间、地区之间、阶层之间的发展不平衡、不充分的问题。例如,中央财政在“十三五”期间支持的山水林田湖草生态保护修复工程试点,主要保障了“两屏三带”<sup>①</sup>国家生态安全战略格局的任务,惠及65个国家级贫困县,国家在保护生态资源的同时也促进落后地区的发展。

(2) 纵向公平分配。纵向公平主要是解决生态补偿的代际公平问题。由于自然资源有限,可持续增长要求当代人不能以损害后代人用于满足所需要的自然资源和环境为条件,要还后代人公平地开发利用自然资源和环境的权利。通过生态补偿可以限制当代人对自然资源的挥霍,为后代人留下一个可持续的生态环境。

党中央、国务院高度重视生态环境保护和修复,特别是党的十八大以来,在全面加强生态环境保护的基础上,持续加大生态修复的力度,持续推进国土绿化、山水林田湖草沙生态保护修复工程、海洋生态修复等重点生态工程,主要用于解决生态环境破坏等历史遗留问题,不仅是对上代人的环境破坏的弥补,也是为了后代人能够享受美好的生态环境,在很大程度上实现了代际公平(见表2)。

表2 中央财政支持的部分生态保护修复工程

项目	牵头单位	基本情况	投入资金
山水林田湖草生态保护修复工程试点	财政部、原国土资源部、原环境保护部	2016—2018年实施了3批25个试点,每批试点3年	中央财政按照首批和第二批每个试点20亿元、第三批每个试点10亿元的标准下拨资金,2016年以来,中央财政已累计下达重点生态保护修复治理资金360亿元

① “两屏三带”是指青藏高原生态屏障、黄土高原-川滇生态屏障、东北森林带、北方防沙带、南方丘陵带。

续前表

项目	牵头单位	基本情况	投入资金
蓝色海湾整治行动 (海洋生态修复)	财政部、原国土资源部、原环境保护部、国家海洋局	2016年开始,连续实施了4年	中央财政按照每个直辖市、省会城市、计划单列市总额4亿元,一般地级市总额3亿元的标准安排奖励资金
水林田湖草沙一体化保护和修复工程项目	财政部办公厅、自然资源部办公厅、生态环境部办公厅	“十四五”期间开展,2021年开始第一批10个项目、2022年第二批9个项目,截至目前共19个项目	中央财政通过竞争性评审方式给予地方项目支持,每个项目中央财政奖补额度不超过20亿元

资料来源:根据财政部等官方网站资料整理。

### 3. 财政监督管理生态补偿资金和政策。

政府作为公共管理者,对生态补偿资金和政策运行情况进行监督,以提升财政治理能力和效果。生态补偿财政监管职能包括中央和地方财政出台生态补偿资金管理办法(见表3)、财政监督管理办法以及绩效评估管理办法等。从财政部等部委的生态补偿财政

政策文件来看,财政监督和绩效考核成为几乎每个政策文件不可或缺的内容。例如,财政部出台的《国家重点生态功能区转移支付办法》提出,对重点生态功能区的生态环境质量的考核就直接决定了奖惩资金的情况,并对资金的用途进行严格管理。

表3 部分财政生态补偿监督管理文件

序号	文件名称	发文单位	文号	生态补偿财政政策内容
1	《中央财政草原生态保护补助奖励资金管理暂行办法》	财政部 农业部	财农〔2011〕532号	财政部门负责安排补奖资金预算,会同农牧部门制定资金分配方案,拨付和发放资金,监督检查资金使用管理情况,组织开展绩效考评等
2	《中央财政林业补助资金管理办法》	财政部 林业局	财农〔2014〕9号	补助资金用于森林生态效益补偿、林业补贴、森林公安、国有林场改革等方面
3	《重点生态保护修复治理资金管理办法》	财政部	财资环〔2021〕100号	治理资金使用要着眼全局,立足中央层面,主要用于共同财政事权事项,支持开展对生态安全具有重要保障作用、生态受益范围较广的重点生态保护修复
4	关于修订《林业草原生态保护恢复资金管理办法》的通知	财政部 国家林草局	财资环〔2021〕76号	国家公园补助用于国家公园勘界、自然资源调查监测等相关支出

资料来源:根据财政部等相关官方网站整理。

政府对生态补偿财政资金和财政政策的监督,既有合规性的监督,如监督检查程序是否到位,也有对财政政策和财政资金绩效的监督。近年来,各级财政通过大力实施预算绩效管理和评估,对相关资金进行监督管理。例如,在跨省流域生态补偿中,中央和地方财政对项目进展进行评估,主要以水质考核为主,根据考核结果给予地方一定的扣缴或者奖励资金。

#### (三) 财政生态补偿实践具有混合性的特点

由于财政具有国家所有者和政府公共管理者的双重身份,在财政生态补偿实践中也体现出混合性特征。

##### 1. 对象的混合性。

我国财政生态补偿无论是针对重点生态功能区(如三江源重点生态功能区)还是针对大江大河(如长江、黄河),都体现出国家意志,体现出国家的生态安全战略和区域协调发展战略。在这些战略目标实现过程中,财政作为国家所有者和公共管理的双重代表身份出现,财政生态补偿的对象和范围出现一定的

交叉重叠。

##### 2. 手段的混合性。

财政生态补偿都是通过预算的方式进行调节的。财政作为所有者代表,更多通过国有资本经营预算和一般公共预算方式参与生态补偿,而作为政府公共管理代表的财政则是通过一般公共预算方式参与生态补偿的。

##### 3. 方式的混合性。

财政生态补偿越来越体现混合性。生态补偿在早期以中央财政为主体,甚至在一定的时期比如退耕还林等生态建设工程只有中央财政参与;随后地方政府开始参与,形成央地合力;当前我国提出生态补偿的有为政府和有效市场相结合的目标,财政生态补偿方式的混合性特征更为突出。

### 五、财政生态补偿应着力解决的理论和实践问题

我国在财政生态补偿方面进行了卓有成效的探

索,但仍面临着一些理论和实践问题需要解决。在财政生态补偿理论层面,要着力解决生态产品价值确认和双重身份下财政职能边界问题。在财政生态补偿实践层面,要着力解决所有者代表的职能部门不清晰以及生态补偿历史遗留复杂的问题,以实现有效率的补偿、体现公平的补偿。

### (一) 应着力解决的理论问题

#### 1. 生态产品价值确认问题。

国内对建立健全生态产品价值实现机制已形成广泛共识,它已成为实现生态补偿的重要方面。生态产品的价值确认是实现生态产品价值的前置性条件,没有对生态产品价值确认,生态补偿无从谈起。遗憾的是,有关生态产品价值确认却面临以下几个理论难题,需要引起学术界的足够重视。

一是生态产品的外延边界。目前,国内学者普遍认为生态产品是来自自然生态系统的物质与服务,但也有学者认为我国对生态产品的外延范围界定的研究还不够(黄宇驰等,2022<sup>[26]</sup>)。

二是生态产品价值的界定。当前,国内关于生态产品价值的讨论比较热烈,但国内尚未形成完全统一的生态产品价值观。国内有学者提出生态产品价值可以分为市场价值(主要包括生态物质产品供给、生态文化产品供给)和非市场价值(主要包括生态服务产品供给)(李宏伟等,2020<sup>[27]</sup>),也有学者提出生态产品价值可以分为直接使用价值和间接使用价值(詹瑜璐和杨建州,2022<sup>[28]</sup>)。总之,国内对生态产品价值的研究仍处于探索和试点阶段。

三是生态产品价值的确认。国内即便就生态产品及其价值形成统一认识,有关生态产品价值的确认也需要理论上进一步研究。首先,不同产权的生态产品的价值确认带来难度。Costanza(2008)<sup>[29]</sup>认为,生态产品可以分为市场产品、俱乐部产品、公共池塘资源、拥挤公共产品以及纯公共产品,Kemkes等(2010)<sup>[30]</sup>则指出,纯粹的市场产品、俱乐部产品等产权属性相对清晰,可以借助市场交易机制实现其价值;而对于公共池塘资源、拥挤产品以及纯公共产品而言,由于其产权难以界定而容易导致相关主体“搭便车”现象,供给主体积极性不高而容易导致“公有地悲剧”的发生。换言之,Kemkes所说的公共产品(实际上是国有自然资源)的价值确认在客观上存在难度。其次,生态产品价值确定在方法论体系上还不完善。生态产品价值核算(GEP)总体上还处于试点阶段,资源资产负债表尚未真正建立,尚

未形成全国性的实体产品价值核算体系和标准,生态环境流域的投资成本收益和风险的识别、确认和计量等基础工作还不能满足生态补偿实践需要。高晓龙等(2022)<sup>[31]</sup>认为我国现有有关生态产品价值实现的概念、核算标准、市场分类、交易平台等尚不完善,仍然只是停留在试点阶段,未来还有很长的路要走。

#### 2. 财政职能边界划分问题。

明确财政作为国家所有者和公共管理者身份的代表,有利于明确财政职能作用的方向,但这样也带来另外一个理论问题就是,作为所有者代表的财政职能边界与作为公共管理者代表的财政职能边界如何划分。客观说来,双重身份下财政职能都有资源配置和公平分配等职能,但实现这些职能的方式是不同的。作为所有者的财政资源配置职能,更多的配置来自庞大的国有财产和国有企业的资源,资源配置方式更多的是“取之于国、用之于国”,其公平分配调节的对象是具有国家战略意义的特定区域(如生态功能区等),而基于公共管理者代表身份的财政职能的资源配置职能方式更多是通过配置来自公共管理管辖权内的企业和其他主体的资源,是典型的“取之于民、用之于民”,其调节公平分配的方式也更加宽泛。这两种财政职能履行的方式会产生不同影响,如何分割上述两种生态补偿财政职能的边界,将是我国财政生态补偿今后要解决的重要理论问题。

### (二) 应着力解决的实践问题

#### 1. 所有者身份代表主体问题。

财政的所有者代表职能在理论上是清晰的,即代表国家履行财政职能。在实践中,作为全民所有制的自然资源代表的相关财政职能究竟应由哪个部门具体代表却并不很清晰。按照国务院的三定方案,自然资源部履行全民所有土地、矿产、森林、草原、湿地、水、海洋等自然资源资产所有者职责和所有国土空间用途管制职责,承担自然资源专项收入征管和专项资金、基金的管理工作<sup>[32]</sup>。国有资本经营预算反映的是各级政府及其所属部门和单位依法利用国家资源提供的特定的公共服务征收、收取、提取、募集的除税收和政府债务收入外的财政收入,反映各级人民政府及其部门、机构履行出资人职责的企业(即一级企业)上缴的国有资本收益。根据国务院“三定”方案可以看出,自然资源部和财政部关于自然资源的所有者代表身份不够清晰,自然资源部作为国家出资人代表有权负责资源性资产企业红利上缴事宜,而财政

部作为预算管理部门则有权根据《预算法》赋予的权利从事预算管理工作。建议通过全国人大立法或者国务院常务会议的形式,明确财政部和自然资源部在涉及自然资源财政专项资金所有者管理方面的权责,以更好地实现生态补偿保护等职能。

## 2. 妥善解决生态补偿历史遗留问题。

妥善处理生态补偿历史遗留问题是财政作为公共管理者履行代际公平调节职能的重要表现,中央财政也对革命老区、民族地区、边疆地区、资源枯竭型城

市等特殊地区实施生态保护活动,以消化其生态环境和经济社会发展的历史欠账,促进地区协调发展。但是,由于上述地区长期向国家和社会无偿或低价提供了大量资源,为国家和经济社会发展作出了巨大贡献,而随着经济社会的发展以及生态文明建设的要求不断提高,这类地区面临转型问题,不少地区资源型城市转型任务十分繁重,建议在我国每年的政府预算中都安排一定规模的特定地区生态补偿规模,并通过绩效评估等方式提升预算绩效水平。

## 参考文献

- [1] Wunder S. Payment for Environmental Services: Some Nuts and Bolts [R]. Bogor: CIFOR, 2005: 3-4.
- [2] Ezzine-de-Blas D, Wunder S, Ruiz-Pérez M, Moreno-Sanchez R P. Global Patterns in the Implementation of Payments for Environmental Services [J]. Plos One, 2016, 11 (3): 2.
- [3] 张晏. 国外生态补偿机制设计中的关键要素及启示 [J]. 中国人口·资源与环境, 2016 (10): 121-129.
- [4] 张诚谦. 论可更新资源的有偿利用 [J]. 农业现代化研究, 1987 (5): 45-51.
- [5] 刘桂环等. 近20年我国生态补偿研究进展与实践模式 [J]. 中国环境管理, 2021 (5): 109-118.
- [6] 邓子基. 财政只能是经济基础的范畴 [J]. 中国经济问题, 1962 (11): 12-22.
- [7] 许廷星. 财政与银行的关系及其在社会主义建设中的作用 [J]. 财经科学, 1982 (3): 34-39, 9.
- [8] 许毅. “国家分配论”的产生和发展 [J]. 财政研究, 1995 (6): 34-39.
- [9] 贾康. 财政的本质及在新时期的职能作用 [J]. 财经问题研究, 1994 (10): 33-38.
- [10] 郭庆旺, 赵志耘. “公共财政论”的再质疑 [J]. 财政研究, 1999 (12): 55-58.
- [11] 邓子基. 借鉴“公共财政论”发展“国家分配论” [J]. 财政研究, 2000 (1): 51-55.
- [12] 楼继伟. 社会主义市场经济条件下的财政职能 [J]. 农村财政与财务, 2006 (10): 4-6.
- [13] 张馨, 杨志勇, 郝联峰, 袁东. 当代财政与财政学主流 [M]. 大连: 东北财经大学出版社, 2000: 358-363.
- [14] 高培勇. 公共经济学 [M]. 北京: 中国人民大学出版社, 2012: 10-16.
- [15] 叶振鹏. 财政理论与实践 [M]. 北京: 中国财政经济出版社, 1998: 17-21.
- [16] 高培勇. 新时代中国财税体制改革的理论逻辑 [J]. 财政研究, 2018 (11): 11-16.
- [17] 高培勇. “基础和支柱说”: 演化脉络与前行态势——兼论现代财税体制的理论源流 [J]. 财贸经济, 2021 (4): 5-19.
- [18] 陈诗一, 祁毓. “双碳”目标约束下应对气候变化的长期财政政策研究 [J]. 中国工业经济, 2022 (5): 25-43.
- [19] 刘尚希. 防范公共风险 刘尚希谈财政改革四十年的逻辑 [J]. 财经界, 2018 (28): 54-55.
- [20] 孙亦军. 财政与资源性资产管理的关系问题研究 [J]. 中央财经大学学报, 2006 (4): 7-10.
- [21] 钟大能. 在西部民族地区完善财政生态补偿机制的对策建议 [J]. 中央财经大学学报, 2006 (5): 22-26.
- [22] 徐丽媛. 生态补偿财税责任立法的国际经验论析 [J]. 山东社会科学, 2017 (3): 168-176.
- [23] 郑云辰, 葛颜祥, 接玉梅, 张化楠. 流域多元化生态补偿分析框架: 补偿主体视角 [J]. 中国人口·资源与环境, 2019 (7): 131-139.
- [24] 马克思. 哥达纲领批判 [M]. 中共中央马克思恩格斯列宁斯大林著作编译局, 译. 北京: 人民出版社, 2012: 12-17.
- [25] 孙毅, 张如石. 补偿经济论 [M]. 北京: 中国财经出版社, 1990: 68-70.
- [26] 黄宇驰, 姚明秀, 王卿, 苏敬华, 王敏. 生态产品价值实现的理论研究与实践进展 [J]. 中国环境管理, 2022 (3): 48-53.
- [27] 李宏伟, 薄凡, 崔莉. 生态产品价值实现机制的理论创新与实践探索 [J]. 治理研究, 2020 (4): 34-42.
- [28] 詹疏璐, 杨建州. 生态产品价值及实现路径的经济学思考 [J]. 经济问题, 2022 (7): 19-26.
- [29] Costanza R. Ecosystem Services: Multiple Classification Systems Are Needed [J]. Biological Conservation, 2018, 141 (2): 350-352.
- [30] Kemkes R J, Farley J, Koliba C J. Determining When Payments Are an Effective Policy Approach to Ecosystem Service Provision [J]. Ecological Economics, 2010, 69 (11): 2069-2074.
- [31] 高晓龙, 张英魁, 马东春, 徐卫华, 郑华, 欧阳志云. 生态产品价值实现关键问题解决路径研究 [J]. 生态学报, 2022 (20): 7-8.
- [32] 自然资源部. 自然资源部职能配置、内设机构和人员编制规定 [EB/OL]. (2018-08-12) [2022-03-20]. [https://www.mnr.gov.cn/jg/sdfa/201809/t20180912\\_2188298.html](https://www.mnr.gov.cn/jg/sdfa/201809/t20180912_2188298.html).

(责任编辑: 孙亦军 张安平)

# 保险公司视角的万能险产品结算利率定价实证研究

## An Empirical Study on Settlement Rate Pricing of Universal Insurance

### Products from the Perspective of Insurance Companies

许 荣 李延胤 方明浩 张俊岩

XU Rong LI Yan-yin FANG Ming-hao ZHANG Jun-yan

**[摘要]** 万能险产品结算利率是否过高是当前保险政策监管与市场关注的焦点。已有研究大多是从理论层面探讨万能险产品利率制定,实证研究较为缺乏。本文基于我国56家寿险公司2014—2018年的相关数据,从寿险公司经营模式、资金需求和风险承担能力等多个保险公司层面特征实证研究影响我国寿险公司万能险产品结算利率定价的影响因素。研究结果表明:(1) 保险公司经营模式与万能险产品结算利率的定价显著相关,偏向资产驱动型经营模式的寿险公司对万能险产品结算利率的定价更高。(2) 寿险公司万能险产品结算利率与其在金融市场资金需求程度正相关。(3) 风险承担能力强的寿险公司偏向于发行高结算利率的万能险产品。此外,保险公司的退保率、审计质量和员工学历结构等其他公司特征也可能对万能险的结算利率产生影响。本研究有助于认识万能险结算利率制定与潜在的金融机构资金需求以及风险管理之间的逻辑关联,从实证研究的视角帮助监管层和实务界更好地认识到万能险的保险属性有可能受到保险公司金融市场融资需求的影响甚至扭曲。

**[关键词]** 万能险 结算利率 资金需求 风险管理

**[中图分类号]** F842 **[文献标识码]** A **[文章编号]** 1000-1549 (2022) 11-0026-13

**Abstract:** Whether the settlement rate of universal insurance products is too high is the focus of current insurance policy supervision and market attention. Most of the existing researches discuss the interest rate formulation of universal insurance products from the theoretical level, but empirical research is relatively lacking. Based on the relevant data of 56 life insurance companies in China from 2014 to 2018, this paper empirically studies the factors influencing the settlement rate pricing of universal insurance products of life insurance companies in China from the perspectives of life insurance companies' business model, capital demand and risk bearing capacity. The results show that: (1) the business model of insurance companies is significantly correlated with the pricing of universal insurance product settlement interest rate, and life insurance companies that prefer asset-driven business model have higher pricing of universal insurance product settlement interest rate. (2) The settlement rate of universal insurance products of life insurance companies is positively correlated with its capital demand in the financial market. (3) Life insurance companies with strong risk bearing capacity tend to issue universal insurance products with high settlement interest rate. In addition, other corporate characteristics of insurance companies, such as surrender rate, audit quality and staff education structure, also affect the settlement rate of universal insurance. This study helps to understand the logical relationship between the setting of universal insurance settlement rate and the potential capital demand and risk management of financial institutions, and helps regulators and practitioners to better understand that the insurance property of universal insurance may be affected or even distorted by the financing demand of insurance companies in the financial market from the perspective of empirical research.

**Key words:** Universal life insurance Settlement rate Financing demand Risk management

**[收稿日期]** 2022-04-11

**[作者简介]** 许荣,男,1975年12月生,中国人民大学财政金融学院暨财政金融政策研究中心、保险研究所教授,博士生导师,研究方向为公司金融实证、保险需求研究;李延胤,男,1997年8月生,中国人民大学财政金融学院博士生,研究方向为金融风险管理;方明浩,男,1991年11月生,中国人民大学经济学院博士后,研究方向为金融与保险;张俊岩,男,1975年6月生,中国人民大学财政金融学院副教授,研究方向为风险管理与保险。本文通讯作者为张俊岩,联系方式为 zhangjunyan@ruc.edu.cn。

**[基金项目]** 中国人民大学中央高校基本科研业务费专项资金(项目编号:18XN1006)。

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

万能险也被称为“可调整保费的终身寿险”，是终身寿险的一个变种，在1979年由美国加利福尼亚人寿保险公司首次推出。由于其相对于传统寿险具有缴费灵活、保障灵活、运作透明、保底收益等优点，一经推出就受到了市场的青睐。为了振兴国内萧条的寿险市场，2000年中国太保推出了中国第一款万能险。我国万能险产品曾一度成为寿险市场中重要的组成部分，在寿险市场发展历程中发挥了重要作用（郭金龙和周华林，2016<sup>[1]</sup>；魏丽和杨建奎，2017<sup>[2]</sup>；周县华等，2019<sup>[3]</sup>）。2015—2016年间，万能险出现了短期爆发式发展，万能险保费收入同比增长高达95.23%和55.10%，万能险保费收入占比分别为31.62%和34.39%。

保险公司保险产品尤其是投资型保险产品费率制订的基本原则是投资收益率超过其给予客户的结算利率，也即存在利差收益。正常情况下保险产品不应该出现费率倒挂的“贴钱”状况。然而，根据新闻报道和银保监会的信息发布，2020年8月，根据非现场监测情况，银保监会发现部分人身险公司万能险账户的财务收益率低于实际结算利率，存在利差损等风险，因此对12家人身险公司进行约谈，责令此12家人身险公司立即整改，根据万能账户实际投资情况科学合理确定实际结算利率，并于2020年9月1日前将整改情况进行书面报告<sup>①</sup>。

在经济和保险理论研究中，保险产品费率确定的基本原则和任何产品定价的基本原理并无本质差别，在自由竞争的市场假设下，保险产品费率由该保险产品的市场供给和需求达到均衡下供求共同决定均衡价格。但是正如经济现实产品价格除了受到供求影响外还会受到多种因素影响，保险产品费率同样有可能受到多种因素影响甚至明显偏离均衡价格，这些影响因素包括由于信息摩擦导致的价格扭曲（Rothschild 和 Stiglitz，1978<sup>[4]</sup>）、由于产品市场摩擦导致的价格竞争或者金融市场摩擦如金融危机或监管限制造成的价格扭曲（Kojien 和 Yogo，2015<sup>[5]</sup>）。本文通过研究中国万能险产品市场费率制定的影响因素，尝试探索保险产品供给者行为也即保险公司行为尤其是其保险资金

运用的特征对万能险产品费率造成何种影响。

万能险的结算利率制定之所以备受市场监管和消费者关注，主要因为万能险兼具保障和投资双重属性。万能险结算利率主要包括最低保证利率与结算利率，按照万能险账户的相关规定，结算利率主要应当由万能险账户的实际投资收益率决定。同时，结算利率不得低于最低保证利率。由图1可知，2012年之前万能险结算利率<sup>②</sup>与十年期国债利率和银行存款利率的差距并不明显。而在2012年之后，特别是在2015—2016年间，多家险企一度推出结算利率超过7%的万能险产品，市场万能险平均结算利率接近5%，远高于同期银行存款利率和十年期国债利率，有相当一部分险企利用高万能险结算利率作为噱头营销，吸引了大量的投保人资金。但已有研究表明，万能险产品承诺的高利率背后潜藏着巨大的风险隐患（展凯和申曙光，2009<sup>[6]</sup>；郭金龙和周华林，2016<sup>[1]</sup>）。过高的万能险结算利率会直接增加承保成本，并通常会推高险企的投资风险，从而危害保险公司偿付能力。20世纪80年代以来美国、日本的寿险公司也曾推出过高利率理财型保险产品，但随着经济环境恶化，不仅许多保单出现了增加保费、减少收益或承保期限等不利于承保人的情况，而且寿险公司资产负债状况也出现了严重失衡，爆发经营危机。因此，研究影响我国保险公司对于万能险结算利率价格制定的关键因素，有可能为防范万能险在中国发展过程中潜在风险，并推动我国万能险产品未来的高质量发展发挥重要作用。

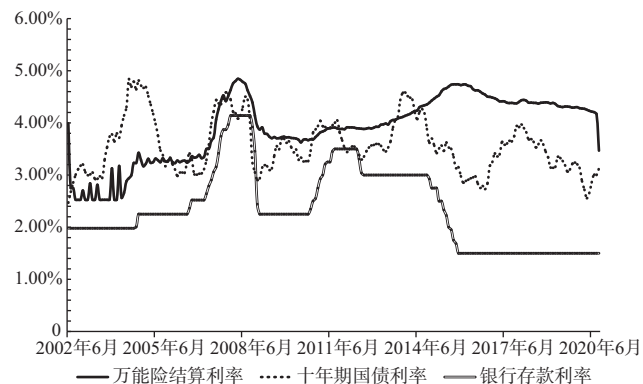


图1 万能险结算利率与固收市场利率的对比走势图

数据来源：万得数据库（Wind）。

① 新华网客户端“银保监会再出重拳整治万能险 超5%收益率或成历史”，<https://baijiahao.baidu.com/s?id=1674864831521966579&wfr=spider&for=pc>。

② 图中的万能险结算利率为同一时期市场上公布的所有万能险产品结算率的平均值。

本文手工收集 2014—2018 年间 56 家中国寿险公司层面的相关数据, 实证研究寿险公司的经营模式、资金需求和风险承担能力等公司特征指标对万能险结算利率定价的影响。研究发现: (1) 寿险公司的经营模式与万能险结算利率显著相关, 偏向资产驱动型经营模式的寿险公司会倾向设定较高的万能险结算利率。(2) 金融市场资金需求更高的寿险公司往往会设定较高利率的万能险结算利率, 这有可能是因为资金需求驱动寿险公司发售高利率万能险以提高保单发售量, 获取更多现金流。(3) 风险承担能力强的公司具有对万能险产品结算利率的定价优势。进一步研究发现寿险公司的其他特征如审计质量、员工学历和退保率也会对寿险公司的万能险结算利率定价产生影响。

本文可能的学术贡献主要体现在: 第一, 已有研究(展凯和申曙光, 2009<sup>[6]</sup>; 周桦, 2013<sup>[7]</sup>; 薛华等, 2015<sup>[8]</sup>; 陈秉正和何宇佳, 2016<sup>[9]</sup>)更多是从理论模型层面推算万能险结算利率的制定过程和影响因素, 并且研究多是基于金融市场和消费者选择权的视角出发。本文研究则从寿险公司自身的经营模式、资金需求、风险承担能力出发, 利用手工收集数据, 实证研究万能险定价的影响因素。据作者有限的知识, 这一研究是目前较少的提供了中国寿险公司万能险结算利率制定的实证证据的研究之一。第二, 本研究有可能帮助更好地认识万能险结算利率制定与潜在的金融机构资金需求以及风险管理之间的逻辑关联。已有研究发现金融机构面临债务压力和资金短缺时将被迫以低价出售资产(Shleifer 和 Vishny, 1992<sup>[10]</sup>)或者采取各种方式进行融资, 但这些研究更多关注银行, 例如银行所面临的财务约束压力(Hellwig, 2009<sup>[11]</sup>)以及存款人提款压力(Calimani 等, 2017<sup>[12]</sup>)等。目前只有较少研究关注保险公司, 例如 Ellul 等(2011)<sup>[13]</sup>研究发现财务约束较为严重也即资金需求特别旺盛的保险公司有可能折价出售保险公司债券。本研究揭示了寿险公司在二级市场上市公司举牌行为有可能显著影响其万能险结算利率制定, 从而为金融机构资金需求影响金融产品定价行为提供了实证证据。第三, 本文研究具有一定的监管参考意义和实务参考价值。如果保险公司因为二级市场举牌等紧急融资需求而变相提高万能险结算利率, 虽然有可能缓解

其融资急需, 但是却在一定程度上歪曲了万能险的长期保障属性, 致使万能险产品变相发挥了短期理财产品的投融资属性, 不仅不能发挥保险的“兜底”保障功能, 反而进一步造成巨大的金融风险 and 潜在的社会不稳定。相关新闻报道已经发现, 在险资举牌较为频繁的 2014 年到 2016 年间, 寿险公司的举牌过程大多有万能险账户的参与。例如, 从 2014 到 2016 年公开披露资金来源的 84 次寿险公司举牌中, 有 68 次举牌的资金有万能险账户的参与。其中, 在宝能系举牌万科的第一阶段, 主要通过前海人寿的 104.22 亿元保险资金大量购买万科股票, 其中就包括万能险账户保费资金 79.6 亿元和传统保费资金 24.62 亿元, 因此前海人寿的万能险发挥了非常重要的融资作用<sup>①</sup>。本文研究有助于从实证研究的视角帮助监管层和实务界更好地认识到万能险的保险属性有可能受到保险公司金融市场融资需求的影响甚至扭曲。如果保险公司将万能险扭曲为理财产品, 在市场上追求短期高收益, 甚至将其变成实现对上市公司控制的工具, 则有可能与万能险的本质特性背道而驰。因此, 本文讨论保险公司经营模式和资金需求等因素与万能险结算利率制定之间的关系, 能够帮助消费者和监管者更好地识别万能险产品的风险, 帮助监管层和保险公司更好地制定万能险结算利率, 从而真正实现“保险姓保”、回归保险保障功能, 实现更好地保障保险消费者、防范化解保险行业系统性风险的目标(许荣等, 2019<sup>[14]</sup>; 方明浩等, 2021<sup>[15]</sup>)。

本文其余部分安排如下: 第二部分是政策背景与研究假设; 第三部分说明了变量选取与研究设计; 第四部分为实证分析结果; 第五部分是稳健性检验; 第六部分为研究总结和建议启示。

## 二、政策背景与研究假设

### (一) 政策背景

万能险的发展与监管政策和宏观环境密切相关, 主要经历了以下几个阶段: 1999 年 6 月起, 传统寿险预定利率上限被监管部门管控为 2.5%, 这给了万能险市场较大的发展空间。同时自 2006 年起, 国内宏观资本市场向好, 助推万能险取得了较快速的发展, 此时万能险结算利率出现稳步上涨。

① 新华社“审视万科事件: 起底宝能系资金链”, [http://www.xinhuanet.com/mrdx/2016-07/08/c\\_135497775.htm](http://www.xinhuanet.com/mrdx/2016-07/08/c_135497775.htm)。

2008年全球金融危机发生,宏观资本市场的震荡极大地影响了万能险的发展。政策方面,为了引导寿险业回归保障本质,保监会从2008年8月开始大力推动寿险业结构调整。随后,保监会出台了一系列政策组合拳,加大对投资型业务的监督检查力度;加强对非寿险投资型产品的监管;通过强化产品审查力度,推动各公司进一步调整和优化产品结构。2011年,保险业开始实施新的会计准则,万能险保费收入大部分在年报中不能计入保费收入,这使得不少保险公司纷纷调整产品策略,扩大分红险、传统险比重,收缩万能险占比。这一阶段万能险出现负增长,同时万能险结算利率处于相对低迷状态。

2012年以后,在持续低利率的大环境下,我国保险投资监管环境逐渐走向宽松。保监会陆续发布保险资金投资新政,进一步放开不动产和股权的投资行业和领域,允许投资银行理财、信托等金融产品、创业板、优先股等。保险费率改革和权益类投资占比的放宽促使我国万能险市场出现了迅猛的发展,万能险结算利率也出现了持续上涨的趋势,并在2016年左右达到了最高值。但在此期间,寿险公司频繁举牌上市公司,干扰公司治理的乱象背后也存在着万能险的支持。

2017年以来,为规范万能险业务发展,落实“保险姓保”的政策理念,银保监会出台了有关规范中短存续期产品、强化人身保险产品监管的多项规定,对万能险的规模、账户管理、保障水平、结算利率等问题进行了规范和监管,并在行业内开展了万能险专项检查。政策的收紧使得各公司纷纷削减自身万能险业务,此时利率出现稳步下降的趋势。

由上文可知,万能险结算利率受监管政策和宏观经济环境影响较大。首先,在政策的作用下,保险公司的投资职能得到了较为充分的发挥,增加了对权益类资产的配置(边文龙和王向楠,2017<sup>[16]</sup>)。而高利率万能险吸引的现金流恰好可以满足因投资权益类资产而产生的资金需求。其次,万能险账户资产管理配置会受到宏观经济环境影响(张绍白,2014<sup>[17]</sup>)进而影响其结算利率。同时低利率的固收市场环境和不稳定的国内股票市场使得更多投资者资金涌入保险市场(魏丽和杨建奎,2017<sup>[2]</sup>),保险公司为了竞争获取这些资金进而采取增大万能险业务占比,提高结算利率的行为。

## (二) 研究假设

### 1. 经营模式对万能险结算利率定价的影响。

2012年及之后保险资金监管政策的放松导致了资产负债管理视角下经营模式的分化(仲赛末和赵桂芹,2018<sup>[18]</sup>),出现了一批非传统的资产驱动型经营模式的寿险公司。传统寿险公司大多为负债驱动型经营模式,这类公司以负债端的承保业务来驱动资产端的投资业务发展,其主要盈利模式在于先吸收负债端的保费,再在资产端进行稳健的财务投资;而资产驱动型经营模式的公司是先在投资端找到高收益项目,再在承保端通过发售保单实现保费快速增长,以达到提高经营收益和扩大资产负债规模的目的。当然这种蕴藏着巨大的市场风险的经营模式也受到了学术界的质疑与担忧(孙立娟和费清,2017<sup>[19]</sup>;秋慧,2019<sup>[20]</sup>)。资产驱动型寿险公司存在一定程度的经营风险,仲赛末和赵桂芹(2018)<sup>[18]</sup>研究发现资产驱动型公司的财务稳健度显著弱于传统经营模式的公司,风险提高而绩效没有显著提升。不仅如此,经营模式对寿险公司流动性创造也会产生影响,但与规模相关。张诗豪和赵桂芹(2020)<sup>[21]</sup>研究发现偏向资产驱动型经营模式的大中型寿险公司,流动性吸收功能较差,而小型寿险公司表现相反。

资产驱动型的寿险公司以投资作为主要驱动力,根据投资业务特点设计相匹配的寿险产品。万能险产品相较于传统寿险产品具有收费透明,有保底收益,投资收益率高等特质,对消费者的吸引力更大。因此资产驱动型的寿险公司大多选择万能险产品为其主要的经营产品以实现投资项目资金的更快完成,满足资金需求压力。以曾经的安邦系为例,2012到2016年其主要以万能险投资型保费为主的“保户储金及投资款”占总保费的比例达70%~95%以上,并利用银行渠道销售快速渗透做大规模。先后举牌或进入前十大股东名单的上市公司达38家,市值达2506亿。因通过万能险账户筹集资金完成了多次蓝筹股的举牌,使得其收益远高于同行业,进而吸引负债端获得更多资金。

高结算利率的万能险更能吸引并满足多数消费者的投资回报需求,在其他因素相同的情况下,销量明显好于低结算利率的万能险产品,有助于帮助寿险公司抢占市场。资产驱动型寿险公司倾向发行高结算利率万能险的主要原因有两点。第一,资产驱动型寿险

公司投资端确定了投资项目后需要短期筹集大量资金,高结算利率万能险产品的高现金流量特征满足了资产驱动型寿险公司的资金流动性需求。第二,2013年后一批新兴险企选择资产驱动型经营模式,这些企业为了扩大自身影响力,实现保险业务规模的迅速扩张,打破传统寿险公司的垄断,往往会选择销售相对结算利率更高的万能险。如前海人寿成立初期就以万能险业务为其“杀手锏”,保费收入连续两年保持100%增幅,仅成立四年就跻身于寿险公司的第11位。

因此本文提出第一个研究假设:

**假设1:**偏向资产驱动型经营模式的寿险公司会倾向提高万能险的结算利率。

## 2. 资金需求对万能险结算利率定价的影响。

韩忠雪和周婷婷(2011)<sup>[22]</sup>研究发现财务约束大即资金需求高的企业缓解融资压力的动机更强。除此之外,企业还可能通过调整其出售的资产和负债定价达到缓解其融资压力的目的。当企业特别是金融机构面临债务压力和资金短缺时,这些机构将被迫以低价出售资产。这种情况在严重依赖短期资本市场的银行表现尤为明显(Shleifer和Vishny,1992<sup>[10]</sup>)。当然相同情况也会发生在保险公司。Ellul等(2011)<sup>[13]</sup>研究发现财务约束较为严重也即资金需求特别旺盛的保险公司有可能折价出售保险公司债券。Nanda等(2019)<sup>[23]</sup>研究发现监管约束会引发保险公司低价抛售债券,当其价格大幅跌破基本价值时,可能会对市场上的其他投资者造成损害。Kojien和Yogo(2015)<sup>[5]</sup>研究了2008年金融危机时期的保险公司为了在供不应求的资金市场筹措资金以解救燃眉之急,因而降低费率获取资金的策略。在2008年11月至2009年2月之间,美国保险公司以16%的折价出售30年期定期年金,以19%折价出售人寿年金,并对万能寿险的折价达到惊人的57%,这表明当时保险公司给予的预定利率远远超出市场该类保险产品的均衡利率。而根据Mitchell等(1999)<sup>[24]</sup>的研究,市场正常状况下的同类保险产品价格往往会有6%到10%的价格加成。

寿险公司通过发行万能险能够很大程度缓解保险公司的融资压力,减轻其流动性资金需求的紧迫程度。通过万能险融资的最大特点是能够较为快速地筹集大量资金满足公司的需求,但万能险的特征决定了结算利率越高的万能险产品越会受到消费者市场的追

捧,相应获得的保费收入越多。

因此本文提出第二个研究假设:

**假设2:**在金融市场具有较紧迫资金需求的寿险公司会倾向于提高万能险的结算利率。

3. 公司风险承担能力对万能险结算利率定价的影响。

公司风险承担能力是指公司对风险的缓冲与覆盖能力(Shane,2000<sup>[25]</sup>),反映了公司的客观能力。公司规模和资产负债率等客观因素决定了公司风险承受能力(许家云等,2015<sup>[26]</sup>;宋光辉等,2016<sup>[27]</sup>),而它们又作为重要的财务信息,向寿险公司万能险结算利率制定者传递了重要的信号。在发行万能险的过程中,利率制定者通过年报或内部披露的相关指标可以判断公司的财务状况和风险承担能力。当公司风险承受能力较弱时,承担高利率万能险保单成本的压力较大,只能发行较低利率万能险产品;反之,公司风险承受能力越强,万能险结算利率制定者会认为公司承担保单发行成本压力越小,从而发行高利率万能险产品抢占市场。

因此本文提出第三个研究假设:

**假设3:**公司风险承受能力强的寿险公司会倾向提高万能险结算利率定价。

## 三、变量选取与研究设计

### (一) 样本数据

由于寿险资金举牌主要集中在2014年及之后,本文最终选取了2014—2018年56家中国寿险公司的非平衡面板数据作为研究样本,其中包括21家外资寿险公司和35家中资寿险公司。寿险公司保费数据来自中国银保监会网站,寿险公司举牌数据来自举牌公司和被举牌公司官方网站的公告信息,经手工整理得到。寿险公司审计数据来自公司年报,经手工整理得到。其他数据来自中国保险年鉴、万得数据库(Wind)和国泰安数据库(CSMAR)。

### (二) 变量定义

#### 1. 被解释变量。

本文把寿险公司万能险产品月度利率进行平均得到年度利率,将各寿险公司的年度万能险结算利率作为被解释变量。

#### 2. 解释变量。

(1) 寿险公司经营模式(*Business\_model*)。寿险

公司万能险投资性保费占比<sup>①</sup>反映了寿险公司的经营策略和产品结构。资产驱动型经营模式的寿险公司对投资型产品的重视程度更高，这些寿险公司会更积极寻求高收益投资项目，开展投资活动。诸多研究通过用万能险投资性保费占比水平表征寿险公司的经营模式（仲赛末和赵桂芹，2018<sup>[18]</sup>；张诗豪和赵桂芹，2020<sup>[21]</sup>），万能险投资性保费占比越高的寿险公司越偏向资产驱动型经营模式。但寿险公司当年发行的万能险数量可能和价格存在着相关性，为了避免这种相关关系可能造成的内生性问题，本文选取万能险投资性保费占比滞后一期变量（第  $t-1$  年）与第  $t$  年的万能险结算利率变量进行回归。

(2) 寿险公司在金融市场的资金需求。本文选择的指标是寿险公司举牌次数 (*Tender\_offer*) 和举牌所需资金 (*Tender\_offer\_money*)，具体的定义为一个自然年内寿险公司发生的举牌次数和总举牌所需资金量与公司规模之比。举牌<sup>②</sup>是触发举牌收购这一行为的简称。2013年后，我国寿险公司频发举牌，保险资金大规模进军国内股票市场，引发了市场的动荡。而在包括收购举牌在内的投资中企业资金流动性发挥着关键作用（Yang等，2019<sup>[28]</sup>），一般而言，寿险公司举牌次数越多，举牌所需资金越大，在金融市场的资金需求越旺盛。

(3) 公司的风险承担能力。首先是公司资本规模 (*Size*)，使用总资产的对数表示。一般来说，资本规模较大的保险公司承担万能险保单发行成本的能力较强。本文认为，资本规模越大的寿险公司会发行越高利率的万能险产品以增强市场竞争力。其次，寿险公司的资产负债率 (*Lev*) 代表了风险指标，度量寿险公司的风险。一般而言，资产负债率越低的保险公司面对高利率万能险带来的资金流动性的压力越小。本文预期公司资产负债率对万能险结算利率定价的影响为负。

本文还选择了可能影响寿险公司万能险定价的其他公司特征变量作为控制变量。一是将是否接受四大会计师事务所 (*Big4*) 和公司学历构成 (*Edu*) 作为评判公司治理情况的重要指标。我们认为，高质量的审计和高学历的员工构成能有效抑制寿险公司定价风险，相应制定的万能险结算利率会较低。二是寿险公司退保率 (*Surrender*) 上升将增大寿险公司的流动性风险（孙祁祥等，2015<sup>[29]</sup>），导致公司财务状况的紧张。我们认为退保率高的寿险公司会倾向于提高万能险的利率以缓解公司财务压力。除此之外，本文还选取公司经营年限 (*Age*) 以及当年 GDP 增速 (*GDP\_growth*) 和无风险利率 (*Interest\_Rate*) 作为可能影响万能险结算利率定价的影响因素。

变量定义见表 1。

表 1 变量定义

变量类型	变量名	符号	定义
被解释变量	万能险结算利率	<i>Rate</i>	寿险公司年平均万能险结算利率
解释变量	经营模式	<i>Business_model</i>	万能险收入投资性部分占总保费比重衡量的滞后一期数据
	举牌次数	<i>Tender_offer</i>	$\log(1 + \text{寿险公司每年的举牌次数})$
	举牌所需资金	<i>Tender_offer_money</i>	$\log(1 + \text{寿险公司每年因举牌需要花费的资金/总资产})$
	资产负债率	<i>Lev</i>	负债与资产比
	资本规模	<i>Size</i>	$\log(\text{资产})$

① 在 2013 年监管机构规定的寿险保费收入的统计口径中，保户投资款新增交费代表了万能险保费收入中投资性质而非保障性质的部分，本文用“保户投资款新增交费”占规模保费（原寿险保费收入、保户投资款新增交费和投连险独立账户新增交费之和）的比例定义万能险投资性保费占比。

② 为了保护中小投资者利益，防止机构大户操纵股价，《证券法》规定：投资者持有或者通过协议、其他安排与他人共同持有有一个上市公司已发行股份达到 5% 或者是 5% 的整倍数时，应向国务院证券监督管理机构、证券交易所作出书面报告，通知该上市公司，并予以公告。

续前表

变量类型	变量名	符号	定义
控制变量	负责审计的公司	<i>Big4</i>	四大会计师事务所审计为1, 非四大会计师事务所审计为0
	学历结构	<i>Edu</i>	公司硕士学历以上员工占比
	退保率	<i>Surrender</i>	本年退保支出/(年初准备金+本年保费收入)
	公司年龄	<i>Age</i>	$\log(1+\text{保险公司成立年数})$
	GDP 增速	<i>GDP_growth</i>	—
	无风险利率	<i>Interest_Rate</i>	—

### (三) 计量模型

考虑到面板数据的结构, 本文计量模型如下:

$$Rate_{i,t} = \beta_0 + \beta_1 \times Business\_model_{i,t-1} + \beta_2 \times Demand_{i,t} + \beta_3 \times Risk\_bearing_{i,t} + \gamma \times Control_{i,t} + \alpha_i + \theta_t + \varepsilon_{i,t} \quad (1)$$

其中,  $Rate_{i,t}$  表示保险公司的万能险结算利率水平,  $Business\_model_{i,t-1}$  表示保险公司经营模式,  $Demand_{i,t}$  表示保险公司的资金需求情况,  $Risk\_bearing_{i,t}$  表示保险公司的风险承担能力,  $Control_{i,t}$  为控制变量, 在估计中控制保险公司的个体固定效应  $\alpha_i$  和年度固定效应

$\theta_t$ ,  $\varepsilon_{i,t}$  为残差项。

### (四) 描述性统计

表2报告了主要变量的描述性统计结果。可以看出, 在样本区间内发行的万能险结算利率均值(中值)为4.47%(4.46%), 标准差为0.59%。万能险投资性保费占比的均值为31.30%, 但最大值为99.10%, 表明在样本期间一些寿险公司的保险业务中万能险投资部分占比非常庞大。寿险公司的举牌次数(*Tender\_offer*)的均值为0.13, 而其标准差为0.44, 举牌资金占比的标准差也达到了0.47。这说明寿险公司在样本期间在资本市场上的表现差异较大。

表2 变量描述性统计

变量名	观察值	均值	标准差	最小值	中值	最大值
<i>Rate</i> (%)	266	4.47	0.59	2.84	4.46	6.01
<i>Business_model</i> (%)	261	31.30	29.10	0.12	21.90	99.10
<i>Tender_offer</i>	266	0.13	0.44	0.00	0.00	2.30
<i>Tender_offer_money</i>	266	0.12	0.47	0.00	0.00	2.64
<i>Surrender</i> (%)	266	7.78	8.90	0.09	4.58	43.20
<i>Lev</i> (%)	266	87.50	7.01	56.10	88.90	98.30
<i>Size</i>	266	10.68	0.70	9.21	10.58	12.41
<i>Big4</i>	266	0.71	0.46	0.00	1.00	1.00
<i>Edu</i> (%)	266	7.60	10.90	0.22	3.65	57.10
<i>Age</i>	266	2.45	0.46	0.69	2.57	3.47
<i>GDP_growth</i> (%)	266	7.00	0.23	6.75	6.95	7.43
<i>Interest_Rate</i> (%)	266	3.54	0.42	2.89	3.59	4.18

注: 为了消除异常值的影响, 对于被解释变量和解释变量, 采取了上下1%的 winsorize 处理。

本文借鉴张诗豪和赵桂芹(2020)<sup>[21]</sup>的方法,以寿险公司资产规模<sup>①</sup>为标准,将全部寿险公司分为大型寿险公司、中型寿险公司以及小型寿险公司。从图2可以看出,从2014年至2018年,整体来看,寿险公司的万能险结算利率先增后减。但大中型寿险公司的万能险结算利率始终高于小型寿险公司的万能险结算利率。

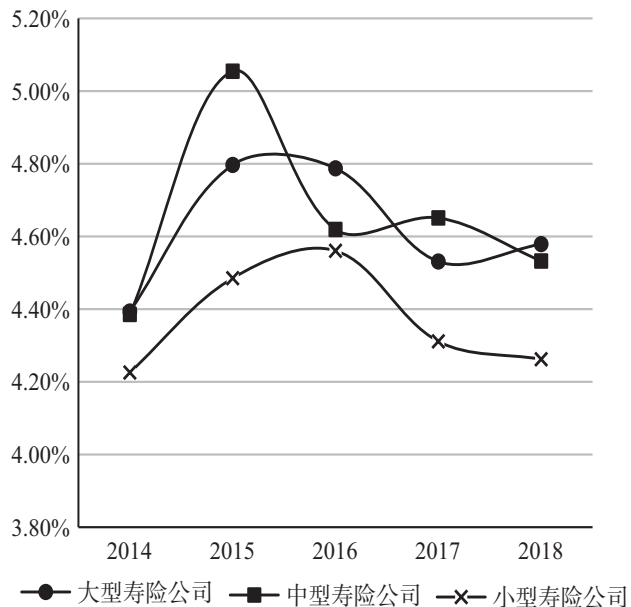


图2 不同规模寿险公司万能险结算利率在2014—2018年的变动趋势

数据来源:万得数据库(Wind)。

#### 四、实证结果分析

本文以寿险公司的万能险平均利率作为被解释变量,以经营模式、资金需求和风险承担能力等公司特征变量为解释变量。在控制其他因素的条件下,对56家寿险公司整体进行回归分析,结果汇总成表3。

从表3可以看出,寿险公司资产驱动经营模式倾向程度与万能险结算利率均呈显著正相关。即保险公司的经营策略越接近资产驱动型,其产品结构越倚重投资型产品业务,万能险结算利率定价越高。本文还发现代表资金需求的寿险公司举牌次数和举牌所需资金量与万能险结算利率存在显著正相关关系,这表明寿险公司举牌次数越多,举牌所需资金量越大,公司面临的资金需求越大,万能险结算利率定价越高。

从回归结果还可以看出,反映公司风险承担能力的重要信号越正向,即资产规模越高,资产负债率越低,寿险公司越有能力承担高利率带来的成本压力,万能险结算利率定价越高。对于其他变量,本文发现:寿险公司退保率越高,资金流动性越差,万能险结算利率定价越高。接受四大会计师事务所审计,硕士以上从业人员占比越高的寿险公司万能险结算利率定价越低,表明其对风险把控越严格。成立时间越短的寿险公司,万能险结算利率定价越高,表明新兴寿险公司希望通过高利率万能险产品吸引消费者的心理更强烈。

表3 万能险定价影响因素的回归分析结果

变量名	Rate		
<i>Business_model</i>	0.305 *** (2.70)		
<i>Tender_offer</i>		0.120 ** (2.34)	
<i>Tender_offer_money</i>			0.133 *** (2.63)
<i>Surrender</i>	0.966 *** (2.97)	1.057 *** (3.21)	1.072 *** (3.26)
<i>Size</i>	0.433 *** (2.80)	0.432 *** (2.75)	0.430 *** (2.75)
<i>Lev</i>	-1.020 ** (-2.03)	-1.139 ** (-2.27)	-0.988 * (-1.95)

① 本文将总资产排行前15%的寿险公司定义为大型寿险公司;将总资产排名在后60%的寿险公司定义为小型寿险公司;其余归类为中等规模寿险公司。

续前表

变量名	Rate		
	<i>Edu</i>	-0.159 (-0.51)	-0.126 (-0.40)
<i>Big4</i>	-0.485 ** (-2.23)	-0.477 ** (-2.16)	-0.469 ** (-2.13)
<i>Age</i>	-0.573 ** (-2.13)	-0.645 ** (-2.38)	-0.628 ** (-2.32)
<i>GDP_growth</i>	-0.230 (-0.98)	-0.080 (-0.34)	-0.061 (-0.26)
<i>Interest_Rate</i>	-0.094 (-0.41)	-0.263 (-1.16)	-0.274 (-1.21)
常数项	4.214 * (1.84)	4.153 * (1.79)	3.899 * (1.69)
年度固定效应	控制	控制	控制
公司固定效应	控制	控制	控制
观测值	261	266	266
<i>R-squared</i>	0.358	0.340	0.345

注：\*\*\*、\*\*、\*分别表示估计系数通过1%、5%、10%的显著性检验，下同。

根据图2可知，大中型寿险公司与小规模寿险公司之间的万能险结算利率存在较大差异。为了探究不同规模寿险公司影响因素的区别，本文分别对其中的

22家大中型寿险公司和34家小规模寿险公司进行分样本回归分析<sup>①</sup>。结果汇总成表4。

表4 不同规模下万能险定价影响因素的回归分析结果

变量名	Rate		
	全样本	小规模	大中型
<i>Business_model</i>	0.305 *** (2.70)	0.259 (1.47)	0.571 *** (3.50)
<i>Surrender</i>	0.966 *** (2.97)	1.150 ** (2.48)	0.313 (0.78)
<i>Size</i>	0.433 *** (2.80)	0.735 *** (3.20)	-0.061 (-0.17)
<i>Lev</i>	-1.020 ** (-2.03)	-1.191 * (-1.80)	0.422 (0.36)
<i>Edu</i>	-0.159 (-0.51)	-0.045 (-0.12)	-0.331 (-0.55)
<i>Big4</i>	-0.485 ** (-2.23)	-0.489 * (-1.78)	—

① 因为小规模寿险公司举牌行为次数很少，故未将举牌行为作为解释变量纳入分样本回归中。



续前表

变量名	Rate		
	全样本	小规模	大中规模
Age	-0.573 ** (-2.13)	-0.738 ** (-2.00)	0.092 (0.22)
GDP_growth	-0.230 (-0.98)	-0.224 (-0.70)	0.011 (0.04)
Interest_Rate	-0.094 (-0.41)	-0.048 (-0.15)	-0.198 (-0.66)
常数项	4.214 * (1.84)	1.501 (0.47)	5.110 (1.30)
年度固定效应	控制	控制	控制
公司固定效应	控制	控制	控制
观测值	261	168	93
R-squared	0.358	0.337	0.600

本文发现在分样本回归中，大中规模寿险公司资产驱动经营模式倾向程度与利率呈显著正相关且系数较大，而小规模公司资产驱动经营模式倾向程度与利率不显著相关且系数较小。这表明在大中型寿险公司中，经营策略对万能险定价影响更明显，而在小型寿险公司中，其影响程度较小。其次，小规模寿险公司风险承担能力即规模和资产负债率与利率的显著性明显高于在大中型寿险公司中的显著性。这说明，小规模寿险公司更重视风险承担能力对结算利率定价的影响。

### 五、稳健性检验

为确保实证结果具备良好的稳健性，本文采用两种方法分别对模型进行稳健性检验：替换主要解释变量后回归和进一步缩尾处理。

#### (一) 替换主要解释变量后回归

如前所述，偏好资产驱动经营模式和在金融市场资金需求旺盛的寿险公司更偏好承保高利率、高现金流的万能险。资产驱动型经营模式不止体现在保险公司业务结构方面，还表现在保险公司利润来源结构方面，即资产驱动型经营模式更注重投资获益和投资端业务发展（尚颖等，2022<sup>[30]</sup>）。因此，本文选用投资

收益占比的滞后一期变量（Invest）即投资收益与营业收入的比重和投资资产占比的滞后一期变量（High risk）<sup>①</sup>表示当期的寿险经营模式（尚颖等，2022<sup>[30]</sup>；仲赛末等，2021<sup>[31]</sup>）。寿险公司投资收益占比和投资资产占比越高，其经营过程中的投资端业务发展越强，经营模式越偏向资产驱动。对于寿险公司资金需求，本文采用了另一个指标——寿险公司作为上市重要股东在二级市场的增持资金量（Overweight），定义为增持金额与公司资产之比的对数。寿险公司作为上市重要股东在二级市场的增持资金量数据来自 wind 数据库。寿险公司在二级市场增持所需要的资金量越大，其资金需求越旺盛。

对比表 3 和表 5，可以发现前文的实证结论较为稳健。即资产驱动型经营模式和金融市场上资金需求越旺盛的公司，万能险结算利率定价越高。

#### (二) 缩尾处理

由于部分解释变量存在极端值，为了进一步验证前文结论的稳健性，本文在对所有变量进行 1% 分位双边缩尾的基础上再进行 3% 分位的双边缩尾处理。对所有变量按 3% 分位双边缩尾处理后得到表 6 的回归结果，主要回归结果得出的结论与前文保持一致。

① (货币资金+交易性金融资产+可供出售金融资产+持有至到期资产+买入返售金融资产+保户质押贷款)/公司总资产

表 5 替换指标后万能险定价影响因素的回归分析结果

变量	Rate				
	全样本	小规模	大中规模	全样本	全样本
<i>Invest</i>	3.158*** (2.95)	2.800 (1.48)	3.777*** (2.82)		
<i>High risk</i>				0.009* (1.80)	
<i>Overweight</i>					0.036** (2.46)
<i>Surrender</i>	0.959*** (2.96)	1.121** (2.43)	0.317 (0.76)	0.928*** (2.83)	1.060*** (3.22)
<i>Size</i>	0.390** (2.52)	0.748*** (3.27)	-0.162 (-0.43)	0.416*** (2.62)	0.423*** (2.69)
<i>Lev</i>	-1.258** (-2.51)	-1.297* (-1.94)	-0.524 (-0.42)	-1.097** (-2.16)	-1.045** (-2.07)
<i>Edu</i>	-0.118 (-0.38)	-0.025 (-0.07)	0.053 (0.08)	-0.166 (-0.52)	-0.145 (-0.46)
<i>Big4</i>	-0.437** (-2.02)	-0.483* (-1.75)	—	-0.502** (-2.26)	-0.463** (-2.10)
<i>Age</i>	-0.588** (-2.20)	-0.811** (-2.20)	-0.080 (-0.19)	-0.662** (-2.43)	-0.589** (-2.16)
<i>GDP_growth</i>	-0.006 (-0.03)	-0.091 (-0.29)	0.300 (0.94)	-0.045 (-0.19)	-0.039 (-0.17)
<i>Interest_Rate</i>	-0.344 (-1.53)	-0.249 (-0.80)	-0.553* (-1.85)	-0.314 (-1.37)	-0.272 (-1.20)
常数项	4.261* (1.86)	1.462 (0.45)	6.869* (1.74)	4.319* (1.86)	-2.355 (-1.03)
年度固定效应	控制	控制	控制	控制	控制
公司固定效应	控制	控制	控制	控制	控制
观测值	261	168	93	266	266
<i>R-squared</i>	0.362	0.337	0.574	0.333	0.362

表 6 再缩尾处理后万能险定价影响因素的回归分析结果

变量名	Rate			
<i>Business_model</i>	0.275** (2.53)			
<i>Tender_offer</i>			0.128** (2.25)	
<i>Tender_offer_money</i>				0.140** (2.48)
<i>Surrender</i>	1.084*** (3.12)		1.164*** (3.33)	1.186*** (3.40)
<i>Size</i>	0.350** (2.37)		0.356** (2.39)	0.349** (2.35)

续前表

变量名	Rate		
<i>Lev</i>	-1.142* (-1.93)	-1.329** (-2.22)	-1.199** (-1.99)
<i>Edu</i>	-0.141 (-0.43)	-0.099 (-0.30)	-0.092 (-0.28)
<i>Big4</i>	-0.510** (-2.44)	-0.506** (-2.38)	-0.501** (-2.36)
<i>Age</i>	-0.504* (-1.95)	-0.571** (-2.19)	-0.557** (-2.14)
<i>GDPgrowth</i>	-0.205 (-0.91)	-0.062 (-0.28)	-0.047 (-0.21)
<i>InterestRate</i>	-0.135 (-0.61)	-0.289 (-1.33)	-0.303 (-1.40)
常数项	5.042** (2.33)	4.941** (2.27)	4.821** (2.21)
年度固定效应	控制	控制	控制
公司固定效应	控制	控制	控制
观测值	261	266	266
<i>R-squared</i>	0.349	0.333	0.337

## 六、总结与建议

万能险产品作为曾经火爆一时的投资型险种，其结算利率的影响因素实证研究尚处于起步阶段。本文尝试从公司层面的角度出发，研究经营模式，资金需求和风险承担能力等其他公司特征指标对万能险结算利率的影响。首先，本文发现寿险公司的资产负债视角下的经营模式对万能险结算利率的制定具有较为重要的影响。寿险公司的经营模式越偏向于资产驱动型，制定的万能险结算利率越高。其次，寿险公司的举牌次数越多，举牌资金量越大，资金需求越紧迫，往往倾向于发布高结算利率的万能险产品。同时，风险承担能力高的寿险公司具备万能险的定价优势。除此之外，审计质量、人员学历结构与退保率等其他公司特征也会对万能险结算利率的制定产生影响。本文的结论具有一定的政策启示。

第一，2012年以来监管政策的宽松促进了资产驱动型寿险公司的发展，而资产驱动型寿险公司会设定较高的万能险结算利率，加上万能险的营销渠道成本费，会给公司带来较大的负债压力。尽管该种经营

模式的寿险公司可能会通过举牌“蓝筹股”的行为获得短期可观利润，但在宏观经济环境突然恶化时，资产驱动型寿险公司很可能出现资产端收益难以弥补负债端成本的情况，继而出现经营危机甚至破产的风险。因此，监管层可以考虑适当限制寿险公司在资产驱动型经营模式上的发展，从而防止过高万能险结算利率带来的成本压力。

第二，寿险公司在金融市场的资金需求迫切程度会促使其设定更高的万能险结算利率。一方面，大多万能险账户资金参与了举牌，这意味着寿险公司投资端权益性资产配置的增加，而权益性资产的投资周期较长；另一方面，负债端较高利率的万能险产品一般周期较短且成本较高。这种“短钱长配”情况和较高负债资金投资成本会给寿险公司带来高现金流压力。

第三，新兴中资寿险公司可能会倾向于采取激进的万能险结算利率定价策略，以扩大自身影响力和规模。相关监管部门需要提高保险公司准入门槛，加强对新兴保险公司的监管。坚决杜绝新成立寿险公司通过高利率万能险“快进快出”套取大额资金，获得超额回报的情况出现。

第四, 提高对万能险结算利率恶性竞争的监管, 防范万能险产品销售误导情况的出现, 加强消费者对万能险产品认识, 促进我国万能险产品向着更健康的方向发展。

## 参考文献

- [1] 郭金龙, 周华林. 保险业的潜在系统性风险 [J]. 中国金融, 2016 (6): 80-82.
- [2] 魏丽, 杨建奎. 万能险的市场选择研究 [J]. 保险研究, 2017 (2): 3-10.
- [3] 周县华, 范庆泉, 覃家明. 我国万能险产品可持续发展的最优保证利率研究 [J]. 保险研究, 2019 (10): 84-100.
- [4] Rothschild M, Stiglitz J. Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information [M] // Uncertainty in Economics. Academic Press, 1978: 257-280.
- [5] Kojien R S J, Yogo M. The Cost of Financial Frictions for Life Insurers [J]. American Economic Review, 2015, 105 (1): 445-475.
- [6] 展凯, 申曙光. 万能险内在收益率测算与影响因素分析 [J]. 保险研究, 2009 (7): 90-97.
- [7] 周桦. 中国万能险投资账户最低收益率保证与退保期权的定价研究 [J]. 系统工程理论与实践, 2013 (3): 557-568.
- [8] 薛华, 郑海涛, 王钰琼. 金融市场突变对万能险定价及偿付能力的影响 [J]. 金融研究, 2015 (4): 176-191.
- [9] 陈秉正, 何宇佳. 基于双驱动型动态最优资产负债管理模型的万能险产品经营分析 [J]. 保险研究, 2016 (5): 13-23.
- [10] Shleifer A, Vishny R W. Liquidation Values and Debt Capacity: A Market Equilibrium Approach [J]. The Journal of Finance, 1992, 47 (4): 1343-1366.
- [11] Hellwig M F. Systemic Risk in the Financial Sector: An Analysis of the Subprime-Mortgage Financial Crisis [J]. De Economist, 2009, 157 (2): 129-207.
- [12] Calimani S, Halaj G, Zochowski D. Simulating Fire-sales in a Banking and Shadow Banking System [R]. ESRB Working Paper Series, 2017.
- [13] Ellul A, Jotikasthira C, Lundblad C T. Regulatory Pressure and Fire Sales in the Corporate Bond Market [J]. Journal of Financial Economics, 2011, 101 (3): 596-620.
- [14] 许荣, 方明浩, 常嘉路. 险资持股能够降低公司违规吗? ——保险投资的公司治理效应研究 [J]. 保险研究, 2019 (12): 48-64.
- [15] 方明浩, 许荣, 赵昶. 险资股权投资的结构改革研究——基于险资投资的倒U型效应分析 [J]. 经济理论与经济管理, 2021 (3): 40-57.
- [16] 边文龙, 王向楠. 投资职能对保险公司风险的影响研究 [J]. 金融研究, 2017 (12): 158-173.
- [17] 张绍白. 万能险中国十五年 [J]. 中国保险, 2014 (11): 44-46.
- [18] 仲赛末, 赵桂芹. 经营模式对寿险公司财务状况的影响——基于资产负债管理视角 [J]. 经济管理, 2018 (9): 155-172.
- [19] 孙立娟, 费清. 经济泡沫、保险自由化与改革: 日本保险业的发展历程 [J]. 现代日本经济, 2017 (2): 79-94.
- [20] 秋慧. 对近年我国寿险公司“资产驱动型”经营模式失败原因的反思 [J]. 经济研究导刊, 2019 (11): 158-161.
- [21] 张诗豪, 赵桂芹. 资本水平、经营模式与流动性创造——对我国寿险公司的实证检验 [J]. 保险研究, 2020 (5): 3-19.
- [22] 韩忠雪, 周婷婷. 产品市场竞争、融资约束与公司现金持有: 基于中国制造业上市公司的实证分析 [J]. 南开管理评论, 2011 (4): 149-160.
- [23] Nanda V, Wu W, Zhou X A. Investment Commonality across Insurance Companies: Fire Sale Risk and Corporate Yieldspreads [J]. Journal of Financial and Quantitative Analysis, 2019, 54 (6): 2543-2574.
- [24] Mitchell O S, Poterba J M, Warshawsky M J, Brown J R. New Evidence on the Money's Worth of Individual Annuities [J]. American Economic Review, 1999, 89 (5): 1299-1318.
- [25] Shane S. Prior Knowledge and the Discovery of Entrepreneurial Opportunities [J]. Organization Science, 2000, 11 (4): 448-469.
- [26] 许家云, 毛其淋, 佟家栋. 出口如何影响了企业的风险承担能力? [J]. 产业经济研究, 2015 (2): 1-14.
- [27] 宋光辉, 钱崇秀, 许林. 商业银行“三性”对其风险承担能力的影响——基于16家上市银行非平衡面板数据的实证检验 [J]. 经济管理, 2016, 38 (9): 135-148.
- [28] Yang J, Guariglia A, Guo J M. To What Extent Does Corporate Liquidity Affect M&A Decisions, Method of Payment and Performance? Evidence from China [J]. Journal of Corporate Finance, 2019, 54: 128-152.
- [29] 孙祁祥, 边文龙, 王向楠. 业务集中度对寿险公司利润和风险的作用研究 [J]. 当代经济科学, 2015 (3): 27-38, 125.
- [30] 尚颖, 贾士彬, 毕书琨. 保险业资产驱动型经营模式会助推系统性金融风险吗? [J]. 保险研究, 2022 (2): 17-32.
- [31] 仲赛末, 赵桂芹, 张诗豪. 寿险机构经营属性分化对金融市场稳定性的影响 [J]. 财经研究, 2021 (3): 125-139.

(责任编辑: 李 晨 张安平)

# 基于预训练模型的 A 股停牌预测研究

## The Research on Prediction of Trading Halts in

### A-Share Based on Pre-trained Model

孙夫雄 谢翔 熊平 梁嘉欣 彭畅

SUN Fu-Xiong XIE Xiang XIONG Ping LIANG Jia-Xin PENG Chang

**[摘要]** 股票停牌是基于提高信息披露程度、抑制股票剧烈波动、消除信息不对称等目的所采取的机制,但个股的停牌行为存在随意性以及其复牌时间不确定性的问题,对投资者的合法权益造成了较大的损害。本文研究从股票波动中挖掘停牌的内在规律,提出了基于预训练模型的股票停牌预测机制。首先预训练模型学习 A 股上市公司股票停牌的共性特征,然后通过预训练模型参数的迁移学习获得个股停牌的特征,进而构建特定个股的停牌预测模型。实证分析选取 2 539 家上市公司作为研究对象,对比分析了多个深度学习组合模型,选择以 TADM (TCN-Attention-Dense Model) 网络结构作为预训练模型具有较好的总体预测效果,在此基础上个股的迁移学习模型表现出更强的样本识别能力且误报率更低。研究发现,虽然诱发个股停牌的因素很多,但仍能从股票波动中预测停牌事件。研究成果为投资者规避停牌风险、减少投资损失或利用个股停牌实现收益最大化等提供借鉴,并可为监管机构对停牌趋势做出前瞻性判断,对宏观调控和政策调整提供参考。

**[关键词]** 股票停牌 预训练模型 迁移学习 深度学习 非平衡样本

**[中图分类号]** F832.48 F276.6 TP181 **[文献标识码]** A **[文章编号]** 1000-1549 (2022) 11-0039-13

**Abstract:** Trading halts in stock market is a mechanism designed to improve the degree of information disclosure, restrain stock volatility and eliminate information asymmetry. However, uncertainty associated with the suspension and resumption of stocks can cause great trouble and loss to investors. The paper studies mechanism of trading halts from perspective of stock fluctuations, and proposes a prediction pre-trained model of trading halts. First, the pre-trained model is used to learn the common features of A-share listed companies' trading halts. These common features are further employed to get features of individual stock with the technique of transfer learning, then the prediction model of specific stocks is constructed. In the experiment, 2 539 listed companies are selected as the research objects. Comparative analysis of multiple deep learning combination models, the pre-trained model adopts the TADM (TCN-Attention-Dense Model), which has a better overall prediction effect. Based on TADM, the transfer learning model of individual stocks holds stronger sample recognition ability and lower false positive rate. It is found that although there are many factors inducing trading halts of individual stock, trading halts can still be predicted from the stock fluctuations, which provides reference for investors to avoid the risk of trading halts and reduce investment losses, or to maximize returns using trading halts, and for regulators to make forward-looking judgments on the trend of trading halts, as well as macro-control and policy adjustment.

**Key words:** Trading halt Pre-trained model Transfer learning Deep learning Unbalanced sample

**[收稿日期]** 2022-06-14

**[作者简介]** 孙夫雄,男,1971年3月生,中南财经政法大学信息与安全工程学院副教授,研究方向为大数据分析、人工智能和信息安全;谢翔,男,1996年7月生,中南财经政法大学信息与安全工程学院硕士研究生,研究方向为金融科技;熊平,男,1974年10月生,中南财经政法大学信息与安全工程学院教授,研究方向为网络安全、数据挖掘和隐私保护;梁嘉欣,女,1997年8月生,中南财经政法大学信息与安全工程学院硕士研究生,研究方向为数据挖掘;彭畅,男,2000年11月生,中南财经政法大学信息与安全工程学院硕士研究生,研究方向为数据挖掘。本文通讯作者为熊平,联系方式为 pingxiang@zuel.edu.cn。

**[基金项目]** 高等学校学科创新引智基地项目(项目编号: B21038);中南财经政法大学研究生拔尖人才培养项目(项目编号: XKRH202101, CXJH202105)。

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

停牌即“停止证券上市”，指的是某一种证券临时停止交易的行为。股票停牌触发后，市场交易将被强制中止，使交易各方有时间获取最新信息重新评估交易价格、调整交易策略，以提升信息披露程度、降低交易信息的不对称性，削弱上市公司风险事项的不确定性影响、维护市场秩序和保护投资者合法权益。但A股停牌的频次高、时间长与随意性等问题侵犯了投资者的交易权，影响了市场效率。部分上市公司利用筹划重大事项等名义停牌，进行股价操纵与内幕交易以攫取非法利益，严重背离了股票停牌制度设立的初衷，降低了价格发现效率。

在2015年“6·15股灾”期间，7月9日这天停牌的上市公司数量有1438家之多，占有上市公司数量比例逾50%。投资者的资金大面积套牢，盈利股票无法兑现浮盈，大量抛压难以释放，整个交易市场被悲观恐慌情绪所笼罩。另外我国A股还存在停牌时间超长且强制复牌程序不配套的问题，例如\*ST新亿（600145）停牌时间从2015年到2020年长达54个月，成为中国证券历史上的最长停牌时间记录，而深深房A（000029）停牌自2016年至2020年时长为51个月。长期停牌极大挫伤了投资者的投资意愿和信心，给交易市场流动性带来严重损害，也为监管层带来了更大的工作挑战，不利于股票市场的长期稳定运行。

中国股市存在的停牌滥用、超长停牌等问题已经引起了学术界、监管层和投资者的重视。国内外学者们从多个角度入手开展了股票停牌相关研究，包括国内外停牌制度对比、停牌时长影响因素、停牌前后股价波动的分析等已发生的确定性事件，但对于股票停牌预测等不确定事件研究尚存在空白。

因此，本文采用预训练模型和迁移学习的深度学习学习方法用于股票停牌的预测，主要贡献如下：一是首次从停牌预测的视角，提出了基于预训练模型的股票停牌预测，首先预训练模型用于学习A股上市公司股票停牌的共性特征，再通过迁移学习机制获得个股停牌的特征，进而实现精确的个股停牌预测，为相关研究提供了新的研究方法 with 实证证据。二是研究成果在一定程度上解决了交易各方的信息不对称问题，使我国中小投资者能有效地规避停牌风险，进而减少投资损失以及实现资产保值，或为部分投资者利用个股

停牌后的股价波动，实现收益最大化等目标提供帮助。三是通过对A股市场不同行业板块的停牌预测分析，使监管机构能够对市场整体停牌趋势做出前瞻性判断，对宏观调控和政策调整提供了参考。

## 二、文献综述

金融数据波动的分析方法主要涵盖了统计学、机器学习以及深度学习，中外学者在该领域皆取得了卓越的研究成果。

在统计学方法研究方面，Doroudyan等（2017）<sup>[1]</sup>研究ARMA-GARCH模型对德黑兰证券交易所价格指数的时间序列进行拟合，实现了金融过程的监控和预警。蔡光辉等（2021）<sup>[2]</sup>提出了基于偏 $t$ 分布的混频Realized GARCH模型，进而推导其参数估计方法，并通过MCS检验判别扩展模型与滚动时间窗预测技术对中国黄金期货市场波动进行预测研究，实验结果显示相较于GARCH模型与Realized GARCH模型，结合高频信息的Realized GARCH模型在样本内估计的预测精度和拟合优度表现更好；在机器学习研究方面，Gupta等（2021）<sup>[3]</sup>利用随机森林技术预测了投资者信心对美国股票市场波动的影响。Vara等（2022）<sup>[4]</sup>比较了卡尔曼滤波器、XGBoost和ARIMA三种主要算法对股票价格的预测，提出了卡尔曼滤波器和XGBoost的混合模型，研究结果显示其预测精度优于单一模型。张剑和王波（2017）<sup>[5]</sup>采用SVM（Support Vector Machine）理论搭建沪深300股指期货量化交易模型，并根据动态预测模型设计了量化交易策略，实验显示量化交易模型整体可以获得较高的投资回报；在深度学习研究方面，Chatzis等（2018）<sup>[6]</sup>比较分析了分类树、支持向量机、随机森林、深度神经网络等技术对股市危机事件预测性能，指出深度神经网络具有显著的分类精度。Althelaya等（2021）<sup>[7]</sup>利用小波变换技术捕捉不同尺度的金融时间序列波动，结合深度学习模型提高了股票预测的精度。Sun等（2021）<sup>[8]</sup>基于时间卷积网络（Temporal Convolutional Network, TCN）对股民的情感信息进行建模，用于对股票和期货价格进行预测，实验结果表明TCN在MAE、MAPE、RMSE和精度方面优于GRU模型。

对于股票停牌的研究，国内外学者基本上侧重于停牌的制度性研究，以及停牌对股价影响的实证研究。Kabir和Engelen（2006）<sup>[9]</sup>评估监管机构是否能够成功利用停牌机制，迫使公司向资本市场披露新的

重大信息, 研究结果显示了停牌机制在传播新信息方面的有效性。Taechapiroontong 等 (2012)<sup>[10]</sup> 考察了亚洲新兴股市停牌期间停牌的有效性以及不同类型投资者或交易员的交易表现。Rahim 等 (2021)<sup>[11]</sup> 以印尼证券交易所为研究对象, 分析投资者对停牌的反应。国内学者也对我国股市进行了卓有成效的实证研究, 如在取消异常波动停牌的政策背景下分别对中小板股价波动、股市流动性与价格发现效率等方面进行了相关研究 (陈舒宁等, 2016<sup>[12]</sup>; 胡婷等, 2017<sup>[13]</sup>; 李洋等, 2018<sup>[14]</sup>)。石阳等 (2019)<sup>[15]</sup> 以重大事项停牌数据为实验对象, 实证检验了随意停牌对投资者利益的影响, 指出了上市公司的部分停牌行为并不能符合投资者保护的初衷, 反而有助于内部私利交易行为, 对投资者利益造成了损害。李明琨等 (2022)<sup>[16]</sup> 从进化博弈分析的角度研究我国熔断机制的有效性, 以及政府监管作用。He 等 (2019)<sup>[17]</sup> 研究了强制和自愿停牌对中国股市股价、波动性和成交量的影响。

综上所述, 深度学习方法在金融时间序列预测上具有较好的运用效果。但针对股票停牌的研究多反映在制度有效性、现有缺陷和制度改进方向, 对股票停牌预测研究尤为匮乏, 因此本文研究了基于预训练模型的股票停牌预测方法。

### 三、相关理论基础

#### (一) 停牌理论分析

市场中如果出现对股票交易有特殊影响的重要信息, 股票停牌能够为投资者主动划割出可以获取充分信息反馈的时间区间, 对股价再预估, 进而对之前与当前股价预期存在差异的投资策略进行调整, 以削弱交易各方的信息不对称, 股票停牌制度制定初衷也在于此。包括股市发生异常波动与企业并购活动在内的多种因素均可能诱发股票停牌。

从停牌原因来看, 停牌发生主要是归结于以下三类事件: (一) 上市公司披露期中业绩报告、年报, 召集股东大会, 重要收购合并, 扩股增资, 公布分红计划, 投资以及股权增减变动等需公布重大信息时; (二) 上市公司就关于对公司产生重大影响的事项被证券监管机关认为应当澄清与公告时; (三) 需就上市公司涉嫌违规开展调查时。

这三类事件的共同核心在于上市公司当前经营中出现了新的重要信息, 为消除不同投资者获取信息的时间差异, 保护市场交易的公平性与有效性,

上市公司需要暂停其股票交易。依据有效市场假说 (Efficient Markets Hypothesis, EHM) 理论 (Fama, 1970<sup>[18]</sup>), 在作为有效市场的股票交易市场中, 全部有价值的信息已经准确、及时并充分地反映在股票价格波动上。因此, 由上市公司经营过程引发的停牌事件虽然直接原因是新信息发生与传播, 但是这种信息与公司的历史经营信息不是割裂的、独立的, 应当是由公司管理层或者监管机构根据历史信息进行决策或者在历史信息中读取到了隐藏信息, 公司经营发生的新信息与历史信息是存在紧密联系的。

上市公司经营信息的直接体现是股票每日行情, 其中包括了股票的交易价格、成交量以及换手率等信息, 衍生出最高价、最低价、开盘价、收盘价、成交量、成交额、股价涨跌幅等交易基础指标。根据基本交易信息, 可以计算股票交易技术分析指标, 这类指标通过综合考量多个基本交易信息, 对价格的波动较为敏感, 在实际交易分析中被广泛应用, 其常见分类如下。

#### 1. 超买超卖型指标。

超买, 是指在买入需求过旺推动下某股票或市场股价陡升, 股价同实质价值发生严重偏离, 制造大量股价泡沫。在历经一轮急升后, 一旦高位承接力较弱, 股票价格出现技术性下跌是有可能的。

超卖, 指受大面积沽售影响而导致某股票或市场股价发生急挫, 股价显著低于实质价值, 但是经过一轮急跌变动, 大幅减少了低位抛售压力, 价格技术性向上反弹可能会出现。

对市场上这种价格和实质价值不匹配的暂时性现象的度量即超买超卖指标, 具体指标包括随机指标 KDJ、威廉指标 WR、动态买卖气指标 ADTM、相对强弱指标 RSI 和布林极限 BB 等。

#### 2. 趋势型指标。

市场人气盛衰与趋势走向是通过趋势型指标进行反映, 均线或者股价通道是判断趋势的方法。通常观点是趋势一经形成则将在特定时段内延续, 这就是利用对趋势的研判推断股价变动趋势的基础。趋势型指标计算过程纳入了股价、成交量以及涨跌指数等数据, 指标具体包括趋向指标 DMI、市场趋势 CYE、平滑异同平均线 MACD、威廉变异离散量 WVAD、筹码引力 PAV 和量价曲线 VPT 等。

#### 3. 能量型指标。

人气的热度、价格动量的潜能能够利用能量型指

标进行测量，而且股票价格的支撑带与压力带也可以被其显示出来，具有为预测分析股价后续走向趋势，研判买卖股票的时机节点提供参考借鉴的作用。一般能量型指标包括：成交量变异率 VR、带状能量线 CR、威廉多空力度线 WAD、市场强弱 CYR 和心理线 PSY 等。

#### 4. 均线型指标。

均线指标又称移动平均线指标，其主要作用是反映价格运行趋势，尤其是其可以根据投资者需求对不同时间尺度的均线进行计算，能够直观展现长期、中期和短期的趋势，如 K 线图。常见均线指标有：MA 均线、顾比均线 GMMA、多均线 BBI、多布林线 BBI-BOLL 和鳄鱼线 ALLIGAT 等。

其他股票交易技术分析指标还包括成交量型、图表型和五彩 K 线等。

#### (二) 相关深度学习理论

本文研究基于的深度学习理论包括时间卷积神经网络、注意力机制、迁移学习和变分自编码器等。

##### 1. 时间卷积神经网络。

Bai 等 (2018)<sup>[19]</sup> 基于卷积神经网络模型提出了

时间卷积网络 (Temporal Convolutional Network, TCN) 模型，针对序列问题采用因果卷积加以改进，模型记忆历史信息过程是通过残差模型进行堆叠卷积层实现的。

单向结构的因果卷积 (causal convolution) 在时间约束上十分严格，表现为神经元的输出仅利用卷积运算与其有因果联系的数据产生，以致限制了时间序列建模的长度。膨胀卷积 (dilated convolution) 允许卷积时的输入存在间隔采样，使得包含长距离依赖关系的较大范围的信息被每个卷积的输出所容纳，网络输出则能提取较大范围的时间序列特征。采样间隔  $d$  (通常以 2 为底的指数形式递增) 称为膨胀因子 (dilation rate)，设 1 维序列输入  $x \in \mathbb{R}^n$  和滤波器  $f: \{0, \dots, k-1\} \rightarrow \mathbb{R}^n$ ，则在  $x_t$  处的膨胀卷积运算  $F$  定义如下：

$$F(x_t) = (x *_{d} f)(x_t) = \sum_{i=0}^{k-1} f(i) \cdot x_{t-d \cdot i} \quad (1)$$

其中： $k$  表示卷积核的大小。假设  $d=1, 2, 4$  且  $k=3$ ，则膨胀因果卷积示意图如图 1 所示。

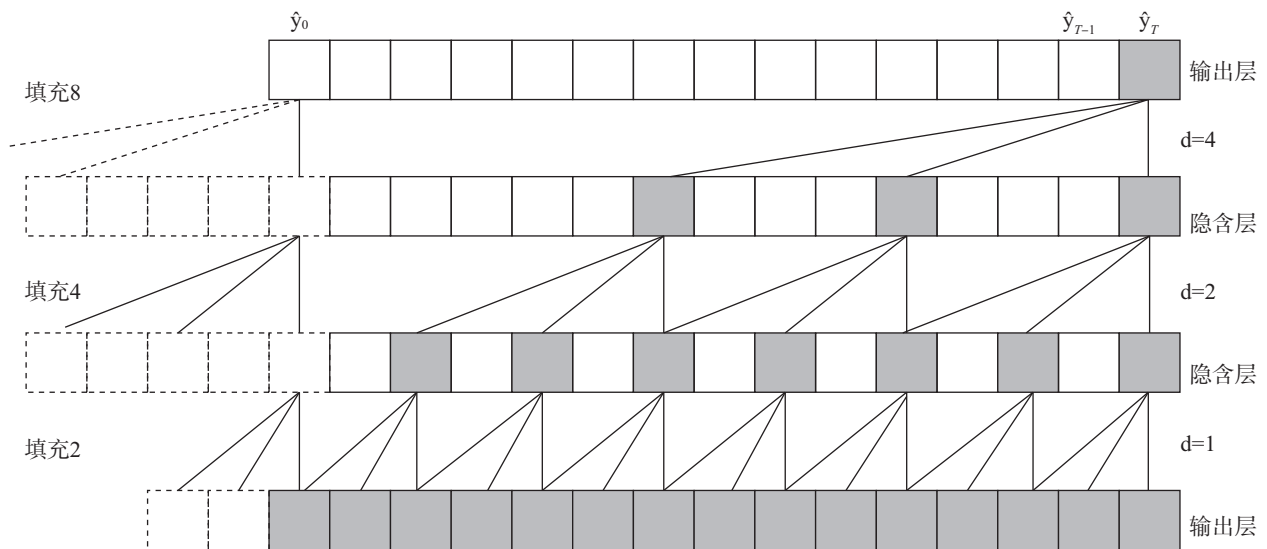


图 1 TCN 示意图

膨胀卷积和因果卷积在 TCN 网络结构的卷积层中进行了融合，由卷积核大小、网络深度以及膨胀系数共同决定了其网络感受野。考虑到对网络深度进一步加深以及提升对时间序列信息的学习能力，原本的卷积层由 TCN 残差模块取代且结构相同，皆包含了膨胀因果卷积层、权重归一化层、ReLU 激活函数、Dropout 层。如果输入与输出的数据维度不一致，则

通过  $1 \times 1$  卷积进行处理。

##### 2. 注意力机制。

注意力 (Attention) 机制来源于对生物观察行为的模仿 (Mnih 等, 2014<sup>[20]</sup>)，即在看到整体信息的时候，会主动发现重要信息并分配更多的注意力到目标部分，以获取更多与之相关的信息。注意力机制能够在稀疏数据中提取重要的信息。自注意机制 (self-



attention) 是注意力机制的一种改进方案 (刘建伟等, 2021<sup>[21]</sup>), 能够对全局的不同特征分配不同的权重, 用于分析数据内部的相关性。Attention 计算示意图如图 2 所示 (Raffel 和 Ellis, 2015<sup>[22]</sup>)。

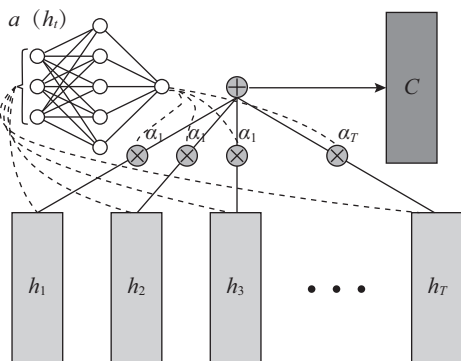


图 2 Attention 计算示意图

图 2 中 Attention 计算过程为: 首先将隐含层的输出向量 \$[h\_1, \dots, h\_T]\$ 组成的矩阵 \$H \in R^{T \times d}\$ 作为注意力层输入, 其中 \$T\$ 表示时间序列长度, \$d\$ 代表输出向

量的维度。然后利用多层全连接层进行非线性转换并通过 Softmax 函数完成归一化, 得到注意力权重矩阵 \$A=[\alpha\_1, \dots, \alpha\_T]\$。最后将输入矩阵乘以对应的注意力权重矩阵, 加和获得向量 \$c\$。对应公式定义如下:

$$A = \text{softmax}(V_\alpha \tanh(W_\alpha H^T)) \tag{2}$$

$$c = AH \tag{3}$$

其中: \$V\_\alpha\$ 和 \$W\_\alpha\$ 是权重系数矩阵, 可以定义 \$V\_\alpha = W\_\alpha^T\$ 进行简化。

### 3. 迁移学习。

将已经发展成熟的知识向近 (相) 似领域迁移的研究方法称作迁移学习 (Kumar 等, 2020<sup>[23]</sup>; Gu 和 Dai, 2021<sup>[24]</sup>)。一般将已有知识表述为源域 (Source domain, \$D\_s\$), 而对近 (相) 似领域又称为目标域 (Target domain, \$D\_t\$), 迁移学习的具体过程就是在相似领域应用其从一个或多个源域中提炼的关联知识处理目标域的新问题, 迁移学习基本原理如图 3 所示。

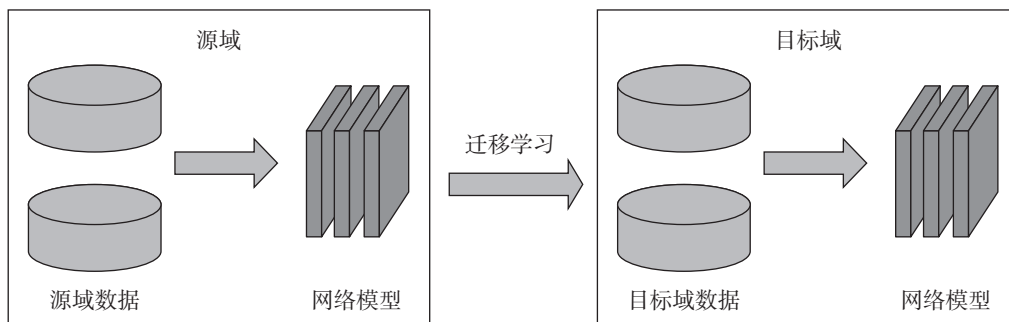


图 3 迁移学习示意图

源域 \$D\_s\$ 一般包含了特征空间 \$x\$ 与边缘分布概率 \$P(X)\$, 其中 \$X=(x\_1, \dots, x\_n)\$ 代表源域数据样本, 预测函数 \$f=P(Y|X)\$ 和标签空间 \$Y=(y\_1, \dots, y\_n)\$ 共同组成任务 \$O\$, 倘若预测函数不知, 仍可通过多组源域与目标域训练数据学习而获得。假定存在源域 \$D\_s\$、目标域 \$D\_t\$ 和有差别的学习任务 \$O\_s\$、\$O\_t\$, 迁移学习目的是通过 \$D\_s\$ 与 \$D\_t\$ 提升归属于任务 \$O\_t\$ 的预测函数 \$f\$ 在目标域 \$D\_t\$ 中的预测表现。

### 4. 变分自编码器。

变分自编码器 (Variational auto-encoder, VAE) 是属于深度生成模型的一种生成式网络结构 (Kingma 和 Welling, 2013<sup>[25]</sup>; Utyamishv 和 Partin-Vaisband, 2019<sup>[26]</sup>; Zhou 等, 2021<sup>[27]</sup>; Liu 等, 2022<sup>[28]</sup>), 其基本结构如图 4 所示, 主要分为编码器 (Encoder) 与解码器 (Decoder) 两部分。

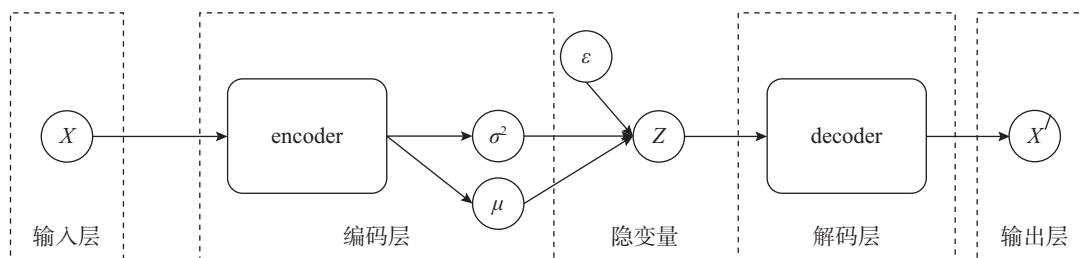


图 4 VAE 模型结构图

图4中数据样本  $x$  经过输入层到达 Encoder, 由 Encoder 生成均值  $\mu$  和方差  $\sigma^2$ , 从正态分布  $N(0, 1)$  中随机采样得到  $\varepsilon$ , 由  $\mu$ 、 $\sigma^2$  和  $\varepsilon$  共同确定隐变量  $z = \mu_i + \varepsilon \cdot \sigma^2$ , 并作为 Decoder 的输入, 由 Decoder 对隐变量  $Z$  进行重构, 输出样本  $x'$ 。定义损失函数  $L$  如下:

$$L = l_1 + l_2 \quad (4)$$

其中  $l_1 = \sum_i (x_i - x'_i)^2$  为均方误差,  $l_2 = -\frac{1}{2} \times \sum_i (2\log\sigma_i + 1 - \sigma_i^2 - \mu_i^2)$  为 KL 散度。

#### 四、研究方法及模型构建

##### (一) 模型总体设计

股票停牌预测模型的研究过程主要包括: (1) 选取部分交易基本指标和技术分析指标, 建立停牌预测指标体系; (2) 数据预处理后, 通过滑动窗口形成时序数据集; (3) 训练 VAE 模型, 生成停牌的模拟时序数据, 以平衡数据集; (4) 预训练模型训练与测试; (5) 迁移学习模型参数, 训练与测试个股停牌预测模型。模型研究框架如图5所示。

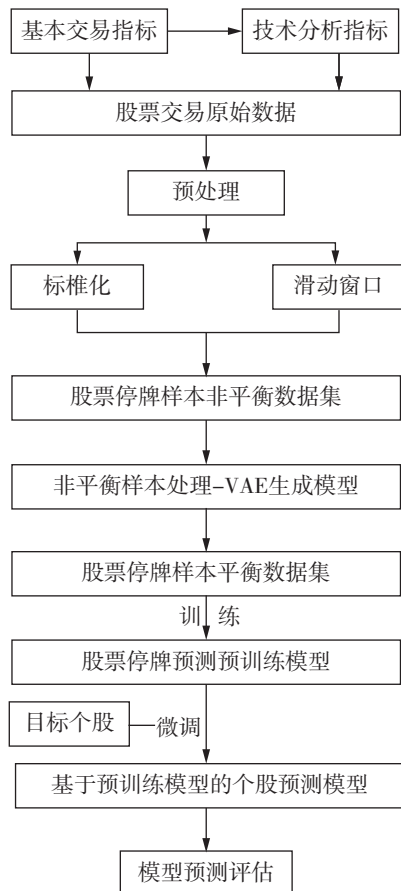


图5 模型研究框架图

##### (二) 模型指标体系

图5中股票交易原始数据包括基本交易指标和技术分析指标, 其中基本交易指标是股票市场交易过程中最直接和原始的记录, 而技术分析指标是利用交易价格、成交量的变动等数据计算得到的能够反映交易规律的综合信息(谢丁, 2009<sup>[29]</sup>)。在综合性、可解释性和客观性原则指导下, 选择了22个指标如表1所示。

表1 模型指标体系

一级指标	二级指标
基本交易指标	开盘价 (Open)
	最高价 (High)
	最低价 (Low)
	收盘价 (Close)
	涨跌额 (Change)
	涨跌幅 (Pct_change)
	成交量 (Volume)
	成交额 (Amount)
	换手率 (Turnover_rate)
	量比 (Volume_ratio)
	市销率 (PS)
交易技术指标	股息率 (Dv_ratio)
	5日均量 (V_ma5)
	20日均量 (V_ma20)
	指数平滑异动平均线 (DEA)
	威廉变异离散量 (WVAD)
	24日相对强弱指标 (RSI3)
	容量比率 (VR)
	随机指标 (KDJ_D)
	6日相对强弱指标 (RSI1)
	情绪指标-心理线 (PSY)
	10日威廉指标 (WR1)

##### (三) 数据预处理

设数据集为  $X = [X_1, \dots, X_N]$ , 其中  $N$  为股票总数,  $X_i$  是第  $i$  只股票的数据集, 分别对  $X_i$  进行预处理。首先选择 Min-Max 进行标准化, 然后采用移动滑窗法生成时序数据集  $S$ , 设定滑动窗口大小为  $t$ , 窗口的滑动步长为  $step$  (即滑动  $step$  个交易日)。滑窗从数据集  $X_i$  的始部向末尾滑动, 同时逐个标记时序的类别。图6显示了当  $step = 1$  时,  $S$  的形成过程。

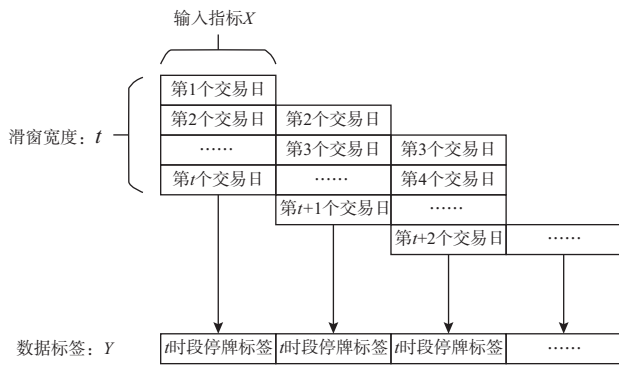


图6 时序数据生成示意图

$$图6中第1个时序数据样本 s_i = \begin{cases} x_{i1} & \cdots & x_{ik} \\ \cdots & \ddots & \cdots \\ x_{i1} & \cdots & x_{ik} \end{cases}$$

其中,  $k=22$  是输入向量维度。  $s_i$  的标签  $y_i \in \{0, 1\}$ , 生成方法如下:

- (1) 设置停牌持续时间阈值为  $\Delta$  天。
- (2) 计算时间间隔 ( $Data(x_{t+1}) - Data(x_t)$ )  $> \Delta$ , 则  $y_t = 1$  表示停牌, 否则  $y_t = 0$  表示正常。

(3) 如果  $y_t = 1$ , 则设置  $s_{t-1}, \dots, s_{t-\epsilon}$  的标签  $y_{t-1}, y_{t-2}, \dots, y_{t-\epsilon}$  为 1, 这是因为在  $\epsilon$  个连续的时序中包含停牌特征。

#### (四) VAE 模型平衡数据集

在股票市场交易中, 股票停牌属于低频行为, 客观上形成了停牌样本的数量大大低于非停牌样本的局面, 分类模型的判别能力会被这种不平衡性严重扭曲。引入 VAE 模型对停牌时序数据进行学习, 进而生成模拟数据以平衡数据集。

##### 1. 模型训练。

将所有标记为停牌的时序数据转为 1 维向量作为 VAE 模型的训练集  $S_{stop} = \{x_1, \dots, x_M\}$ , 其中  $M$  为数据集大小。VAE 的训练如算法 1 所示。

##### 算法 1: 训练 VAE。

输入:  $S_{stop}$ ;  $e$ : 迭代次数。

输出: 已训练 VAE。

- ① 随机初始化各层权重矩阵  $W$  与偏置向量  $b$ ;
- ② for  $epoch = 1 : e$  do
- ③ for  $i = 1 : M$  do
- ④ 数据  $x_i$  输入 VAE 的编码器, 生成  $\mu_i$  和  $\sigma_i^2$ ;
- ⑤ 生成标准正态分布的随机变量  $\varepsilon$ ;
- ⑥ 将  $z = \mu_i + \varepsilon \cdot \sigma_i^2$  输入到 VAE 的解码器, 得到重构向量;
- ⑦ 最小化公式 (4) 的损失函数, 更新  $W$  与  $b$ ;

⑧ end for

⑨ end for

##### 2. 生成模拟数据。

利用已训练的 VAE 模型生成模拟停牌数据  $x'$ 。为了能够评估  $x'$  的质量, 需要度量  $x$  和  $x'$  相似度。通常向量相似度计算的方法包括曼哈顿距离、欧式距离和余弦相似度等, 经过实验测试和对比, 本文选用余弦相似度, 定义如下:

$$\cos(x, x') = \frac{\sum_1^n (x_i \times x'_i)}{\sqrt{\sum_1^n (x_i)^2} \times \sqrt{\sum_1^n (x'_i)^2}} \quad (5)$$

VAE 生成模拟数据过程如算法 2 所示。

##### 算法 2: VAE 生成模拟数据。

输入: VAE;  $S_{stop}$ ;  $e$ : 迭代次数;  $\lambda$ : 相似度阈值;  $\gamma$ : 比较次数阈值;  $m$ : 随机抽取样本次数。

输出: 停牌模拟数据集  $S'_{stop}$ 。

- ① 加载 VAE;
- ② for  $epoch = 1 : e$  do
- ③ VAE 的解码器生成新样本  $x'$ ;
- ④  $cnt = 0$ ;
- ⑤ for  $i = 1 : M$  do
- ⑥ 从  $S_{stop}$  中随机抽取 1 个样本  $x$ ;
- ⑦ 根据公式 (5), 计算相似度  $Sim = \cos(x, x')$ ;
- ⑧ if  $Sim \geq \lambda$  then
- ⑨  $cnt++$ ;
- ⑩ end if
- ⑪ end for
- ⑫ if  $cnt \geq \gamma$  then
- ⑬ 接受  $x'$ , 加入到  $S'_{stop}$  中;
- ⑭ else
- ⑮ 丢弃  $x'$ ;
- ⑯ end if
- ⑰ end for

##### (五) 预训练模型

神经网络模型 TCN、BiLSTM、CNN 和 Attention 等在时间序列分析中皆有良好的表现 (Fan 等, 2021<sup>[30]</sup>; Wang 等, 2021<sup>[31]</sup>; Yang 和 Wang, 2022<sup>[32]</sup>; Abbasi-mehr 和 Paki, 2022<sup>[33]</sup>)。鉴于股票停牌时间序列复杂性, 单一模型在时空特征提取方面存在一定局限性, 本文研究了多种深度学习组合模型包括:

1. TADM: 包括输入层、TCN 层、Attention 层、

Dense 层、输出层。

2. TBADM: 包括输入层、TCN 层、BiLSTM 层、Attention 层、Dense 层、输出层。

3. CBADM: 包括输入层、1D-CNN 层、BiLSTM 层、Attention 层、Dense 层、输出层。

4. BADM: 包括输入层、3 层 BiLSTM、Attention 层、Dense 层、输出层。

5. BDM: 包括输入层、3 层 BiLSTM、Dense 层、输出层。

(六) 个股停牌预测模型

根据基于模型/参数的迁移方法 (Model/ Parameter Based Transfer Learning), 通过预训练-微调方式将预训练模型的部分参数进行迁移, 构建基于预训练模型的个股停牌预测模型, 如图 7 所示。

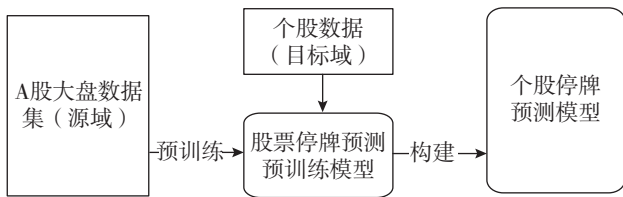


图 7 模型预训练-微调过程

神经网络的浅层完成一般特征 (general feature) 的学习, 而深层需完成和具体任务关联的特殊特征 (specific features) 的学习。伴随网络层次的深入, 模型的学习表征逐渐由一般特征平滑到特殊特征。由于源域数据集和目标数据集的类别一致性, 参数迁移时采用 all-to-all 迁移策略即目标域的第  $m$  层由源域中的第  $n$  层迁移得到。以 TADM 为例, 将输入层到注意力层参数逐层迁移到个股停牌预测模型对应各层之中, 如图 8 所示。

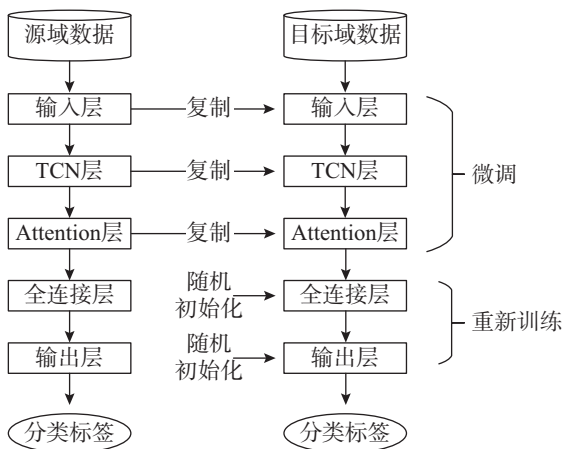


图 8 模型预训练-微调过程

参数迁移完成后, 个股预测模型利用预训练模型

强大的股票停牌特征提取能力从目标域数据中分析潜在停牌信息, 鉴于目标个股自身的波动特征, 需要个股预测模型对其提取以提升预测的准确度, 将模型全连接层与输出层随机初始化后再重新进行训练。

个股的训练集  $S^{(i)}$  存在极度不平衡的问题, 如股票 002352 的正常类 1 702 条, 而停牌类只有 10 条。设第  $i$  只股票的停牌时序数据集为  $S_{stop}^{(i)}$ , 在算法 2 的基础上, 设计个股的模拟数据生成算法如下:

算法 3: VAE 生成个股模拟数据。

输入: VAE;  $S_{stop}^{(i)}$ ;  $e$ : 迭代次数;  $\lambda$ : 相似度阈值;  $m$ : 生成模拟样本数量。

输出: 停牌模拟数据集  $S_{stop}^{(i)'}$ 。

- ① 加载 VAE;
- ② for  $n = 1 : \text{length}(S_{stop}^{(i)})$  do
- ③  $cnt = 0$ ;
- ④  $x = S_{stop}^{(i)}[n]$ ;
- ⑤ for epoch = 1 :  $e$  do
- ⑥ VAE 的解码器生成新样本  $x'$ ;
- ⑦ 根据公式 (5), 计算相似度  $Sim = \cos(x, x')$ ;
- ⑧ if  $Sim > \lambda$  then
- ⑨  $cnt ++$ ;
- ⑩ 接受  $x'$ , 加入到  $S_{stop}^{(i)'}$  中;
- ⑪ end if
- ⑫ if  $cnt \geq m$  then
- ⑬ break;
- ⑭ end if
- ⑮ end for
- ⑯ end for

个股的每个停牌时间序列生成  $m$  条模拟数据, 由于模型 VAE 是由所有停牌类数据训练而得, 个股的部分停牌时间序列可能背离整体特征导致无法在特定  $\lambda$  下产生模拟数据。

(七) 评价标准

通常的模型评估指标包括准确率 (Accuracy)、精准率 (Precision)、召回率 (Recall) 以及 F1 值 (F1-score), 四个比率值的区间为  $[0, 1]$ , 值越趋近于 0 表示预测表现越差, 越趋近于 1 表示预测表现越好。

对所有股票的停牌类样本分类正确率反映了预训练模型提取 A 盘整体停牌特征的能力, 因此设置  $Q$  评分如下:

1. 将停牌类别 ( $y = 1$ ) 的分类正确率 TP (True

Positive) 划分 3 个区间:  $L_1 = [0.8, 1.0]$ ,  $L_2 = [0.5, 0.8)$  和  $L_3 = [0.0, 0.5)$ , 分别设置权重  $w_1 = 1.0$ ,  $w_2 = 0.6$ ,  $w_3 = 0.2$ 。

2. 统计落在  $L_1$ ,  $L_2$  和  $L_3$  中的股票数:  $c_1$ ,  $c_2$  和  $c_3$ 。

3. 定义  $Q$ :

$$Q = \frac{w_1 c_1 + w_2 c_2 + w_3 c_3}{N} \quad (6)$$

其中  $N$  是股票总数。

### 五、实验及结果分析

#### (一) 实验数据

以我国 A 股市场 2 539 家上市公司作为研究对象, 其中 2 家公司未发生停牌事件。从 Tushare 平台采集股票交易基本数据, 包括日线行情与每日指标数据, 并由其计算技术分析指标。采集时间区间为 2013 年 5 月 13 日至 2021 年 9 月 30 日。统计各指标数据并进行数据的计算处理后, 获得了共计 4 390 024 条记录的实验数据集, 每一条数据中包括 12 个股票交易基本数据和 10 个技术分析指标, 合计 22 个属性特征。

本文主要利用市场中长期信息研究 A 股的停牌问题, 因此采用移动滑窗法生成股票停牌时间序列数据集, 各参数设置分别为: 时间滑窗宽度  $t = 15$  个交易日, 停牌持续时间阈值  $\Delta = 7$  天, 窗口的滑动步长  $step$  为 1, 参数  $\epsilon = 7$ 。实验共生成了 3 764 841 条时间序列数据, 其中正常类 3 740 960 条, 停牌类 23 881 条, 数据维度为  $[15, 22]$ 。为减少正常类的冗余和降低样本不平衡程度, 每隔 10 个交易日重新抽样, 生成训练集的正常类为 373 772 条。

#### (二) VAE 模型平衡训练集

选取所有停牌时序数据作为 VAE 模型训练集, 并将时间序列转换为 1 维向量, 其维度为  $330 = 15 \times 22$ 。设置 VAE 模型参数结构为:  $330 - 128 - 64 - 32 - [4, 4] - 32 - 64 - 128 - 330$ , 即输入层与输出层包含 330 个神经元, Encoder 层包含 3 个隐含层, 神经元个数分别是 128、64 和 32, 而 Decoder 层也包含 3 个隐含层, 神经元数为 32、64 和 128。隐变层的  $\mu$  和  $\sigma^2$  分别有 4 个神经元, 计算  $z = \mu + \epsilon \cdot \sigma^2$  得到隐变量输入到 Decoder 层。

实验中设置算法 1 中的参数  $M = 23 881$ ,  $e = 100$ 。

算法 2 中的参数  $\lambda = 0.85$ ,  $m = 10$ ,  $\gamma = 6$ ,  $e = 1 000 000$ 。已训练 VAE 模型生成模拟停牌数据 71 065 条, 平衡前后训练集类别样本的分布见表 2 所示。

表 2 平衡前后训练集样本分布

类别	平衡前	平衡后
正常	373 772	373 772
停牌	23 881	94 946

为了验证 VAE 模型的有效性, 采用 BDM 模型测试平衡前后的训练集, 实验结果如表 3 所示。

表 3 训练集平衡前后的对比实验结果

模型	类别	准确率	精准率	召回率	F1 值
平衡前	停牌	0.961	0.721	0.512	0.599
	正常		0.971	0.988	0.979
平衡后	停牌	0.964	<b>0.954</b>	<b>0.861</b>	<b>0.905</b>
	正常		0.966	0.989	0.977

表 3 显示采用 VAE 模型对样本类别不平衡问题的处理对提升模型分类效果十分明显。

#### (三) 预训练模型实验及分析

抽取平衡前训练集的 2% 作为预训练模型的测试集, 将剩余数据进行类别平衡后, 取其 80% 为训练集, 剩下 20% 为验证集。

##### 1. 模型参数设置。

模型 TADM、TBADM 和 CBADM 中的 BiLSTM 层、TCN 层、1D-CNN 层的主要参数设置如表 4 所示。

表 4 主要参数

层名	参数	值
TCN	filters	64
	kernel_size	3
	stacks	1
	dilations	(1, 2, 4, 8, 16, 32, 64)
	activation	ReLU
BiLSTM	units	64
	activation	TanH
1D-CNN	filters	64
	kernel_size	1
	activation	ReLU

模型 BADM 和 BDM 中 3 层 BiLSTM 层的 units 分别为: 64, 32 和 16, 激活函数设为 TanH。

### 2. 对比实验及结果分析。

模型训练参数设为: epochs = 100, batch size = 32 和 dropout = 0.2, 采用 Adam 优化器, 学习率设为 0.001。5 个模型在测试集上的实验结果如表 5 所示。

表 5 模型实验结果

模型	准确率	精准率	召回率	F1 值	Q
TADM	0.954	0.933	0.923	0.928	0.902
TBADM	0.970	0.974	0.930	0.950	0.491
CBADM	0.967	0.972	0.924	0.946	0.539
BADM	0.969	0.978	0.925	0.948	0.527
BDM	0.964	0.960	0.926	0.942	0.648

表 5 显示 TBADM 的模型性能最佳, 但 TCN 已经拥有了 CNN 和 LSTM 的优点, 能提取较大范围的时间序列特征, 而进一步使用 BiLSTM 层学习 TCN 输出的双向特征以及注意力机制突出了其输出特征向量, 导致其过拟合严重。CBADM、BADM 和 BDM 的模型性能相似, 也存在过拟合问题, 源于受限于 BiLSTM 层对时间序列建模的长度。TADM 的模型性能虽低, 但其学习 A 盘整体停牌特征能力是最强的, 因此本文以 TADM 作为预训练模型。用模型 TADM 分别对 2 539 个股的时序数据 (不包含模拟停牌数据) 进行分类测试, 表 6 统计了落在不同 TP 区间中的股票数量。

表 6 预测结果统计

区间	个数	百分比
0~40%	0	0
40%~50%	6	0.24%
50%~60%	42	1.66%
60%~70%	181	7.13%
70%~80%	386	15.21%
80%~90%	823	32.44%
90%~100%	367	14.47%
100%	732	28.85%

表 6 中股票数合计 2 537 只, 有 2 只股票没有停牌记录, 结果显示 TP ≥ 80% 的股票数合计占比超过总体股数的 75% 以上。

### 3. 模型性能的行业分析。

进一步分析预训练模型 TADM 在各个行业中的预测性能。依据证监会行业划分标准, 在所有 79 个行业中选取股票数超过 100 的 6 个行业, 如表 7 所示。

表 7 行业股票分布统计

行业分类	股票数	占比
C39: 计算机、通信和其他电子设备制造业	188	7.41%
C26: 化学原料和化学制品制造业	156	6.15%
C27: 医药制造业	144	5.68%
C38: 电气机械和器材制造业	142	5.60%
K70: 房地产业	114	4.49%
C35: 专用设备制造业	107	4.22%

图 9 给出了落在区间  $L_1$ 、 $L_2$  和  $L_3$  的 6 个行业股票数占比直方图。

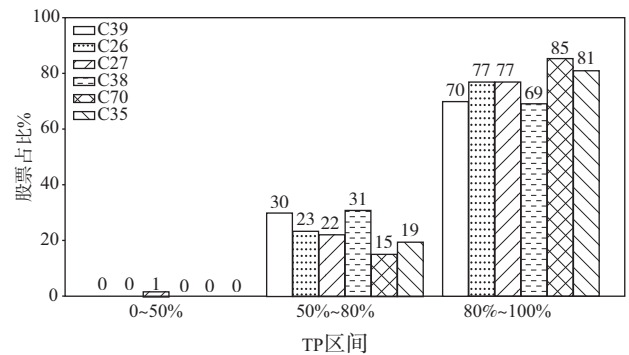


图 9 不同行业的模型性能对比直方图

图 9 显示针对 6 个行业的股票预测, TADM 模型皆具有良好的表现。只有一只股票的 TP 低于 50% (46.15%), 是属于行业 C27 中某药业, 曾一直被怀疑存在严重的财务舞弊, 最终因其造假被证监会处罚, 模型对其预测 TP 结果说明虚假财务数据误导了投资者, 背离了模型从正常交易中发掘的停牌规律, 所以无法进行有效预测。另外, 不同行业受到的政策市场等影响不一样, 因此它们的停牌规律是有差异的, TADM 模型的表现也就略有不同, 如表 8 所示。

表 8 不同行业的模型性能分析

行业分类	股票数 (TP = 100%)	TP 均值	方差
C39	48 (25.5%)	84.06%	0.0172
C26	47 (37.6%)	86.65%	0.0130
C27	43 (29.86%)	87.24%	0.0128
C38	39 (27.47%)	85.26%	0.0153
K70	<b>34 (29.82%)</b>	<b>88.05%</b>	<b>0.0107</b>
C35	<b>34 (31.76%)</b>	<b>87.78%</b>	<b>0.0109</b>

表 8 显示 K70 和 C35 的 TP 均值较高且方差较小, 说明 TADM 模型在这 2 个行业中表现更佳。

### (四) 个股停牌预测模型实验及分析

基于 TADM 预训练模型的迁移学习构建个股的停

牌预测模型 STFTADM (Single Transfer Learning Model base on TADM)。TADM 模型的参数共计 174 209 个,设计重新训练参数的数量为 1 921,即只有 1%的参数参与学习,如图 8 所示。

1. 个股选择与平衡处理。

为了较全面评估迁移学习的效果,实验中分别从区间  $L_1$ 、 $L_2$  和  $L_3$  中选取股票。选取表 6 中落在  $L_1$  区间的所有 6 只股票。被选取股票的相关信息如表 9 所示。

表 9 选取的股票信息

TP 区间	股票代码	TP	停牌次数	正常次数
$L_1$	002352	40.00%	10	1 702
	603843	40.00%	10	809
	600462	46.15%	13	1 108
	600518	46.15%	13	1 376
	600753	47.06%	17	1 258
	600209	47.06%	17	908
$L_2$	002571	50.00%	14	1 208
	000100	55.55%	9	1 673
	002356	64.28%	14	951
	600289	71.43%	14	1 296
	600603	78.57%	14	1 373
$L_3$	002828	85.71%	7	924
	600332	90.90%	11	1 710

表 9 显示个股的训练集存在极度不平衡,采用算法 3 进行处理,设置参数  $\lambda=0.83$ ,  $m=50$ ,  $e=10\ 000$ 。生成的个股模拟停牌数据如表 10 所示。

表 10 个股模拟停牌数据

股票代码	模拟停牌次数	停牌次数	模拟失败数
002352	408	10	2
603843	306	10	4
600462	612	13	1
600518	408	13	5
600753	663	17	4
600209	612	17	5
002571	510	14	4
000100	408	9	1
002356	663	14	1
600289	510	14	4
600603	561	14	3
002828	204	7	3
600332	510	11	1

表 10 显示个股的个别停牌类数据未能产生模拟数据,这是由于模型 VAE 是由 A 股所有停牌类数据训练而得,个股的部分停牌时间序列可能背离整体特征,导致其在阈值 0.83 下无法产生对应的模拟数据。

2. 模型训练与结果分析。

将个股平衡后的数据集依照 6 : 2 : 2 的比例依次划分为训练集、验证集与测试集。模型 STFTADM 训练参数设为: epochs = 25, batch size = 1 和 dropout = 0.2,采用 Adam 优化器,学习率的设置:  $L_1$  中的股票为 0.005;  $L_2$  为 0.001;  $L_3$  为 0.000 5。各个模型 STFTADM 测试对应个股的时序数据(不包含模拟停牌数据),与对应的预训练模型测试的对比结果如表 11 所示。

表 11 个股模型实验结果对比

股票代码	模型	类别	精准率	召回率	F1 值
002352	TADM	停牌	0.088 9	0.400 0	0.145 5
		正常	0.996 4	0.975 9	0.986 0
	STFTADM	停牌	1.000 0	0.800 0	0.888 9
		正常	0.998 8	1.000 0	0.999 4
603843	TADM	停牌	0.266 7	0.400 0	0.320 0
		正常	0.992 5	0.986 4	0.989 5
	STFTADM	停牌	0.909 1	1.000 0	0.952 4
		正常	1.000 0	0.998 8	0.999 4
600462	TADM	停牌	0.117 6	0.461 5	0.187 5
		正常	0.993 5	0.959 4	0.976 1
	STFTADM	停牌	0.857 1	0.923 1	0.888 9
		正常	0.999 1	0.998 2	0.998 6
600518	TADM	停牌	0.095 2	0.461 5	0.157 9
		正常	0.994 7	0.958 6	0.976 3
	STFTADM	停牌	0.750 0	0.923 1	0.827 6
		正常	0.999 3	0.997 1	0.998 2
600753	TADM	停牌	0.142 9	0.470 6	0.219 2
		正常	0.992 6	0.961 8	0.977 0
	STFTADM	停牌	0.764 7	0.764 7	0.764 7
		正常	0.996 8	0.996 8	0.996 8
600209	TADM	停牌	0.250 0	0.470 6	0.326 5
		正常	0.989 9	0.973 6	0.981 7
	STFTADM	停牌	0.833 3	0.882 4	0.857 1
		正常	0.997 8	0.996 7	0.997 2
002571	TADM	停牌	0.122 8	0.500 0	0.197 2
		正常	0.994 0	0.958 6	0.976 0
	STFTADM	停牌	1.000 0	0.928 6	0.963 0
		正常	0.999 2	1.000 0	0.999 6

续前表

股票代码	模型	类别	精准率	召回率	F1 值
000100	TADM	停牌	0.131 6	0.555 6	0.212 8
		正常	0.997 6	0.980 3	0.988 8
	STFTADM	停牌	1.000 0	0.777 8	0.875 0
		正常	0.998 8	1.000 0	0.999 4
002356	TADM	停牌	0.187 5	0.642 9	0.290 3
		正常	0.994 5	0.959 0	0.976 4
	STFTADM	停牌	0.875 0	1.000 0	0.933 3
		正常	1.000 0	0.997 9	0.998 9
600289	TADM	停牌	0.217 4	0.714 3	0.333 3
		正常	0.996 8	0.972 2	0.984 4
	STFTADM	停牌	0.866 7	0.928 6	0.896 6
		正常	0.999 2	0.998 5	0.998 8
600603	TADM	停牌	0.220 0	0.785 7	0.343 7
		正常	0.997 8	0.971 6	0.984 5
	STFTADM	停牌	0.928 6	0.928 6	0.928 6
		正常	0.999 3	0.999 3	0.999 3
002828	TADM	停牌	0.260 9	0.857 1	0.400 0
		正常	0.998 9	0.981 6	0.990 2
	STFTADM	停牌	0.777 8	1.000 0	0.875 0
		正常	1.000 0	0.997 8	0.998 9
600332	TADM	停牌	0.204 1	0.909 1	0.333 3
		正常	0.999 4	0.977 2	0.988 2
	STFTADM	停牌	0.846 2	1.000 0	0.916 7
		正常	1.000 0	0.998 8	0.999 4

表 11 的实验结果显示个股的 STFTADM 模型相对于 TADM 的测试效果有着较大幅度的提高,如股票 002356 的停牌分类正确率由 64.285 7% 提升到 100%。表 11 中东方银星(股票代码:600753)的预测结果较差,其 STFTADM 模型的 TP = 76.470 6%,该公司曾亏损约 5 002 万元,经营性现金流大幅下降,近年才由盈转亏。康美药业(股票代码:600518)的 STFTADM 模型的 TP = 75%,该公司曾存在严重的财务舞弊。贝肯能源(股票代码:002828)的 STFTADM 模型的 TP = 77.78%,该公司作为民营企业,其收入对中石油集团单一客户依赖性较高,存在业绩下滑的风险。

相较于预训练模型 TADM,个股的模型 STFTADM 表现出更强的样本识别能力且误报率较低。这是因为利用了预训练模型的网络层参数,能较好地捕获停牌样本的时间序列特征,同时基于预训练模型在无形中拓展了目标个股数据集,使得模型更加鲁棒,增强了

模型的泛化能力。

## 六、结论

本文基于停牌预测的视角,选取 2013—2021 年 A 股市场 2 539 家上市公司的数据,采用预训练模型和迁移学习等深度学习方法对股票停牌事件进行预测研究。研究结果表明:(1)虽然现行股票停牌存在随意性强、时间长以及诱因复杂性问题,但仍能从中长期金融数据波动中提取其中关于停牌的特征模式,引入深度学习模型能学习到所有上市公司股票停牌的共性特征,在此基础上通过迁移学习机制获得个股停牌特征,进而实现精确预测。(2)在股票市场交易中,股票停牌属于低频行为,本文提出的基于 VAE 的停牌模拟数据生成算法能有效模拟停牌时序数据,极大改善了深度学习分类模型的性能。(3)预训练模型能较好地预测 A 股市场各个行业板块的停牌事件,约 75% 的股票停牌预测准确度在 80% 以上,六个行业板块的实验结果分析显示房地产业和专用设备制造业的预测效果最佳。在此基础上的个股迁移学习模型测试结果表现出更高的预测精度。(4)预测效果较差的上市公司往往存在较严重的财务舞弊或经营不善等问题,其中长期金融数据波动背离了模型从正常交易中发掘的停牌规律。

A 股市场的随意停牌现象不仅限制了投资者正常交易的权利,而且还存在着信息不对称下内部人对于股东权益侵蚀的可能(石阳等,2019<sup>[15]</sup>)。本文研究能够对未来可能发生的股票停牌做出较为准确的判断,能为资本市场提供更好的信息服务。基于此,本文提出以下相关建议:(1)我国中小投资者在信息不对称的情况下,一方面可以借鉴个股停牌预测结果,有效地规避停牌风险,进而减少投资损失以及实现资产保值;另一方面也可为部分投资者利用个股停牌后的股价波动,实现收益最大化等目标提供帮助。(2)个股停牌预测结果在一定程度上反映了上市公司的财务或经营状况,投资者可以适当规避停牌预测结果较差的个股,以优化投资方案。(3)本文的研究为证监会与交易所预测停牌行为提供了可能,监管部门可以动态监控 A 股市场各个行业板块的停牌事件,及时对整体停牌趋势做出前瞻性判断,为宏观调控和政策调整提供了参考。



## 参考文献

- [1] Doroudyan M H, Owlia M S, Sadeghi H, et al. Monitoring Financial Processes with ARMA-GARCH Model Based on Shewhart Control Chart (Case Study: Tehran Stock Exchange) [J]. *International Journal of Engineering, Transactions B: Applications*, 2017, 30 (2): 270-280.
- [2] 蔡光辉, 徐君, 应雪海. 非对称厚尾分布的混频 Realized GARCH 模型构建 [J]. *统计与决策*, 2021 (2): 130-135.
- [3] Gupta R, Nel J, Pierdzioch C. Investor Confidence and Forecastability of US Stock Market Realized Volatility: Evidence from Machine Learning [J]. *Journal of Behavioral Finance*, 2021: 1-12.
- [4] Vara P V, Srinivas G, Venkataramana L Y, et al. Prediction of Stock Prices Using Statistical and Machine Learning Models: A Comparative Analysis [J]. *The Computer Journal*, 2022 (5): 1338-1351.
- [5] 张剑, 王波. 基于 SVM 的沪深 300 股指期货期货量化交易策略 [J]. *数学理论与应用*, 2017 (2): 112-121.
- [6] Chatzis S P, Siakoulis V, Petropoulos A, et al. Forecasting stock market crisis events using deep and statistical machine learning techniques [J]. *Expert Systems with Applications*, 2018, 112 (DEC.): 353-371.
- [7] Althelaya K A, Mohammed S A, El-Alfy E. Combining Deep Learning and Multiresolution Analysis for Stock Market Forecasting [J]. *IEEE Access*, 2021 (9): 13099-13111.
- [8] Sun M T, Kitchat K, Hsieh L C. TCN-based Futures Prediction Using Financial Indices, Bargain Chips, and Forum Messages [C]//2021 International Conference on Technologies and Applications of Artificial Intelligence (TAAI). *IEEE*, 2021: 72-77.
- [9] Kabir R, Engelen P J. Empirical Evidence on the Role of Trading Suspensions in Disseminating New Information to the Capital Market [J]. *Social Science Electronic Publishing*, 2006, 33 (7/8): 1142-1167.
- [10] Taechapiroontong N, Charoenwong C, Chiraphol C N, et al. The Effectiveness of Trading Halts and Investor Trading Performance [J]. *International Research Journal of Finance and Economics*, 2012 (102): 191-209.
- [11] Rahim R, Sulaiman D, Husni T, et al. Investor Behavior Responding to Changes in Trading Halt Conditions: Empirical Evidence from the Indonesia Stock Exchange [J]. *The Journal of Asian Finance, Economics and Business*, 2021, 8 (4): 135-143.
- [12] 陈舒宁, 张维, 何枫, 等. 中小板交易异动停牌制度有效性研究 [J]. *系统工程学报*, 2016 (6): 831-839.
- [13] 胡婷, 惠凯, 彭红枫. 异常波动停牌对股价波动性和流动性的影响研究——来自我国取消异常波动停牌的自然实验 [J]. *金融研究*, 2017 (9): 146-160.
- [14] 李洋, 王春峰, 房振明, 等. 异常波动停牌与价格发现效率 [J]. *管理科学*, 2018 (6): 33-45.
- [15] 石阳, 刘瑞明, 王满仓. 上市公司随意停牌与投资者利益——来自中国资本市场的证据 [J]. *经济研究*, 2019 (1): 36-51.
- [16] 李明琨, 葛艺博, 李慧. 有限理性影响下的证券交易异常行为博弈与政府监管研究 [J]. *管理评论*, 2022 (1): 51-58.
- [17] He Q, Gan J, Wang S, et al. The Effects of Trading Suspensions in China [J]. *The North American Journal of Economics and Finance*, 2019 (50): 100985.
- [18] Fama E F. Efficient Capital Markets: A Review of Theory and Empirical Work [J]. *Journal of Finance*, 1970, 25 (2): 383-417.
- [19] Bai S, Kolter Z, Koltun V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling [J]. *ArXiv Preprint ArXiv: 1803.01271*, 2018.
- [20] Mnih V, Heess N, Graves A, et al. Recurrent Models of Visual Attention [J]. *Advances in Neural Information Processing System*, 2014, 3.
- [21] 刘建伟, 刘俊文, 罗雄麟. 深度学习中注意力机制研究进展 [J]. *工程科学学报*, 2021 (11): 1499-1511.
- [22] Raffel C, Ellis D. Feed-Forward Networks with Attention Can Solve Some Long-term Memory Problems [J]. *ArXiv Preprint ArXiv: 1512.08756*, 2015.
- [23] Kumar P, Hati A S, Sanjeevikumar P, et al. Amalgamation of Transfer Learning and Deep Convolutional Neural Network for Multiple Fault Detection in SCIM [C]//20th IEEE Intl. Conf. on Environment and Electrical Engineering, IEEE I&CPS 20, Madrid (Spain). *IEEE*, 2020.
- [24] Gu Q, Dai Q. A Novel Active Multi-source Transfer Learning Algorithm for Time Series Forecasting [J]. *Applied Intelligence*, 2021, 51 (2): 1-25.
- [25] Kingma D P, Welling M. Auto-encoding Variational Bayes [J]. *Arxiv Preprint ArXiv: 1312.6114*, 2013.
- [26] Utyamishv D, Partin-Vaisband I. Progressive VAE Training on Highly Sparse and Imbalanced Data [J]. *ArXiv Preprint ArXiv: 1912.08283*, 2019.
- [27] Zhou Y, Liang X, Zhang W, et al. VAE-based Deep SVDD for Anomaly Detection [J]. *Neurocomputing*, 2021, 453: 131-140.
- [28] Liu C, Antypenko R, Sushko I, et al. Intrusion Detection System after Data Augmentation Schemes Based on the VAE and CVAE [J]. *IEEE Transactions on Reliability*, 2022, 71 (2): 1000-1010.
- [29] 谢丁. 趋向类技术分析指标的实证研究及评测 [J]. *金融经济*, 2009 (11): 38-40.
- [30] Fan J, Zhang K, Huang Y, et al. Parallel Spatio-temporal Attention-based TCN for Multivariate Time Series Prediction [J]. *Neural Computing and Applications*, 2021: 1-10.
- [31] Wang Y, Zhang Z, Feng L, et al. A New Attention-based CNN Approach for Crop Mapping Using Time Series Sentinel-2 Images [J]. *Computers and Electronics in Agriculture*, 2021, 184: 106090.
- [32] Yang M, Wang J. Adaptability of Financial Time Series Prediction Based on BiLSTM [J]. *Procedia Computer Science*, 2022, 199: 18-25.
- [33] Abbasimehr H, Paki R. Improving Time Series Forecasting Using LSTM and Attention Models [J]. *Journal of Ambient Intelligence and Humanized Computing*, 2022, 13 (1): 673-691.

# “监督效应” VS “掏空效应”：高管纵向兼任与审计决策

“Monitoring Effect” VS “Tunneling Effect”：

Vertical Interlocks of Executives and Audit Decisions

宋衍蘅 毕煜晗 文雯

SONG Yan-heng BI Yu-han WEN Wen

**[摘要]** 高管纵向兼任有助于大股东更好地监督上市公司（监督效应），还是更方便给大股东输送利益（掏空效应）？本文采用2004—2018年中国沪深A股上市公司的数据，从审计决策视角探究了高管纵向兼任的经济后果。实证研究发现，高管纵向兼任发挥了“监督效应”：审计师减少了存在高管纵向兼任公司的审计投入、降低了审计费用、更有可能出具标准无保留审计意见。机制检验表明，高管纵向兼任通过提高内部控制质量、降低股东与管理层之间的代理成本发挥了良好的监督效应，进而影响了审计决策。拓展性研究显示，高管纵向兼任对审计决策的影响在大股东控制力较弱样本中更加显著；相比于总经理兼任，董事长纵向兼任对审计决策的影响更强；区分产权性质后发现，高管纵向兼任对审计决策的影响在国有企业中更加显著。本文的研究结果拓展了高管纵向兼任的经济后果和审计决策的影响因素研究，对于完善公司治理体系亦有借鉴意义。

**[关键词]** 高管纵向兼任 审计投入 审计收费 审计意见

**[中图分类号]** F239.0 **[文献标识码]** A **[文章编号]** 1000-1549 (2022) 11-0052-13

**Abstract:** Do vertical interlocked executives help controlling shareholders better monitor listed companies (monitoring effect), or it is more convenient to convey benefits to controlling shareholders (tunneling effect)? Using the data of Chinese A-share listed companies from 2004 to 2018, this paper empirically examines the economic consequences of vertical interlocks of executives from the perspective of audit decisions. The results indicate that vertical interlocks of executives play a “monitoring” role from the auditors’ point of view. In detail, auditors reduce audit investment, charge lower audit fees and are more likely to issue clean audit opinions for these firms. Channel tests show that vertical interlocks of executives affect audit decisions by improving internal control quality and reducing agency costs between shareholders and managers. Further analysis shows that the influence of vertical interlocks of executives on audit decisions is more pronounced in companies with fewer large shareholder control rights. Compared with vertical interlocked CEOs, vertical interlocked chairmen have a more significant impact on audit decisions. Distinguishing the nature of property rights, the influence of vertical interlocks of executives on audit decisions is more pronounced in state-owned companies. This paper extends the literature on the economic consequences of vertical interlocks of executives and the determinants of audit decisions, which also have implications for improving corporate governance system.

**Key words:** Vertical interlock of executive Audit investment Audit fee Audit opinion

**[收稿日期]** 2020-11-10

**[作者简介]** 宋衍蘅，女，1973年3月生，北京外国语大学国际商学院副教授，研究方向为审计与资本市场；毕煜晗，女，1995年8月生，北京师范大学经济与工商管理学院博士研究生，研究方向为审计与资本市场；文雯，女，1991年9月生，北京外国语大学国际商学院副教授，研究方向为公司财务与公司治理。本文通讯作者为文雯，联系方式为 wenwen\_bfsu@bfsu.edu.cn。

**[基金项目]** 国家自然科学基金项目“中国多梯队审计市场竞争格局及其经济后果研究——基于不同会计师事务所竞争策略的比较分析”（项目编号：72072015）；国家自然科学基金项目“上市公司精准扶贫行为的同群效应：作用机制与经济后果”（项目编号：72002014）；北京外国语大学“双一流”重大（点）标志性项目研究成果“后疫情时期全球化风险研究：金融安全与商务风险视角”（项目编号：2022SYLZD001）。

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

高管纵向兼任是指“上市公司董事长或总经理在控股股东或其实际控制人处任职”的情形（潘红波和韩芳芳，2016<sup>[1]</sup>；闫珍丽等，2019<sup>[2]</sup>）。高管纵向兼任是控股股东与上市公司之间的一种纵向联结形式，有助于加强控股股东对上市公司的监督，但是，在存在利益冲突的情况下，也可能加剧控股股东对中小股东的利益侵占。为解决这个问题，中国证监会要求“上市公司的高级管理人员在控股股东不得担任除董事、监事以外的其他行政职务<sup>①</sup>”。也就是说，中国证监会一方面认可控股股东在上市公司发挥监督作用，另一方面，又试图规避可能产生的利益侵占问题。那么，高管纵向兼任是否可以更好地发挥监督作用？又是否会带来利益侵占问题？这仍然是一个有待实证检验的问题。

我国自2006年起，确定了风险导向型的审计模式，审计师需全面、客观地识别被审计单位的经济活动风险，保证不存在重大错报风险，审慎做出审计决策。本文探究高管纵向兼任对审计决策的影响，从审计投入、审计费用、审计意见三个方面探讨公司会计信息的鉴证者——审计师对高管纵向兼任现象的理解。

从理论上，高管纵向兼任对审计决策将产生怎样的影响并不确定。一方面，高管纵向兼任是大股东加强对上市公司监督的一种方式，可以发挥“监督效应”。大股东在上市公司兼职，可以通过参加上市公司的董事会或日常经营管理，加强对高管的监督，抑制高管的机会主义行为，从而减轻了公司大股东和管理层间信息不对称程度。高管纵向兼任可以加强大股东与上市公司之间的联系，使得上市公司获得大股东的更多支持和资源。现有研究发现，高管纵向兼任提高了上市公司的会计信息质量（潘红波和韩芳芳，2016<sup>[1]</sup>）和投资效率（韩金红和余珍，2019<sup>[3]</sup>），减少了高管的短视行为，提高了企业的风险承担水平（佟爱琴和李孟洁，2018<sup>[4]</sup>），并最终提升了公司价值（Arnoldi等，2013<sup>[5]</sup>）。也正因为如此，国资委在央企改革过程中力推集团公司董事长兼任上市公司

“一把手”<sup>②</sup>，以更好地实现国有资产保值增值的目标。

另一方面，高管纵向兼任还可能助长大股东对中小股东的利益攫取行为，发挥“掏空效应”。现代企业的金字塔型控股使得控制权与现金流权相偏离，大股东较大的控制权使其有能力侵占公司资源，而较小的现金流权则降低了侵占成本。高管纵向兼任作为大股东加强控制权的一种方式，方便了大股东直接干预上市公司决策，也为大股东掏空上市公司资源开了一扇门。现有文献发现，高管纵向兼任增强了大股东的机会主义行为，增加了公司现金被滥用的风险（Chen和Yang，2021<sup>[6]</sup>），同时抑制了企业创新（闫珍丽等，2019<sup>[2]</sup>）。2019年曝光的康得新122亿资金离奇“失踪”事件，将大股东与上市公司在资金管理使用上的混乱局面公之于众，重新将大股东掏空问题推到前台。

那么，与上市公司进行直接沟通和交流、为上市公司会计信息提供鉴证服务的专业人士——审计师又是如何看待高管纵向兼任的呢？尚无文献针对这一问题展开讨论。本文以2004—2018年中国非金融类A股上市公司为研究样本，从审计投入、审计费用和审计意见三个角度探讨高管纵向兼任对审计决策的影响。本文的主要研究结论是：（1）审计师的行为表现认可了高管纵向兼任的“监督效应”，主要表现在审计师对存在高管纵向兼任的公司审计投入更少，审计收费更低，出具标准无保留审计意见的可能性更高。（2）机制检验结果表明，高管纵向兼任提高了上市公司的内部控制质量、降低了股东与管理层之间的代理成本，从而发挥了良好的“监督效应”。（3）拓展性研究显示，大股东控制权较弱的企业和国有企业中，高管纵向兼任对审计决策的影响更加显著；区分纵向兼任高管的职位特征后发现，董事长纵向兼任对审计决策的影响更强。

本文的研究贡献主要体现在以下三个方面：第一，本文从审计师视角拓展了高管兼任的经济后果研究。以往研究较多关注了高管横向兼任多家企业的管理职位（如连锁董事）的影响和溢出效应，对中国企业中普遍存在的高管纵向兼任问题关注较少。仅有

① 资料来源：中国证监会《上市公司治理准则》（2002年出台、2018年修订）第六章第六十九条。同时，该准则规定“……控股股东高级管理人员兼任上市公司董事、监事的，应该保证有足够的时间和精力承担上市公司的工作”。

② 如中国石油天然气集团公司董事长、党组书记王宜林，兼任中石油上市公司董事长，中石化集团董事长戴厚良兼任中国石化（上市公司）董事长，国家能源集团董事长、党组书记王祥喜兼任中国神华能源股份有限公司董事长等。

的几篇文献主要关注了高管纵向兼任对上市公司投资决策和绩效的影响,如企业创新、会计信息质量和企业价值等(Arnoldi等,2008<sup>[5]</sup>;潘红波和韩芳芳,2016<sup>[1]</sup>;闫珍丽等,2019<sup>[2]</sup>)。但是,资本市场参与者众多,很多投资者并没有对公司进行深入调研,其决策和看法不一定客观理性。本文从审计专业人士的角度重新审视这个问题,考察专业信息鉴证者深入公司调查以后的判断结果,希望可以提供一个不同维度的经验证据。第二,本文探究了高管纵向兼任发挥监督作用的机制,即主要是通过提高上市公司内部控制质量和降低股东与管理层之间的代理成本实现的,为加深对这个问题的理解提供了经验证据。第三,本研究丰富了有关审计决策的影响因素研究。已有关于审计决策的研究大多关注公司财务风险和审计风险,本文探讨了公司治理方式的可能影响,拓展了审计决策作用机理领域的研究。

## 二、文献回顾与研究假设

现有公司治理研究关注两类代理问题:其一是股东(委托人)与管理层(代理人)之间的第一类代理问题;其二是拥有实际控制权和私有信息的大股东和无法进入董事会、只能通过公开信息进行决策的中小股东之间的第二类代理问题(郑志刚和孙娟娟,2009<sup>[7]</sup>)。高管纵向兼任的直接目的是大股东通过在上市公司兼任职务,来强化其对上市公司的控制力和监督能力,发挥“监督效应”,降低股东与管理层之间的代理成本。但是,高管纵向兼任反过来也会成为大股东掏空上市公司的“帮凶”,加剧大小股东之间的代理问题,形成“掏空效应”。本文综述高管纵向兼任领域两种观点的文献,并分析其对审计决策的影响。

### (一) 高管纵向兼任的“监督效应”

在第一类代理问题中,主要的利益冲突方是股东和管理层(Jensen和Meckling,1976<sup>[8]</sup>),降低此类代理成本的主要方式有两种,一种是通过股权激励方式使高管的利益与股东趋于一致,另一种是通过公司治理机制监督管理层行为。由于大股东很难通过分散投资减少风险,因此,大股东对管理层的监督就显得尤为必要,而大股东手中集中掌握的股权也提高了其监督管理层的能力(La Porta等,1999<sup>[9]</sup>)。高管纵向兼任是大股东深入公司内部,通过参与公司管理和相关决策来监督管理层行为的一种方式。

已有支持“监督效应”的经验证据表明,高管纵向兼任能加强公司治理,提升企业价值。例如,La Porta等(1999)<sup>[9]</sup>发现,在股东权益保护不尽完善的国家中,家族企业更倾向于委派家族成员担任高管职位;Arnoldi等(2013)<sup>[5]</sup>认为纵向兼任董事长有利于企业价值的提升;潘红波和韩芳芳(2016)<sup>[1]</sup>认为高管纵向兼任有助于提高企业会计信息质量,是大股东加强对上市公司管理者监督的一种有效手段;佟爱琴和李孟洁(2018)<sup>[4]</sup>发现,高管纵向兼任能够抑制管理层的短视行为,并为企业带来更多的资源,从而提高企业风险承担水平。此外,高管个人声誉与上市公司绩效还可能通过纵向兼任联结起来,制约大股东的掏空行为,变相实现“监督效应”。如姜付秀等(2017)<sup>[10]</sup>发现,控股家族退居“幕后”,即家族成员不再担任董事长以后,上市公司出现问题导致家族声誉受损的成本降低,反而会增加控股股东的资金占用、关联交易等私利行为,使公司的业绩表现更差。

上述研究表明,高管纵向兼任可以在一定程度上发挥“监督效应”。本文认为,高管纵向兼任加强大股东对上市公司的监督,有效降低股东与管理层之间的代理成本,因而可能对审计决策产生影响。首先,从审计投入的角度,如果高管纵向兼任发挥了“监督效应”,那么公司的内部控制水平相应提高。审计师在执行审计程序前,需先确定内部控制的可信赖程度,然后根据公司的内部控制水平确定实质性测试范围(杨肃昌和马亚红,2020<sup>[11]</sup>)。对于内部控制水平较高的公司,审计师会减少审计范围、执行更少的审计程序,因此审计师的审计投入会减少。其次,从审计收费的角度,Simunic(1980)<sup>[12]</sup>将审计收费分为两部分,固有费用和风险溢价。审计投入的减少会降低固有费用,高管纵向兼任的“监督效应”使得审计师对公司的治理环境的判断比较好,对公司存在重大错报风险的疑虑也会比较少,由此会收取更低的风险溢价,因此,审计师的审计收费会降低。再次,从审计意见的角度,根据现行审计准则,审计师有责任对财务报表整体是否存在由于舞弊或错误导致的重大错报做出合理保证。如果审计师认为高管纵向兼任具有“监督效应”,审计师对公司的风险判断会比较低,审计意见会相对乐观,审计师也会出具更多的标准无保留意见。综上,本文提出了假设H<sub>1a</sub>:

**H<sub>1a</sub>:** 如果“监督效应”成立,当上市公司存在高管纵向兼任时,审计师会减少审计投入、收取更低

的审计费用、更有可能出具标准无保留审计意见。

(二) 高管纵向兼任的“掏空效应”

在第二类代理问题中，主要的利益冲突方是大股东和中小股东。大股东可能通过金字塔结构、交叉持股等方式，使其在上市公司中的现金流权和控制权相偏离，进而通过相对较小的现金流权获得较大的控制权来控制上市公司（La Porta 等，1999<sup>[9]</sup>）。在这种情况下，大股东有动机利用自己的控制权来侵占中小股东利益。高管纵向兼任有利于大股东对上市公司的控制，也可能成为大股东侵占中小股东利益的一种工具。

部分文献认为高管纵向兼任发挥了“掏空效应”。例如，郑杲娉等（2014）<sup>[13]</sup>发现，存在高管纵向兼任的公司发生了更多与担保相关的关联交易，而且这些关联交易规模更大，信息披露质量更低，因此，高管纵向兼任显著降低了公司价值；闫珍丽等（2019）<sup>[2]</sup>发现，高管纵向兼任更可能是大股东掏空上市公司的一种手段，增加了关联交易和资金占用等利益攫取行为，显著降低了企业的创新水平；冉秋红和周宁慧（2018）<sup>[14]</sup>同样发现，存在高管纵向兼任的公司，其创新活力更低；Chen 和 Yang（2021）<sup>[6]</sup>发现，从投资者的角度来看，高管纵向兼任增加了第二类代理成本，增加了公司持有现金的风险，进而降低了公司现金的市场价值。

上述经验证据表明，高管纵向兼任还可能是大股东侵占中小股东利益的一个途径，可能会带来“掏空效应”，增加公司的第二类代理成本。本文认为，如果审计师认为高管纵向兼任具有“掏空效应”，那么其对这类公司出现重大错报和舞弊风险可能性的判断就比较大、对公司治理环境的判断比较差、对公司内部控制有效性的判断比较低，从而会增加对公司审计风险的判断。首先，从审计投入的角度，审计师会通过增加审计程序、扩大审计范围等方式来降低审计风险，增加审计投入。其次，从审计收费的角度，由于审计投入的增加，审计收费的固有费用会增加，审计师对风险的评估使其收取更多的风险溢价来规避审计失败或被诉风险，从而增加审计收费。再次，从审计意见的角度，审计师对高风险公司发表更加严格的审计意见，更有可能出具非标准无保留意见的审计报告，

向外界释放该公司财务报告不可靠的信号。综上所述，本文提出了与“监督效应”假设完全相反的“掏空效应”假设 H<sub>1b</sub>：

**H<sub>1b</sub>**：如果“掏空效应”成立，当上市公司存在高管纵向兼任时，审计师会增加审计投入、收取更高的审计费用、更有可能出具非标准无保留审计意见。

三、研究设计

(一) 样本与数据来源

本文的初始样本为 2004—2018 年中国 A 股上市公司。由于高管纵向兼任数据最早可获得的年份为 2004 年，因此，本文的样本区间从 2004 年开始。借鉴已有研究，进行如下样本处理：（1）剔除金融行业上市公司；（2）剔除已退市和 ST 公司；（3）剔除净资产为负的公司；（4）剔除研究数据有缺失的样本。最终分别得到 26 404、24 915 和 23 355 个公司-年度观测值<sup>①</sup>。在此基础上，本文对连续变量进行了 1% 和 99% 水平上的缩尾处理，以消除异常值对实证结果的影响。研究所用的高管兼任数据、财务与公司治理等数据均来自 CSMAR 数据库。

(二) 研究模型与变量定义

1. 高管纵向兼任与审计投入。

借鉴 Knechel 和 Payne（2001）<sup>[15]</sup>、翟胜宝等（2017）<sup>[16]</sup>、粟立钟等（2019）<sup>[17]</sup>的研究，本文采用 OLS 模型（1）检验高管纵向兼任对审计投入的影响：

$$\begin{aligned}
 AINVEST_{i,t} = & \alpha_0 + \alpha_1 AM_{i,t} + \alpha_2 SIZE_{i,t} + \alpha_3 LEV_{i,t} \\
 & + \alpha_4 ROA_{i,t} + \alpha_5 INV_{i,t} + \alpha_6 REC_{i,t} \\
 & + \alpha_7 ACH_{i,t} + \alpha_8 BIG10_{i,t} + \alpha_9 GROWTH_{i,t} \\
 & + \alpha_{10} CUR_{i,t} + \alpha_{11} OCF_{i,t} + \alpha_{12} AGE_{i,t} \\
 & + \alpha_{13} TOP1_{i,t} + \alpha_{14} ABSACC_{i,t} \\
 & + \sum IND + \sum YEAR + \varepsilon \quad (1)
 \end{aligned}$$

其中，下标 *i* 代表个体公司，*t* 代表年份。*AM* 代表高管纵向兼任哑变量，与现有文献（闫珍丽等，2019<sup>[2]</sup>；潘红波和韩芳芳，2016<sup>[1]</sup>）一致，本文将将其界定为上市公司董事长或总经理在控股股东或其实际控制人处任职。如果存在上述兼任现象，则虚拟变量 *AM* 取值为 1，否则为 0。*AINVEST* 代表审计投入，

① 因为模型的被解释变量和控制变量不同，因而样本量也存在差异，审计投入、审计收费、审计意见的样本量分别为 26 404、24 915 和 23 355。

参考 Mao 和 Yu (2015)<sup>[18]</sup>、翟胜宝等 (2017)<sup>[16]</sup> 的研究, 采用审计延迟来衡量审计投入。具体来说, 审计投入等于年末至次年审计报告签署日之间的天数取自然对数。如果  $H_{1a}$  成立, 预期  $\alpha_1$  的估计系数显著为负; 如果  $H_{1b}$  成立, 预期  $\alpha_1$  的估计系数显著为正。

2. 高管纵向兼任与审计收费。

参考 Hay 等 (2006)<sup>[19]</sup>、翟胜宝等 (2017)<sup>[16]</sup>、陈婧等 (2018)<sup>[20]</sup> 的研究, 本文采用 OLS 模型 (2) 检验高管纵向兼任对审计收费的影响:

$$\begin{aligned}
 AFEE_{i,t} = & \beta_0 + \beta_1 AM_{i,t} + \beta_2 SIZE_{i,t} + \beta_3 LEV_{i,t} + \beta_4 ROA_{i,t} \\
 & + \beta_5 AINVEST_{i,t} + \beta_6 INV_{i,t} + \beta_7 REC_{i,t} \\
 & + \beta_8 BIG10_{i,t} + \beta_9 GROWTH_{i,t} + \beta_{10} OCF_{i,t} \\
 & + \beta_{11} AOP_{i,t} + \beta_{12} EHOLD_{i,t} + \beta_{13} COMP_{i,t} \\
 & + \beta_{14} SEO_{i,t} + \beta_{15} MA_{i,t} + \beta_{16} MI_{i,t} \\
 & + \sum IND + \sum YEAR + \varepsilon
 \end{aligned} \quad (2)$$

其中,  $AFEE$  代表审计费用。借鉴伍利娜 (2003)<sup>[21]</sup> 和 O'Keefe 等 (1994)<sup>[22]</sup> 的做法, 用公司当年度收取的审计费用的自然对数来表示。如果  $H_{1a}$  成立, 预期

$\beta_1$  的估计系数显著为负; 如果  $H_{1b}$  成立, 预期  $\beta_1$  的估计系数显著为正。

3. 高管纵向兼任与审计意见。

参考翟胜宝等 (2017)<sup>[16]</sup>、王娟和潘秀丽 (2018)<sup>[23]</sup> 的研究, 本文采用 Logit 模型 (3) 检验高管纵向兼任对审计意见的影响:

$$\begin{aligned}
 AOP_{i,t} = & \gamma_0 + \gamma_1 AM_{i,t} + \gamma_2 SIZE_{i,t} + \gamma_3 LEV_{i,t} + \gamma_4 ROA_{i,t} \\
 & + \gamma_5 AFEE_{i,t} + \gamma_6 INV_{i,t} + \gamma_7 REC_{i,t} + \gamma_8 ACH_{i,t} \\
 & + \gamma_9 BIG10_{i,t} + \gamma_{10} GROWTH_{i,t} + \gamma_{11} CUR_{i,t} \\
 & + \gamma_{12} AGE_{i,t} + \gamma_{13} TOP1_{i,t} + \gamma_{14} ABSACC_{i,t} \\
 & + \gamma_{15} COMP_{i,t} + \gamma_{16} INDEP_{i,t} + \gamma_{17} Z_{i,t} \\
 & + \gamma_{18} AINVEST_{i,t} + \sum IND + \sum YEAR + \varepsilon
 \end{aligned} \quad (3)$$

其中,  $AOP$  代表审计意见。如果当年审计师出具了非标准无保留审计意见, 则  $AOP$  取值为 1, 否则为 0。如果  $H_{1a}$  成立, 预期  $\gamma_1$  的估计系数显著为负; 如果  $H_{1b}$  成立, 预期  $\gamma_1$  的估计系数显著为正。本文的变量定义见表 1。

表 1 变量定义表

变量类型	变量符号	变量名称	变量说明
解释变量	$AM$	高管纵向兼任	哑变量, 如果上市公司董事长或总经理在控股股东或实际控制人处任职则赋值为 1, 否则为 0
被解释变量	$AINVEST$	审计投入	以审计延迟代替, 等于年末至次年审计报告签署日之间的天数取自然对数
	$AFEE$	审计费用	公司审计收费的自然对数
	$AOP$	审计意见	虚拟变量, 如果当年审计师出具了非标准无保留审计意见则赋值为 1, 否则为 0
控制变量	$SIZE$	公司规模	公司期末资产总额的自然对数
	$LEV$	负债水平	公司期末负债总额除以期末资产总额
	$ROA$	盈利能力	公司净利润除以期末资产总额
	$INV$	存货占比	公司期末存货净额除以期末资产总额
	$REC$	应收账款占比	公司期末应收账款净额除以期末资产总额
	$CUR$	流动资产占比	公司期末流动资产除以期末资产总额
	$OCF$	经营现金流占比	公司经营现金流量除以期末资产总额
	$ABSACC$	应计利润占比	应计利润与期末资产总额的比值, 再取其绝对值; 应计利润等于净利润减去经营活动现金流量
	$Z$	Z 指数	公司破产风险指数, 来源于 Choice 数据库
	$GROWTH$	营业收入增长率	公司当年营业收入相对于上一年营业收入的增长率
	$SEO$	股票增发	哑变量, 如果公司增发了股票则赋值为 1, 否则为 0
	$MA$	是否发生并购	哑变量, 如果公司发生了并购则赋值为 1, 否则为 0
$AGE$	公司上市年数	公司上市年限加 1 取自然对数	
$TOP1$	股权集中度	第一大股东持股比例	

续前表

变量类型	变量符号	变量名称	变量说明
控制变量	COMP	高管薪酬	董事、监事、高管前三名薪酬的自然对数
	EHOLD	高管持股	高管持股数占公司总股数的比例
	INDEP	董事会独立性	独立董事人数除以董事会总人数
	ACH	会计师事务所变更	哑变量, 如果公司发生了会计师事务所变更取值为1, 否则为0
	BIG10	十大会计师事务所	哑变量, 参考中国注册会计师协会公布的业务收入排名取前十大, 如为前十大会计师事务所审计则为1, 否则为0
	MI	市场化指数	取自中国市场化指数(樊纲等, 2016 <sup>[24]</sup> )
	IND	行业	行业哑变量, 依据证监会2012年行业分类
	YEAR	年份	年份哑变量

#### 四、实证结果及分析

##### (一) 描述性统计

由表2主要变量描述性统计可知, 存在高管纵向兼任的样本约占全样本的48.0%。样本公司审计投入(AINVEST)的均值为4.491, 相当于样本公司审计时滞平均为89天, 25分位数和75分位数分别为4.394和4.700, 分别相当于81天和110天, 说明

审计师对不同公司的审计投入具有较大差异。审计费用(AFEE)均值为13.456, 换算成金额约为69.8万元, 25分位数和75分位数分别为13.017和13.816, 相当于约45万元和100万元, 表明不同公司的审计费用存在明显差异。审计意见(AOP)的均值为0.033, 说明约有3.3%的公司被出具非标准无保留意见的审计报告。本文其他变量的分布均在合理范围之内, 不再赘述。

表2 主要变量描述性统计

变量	均值	标准差	最小值	25分位数	中位数	75分位数	最大值
AM	0.480	0.500	0.000	0.000	0.000	1.000	1.000
AINVEST	4.491	0.271	3.332	4.394	4.543	4.700	4.771
AFEE	13.456	0.638	12.206	13.017	13.385	13.816	15.579
AOP	0.033	0.179	0.000	0.000	0.000	0.000	1.000
SIZE	21.881	1.274	19.378	20.951	21.721	22.615	25.823
LEV	0.438	0.208	0.050	0.272	0.437	0.598	0.889
ROA	0.039	0.056	-0.201	0.014	0.037	0.067	0.198
INV	0.155	0.143	0.000	0.062	0.120	0.199	0.725
REC	0.115	0.102	0.000	0.033	0.091	0.169	0.463
CUR	0.566	0.212	0.087	0.415	0.581	0.731	0.965
OCF	0.045	0.075	-0.189	0.005	0.044	0.088	0.255
ABSACC	0.057	0.058	0.001	0.018	0.040	0.075	0.326
Z	6.873	9.703	0.139	1.970	3.597	7.181	61.479
GROWTH	0.205	0.478	-0.597	-0.010	0.126	0.300	3.195
SEO	0.108	0.310	0.000	0.000	0.000	0.000	1.000
MA	0.710	0.454	0.000	0.000	1.000	1.000	1.000
AGE	1.982	0.888	0.000	1.386	2.197	2.708	3.219
TOP1	0.360	0.153	0.090	0.239	0.339	0.467	0.757
COMP	14.161	0.780	12.102	13.685	14.176	14.651	16.212
EHOLD	0.063	0.135	0.000	0.000	0.000	0.037	0.609

续前表

变量	均值	标准差	最小值	25 分位数	中位数	75 分位数	最大值
<i>INDEP</i>	0.369	0.052	0.286	0.333	0.333	0.400	0.571
<i>ACH</i>	0.083	0.276	0.000	0.000	0.000	0.000	1.000
<i>BIG10</i>	0.477	0.499	0.000	0.000	0.000	1.000	1.000
<i>MI</i>	7.826	1.783	2.920	6.570	8.030	9.300	10.270

注：变量定义参见表1，所有连续变量均进行1%和99%的缩尾处理。

(二) 多元回归结果分析

表3报告了高管纵向兼任对审计决策影响的多元回归分析结果。如表3所示，列(1)中，高管纵向兼任(*AM*)的估计系数为-0.011，且在5%的水平上显著为负，说明审计师显著降低了存在高管纵向兼任公司的审计投入；列(2)中，高管纵向兼任(*AM*)的估计系数为-0.022，且在5%的水平上显著为负，说明在控制了其他因素以后，存在高管纵向兼任公司的审计费用仍然显著较低；列(3)中，高管纵向兼任(*AM*)的估计系数为-0.217，且在10%的显著性水平上显著为负，说明存在高管纵向兼任公司获得了更多的标准无保留审计意见，被出具非标准无保留意见的可能性降低了19.5%（由未报告的odds ratio计算得来）。以上研究结果说明，审计师对存在高管纵向兼任公司降低了审计投入、收取了更低的审计费用、更有可能出具标准无保留意见的审计报告。因此，从审计师的角度来看，高管纵向兼任更多地发挥了“监督效应”，而不是“掏空效应”，支持了假设H<sub>1a</sub>。

表3 高管纵向兼任与审计决策

变量	多元线性回归		Logistic 回归
	<i>AINVEST</i>	<i>AFEE</i>	<i>AOP</i>
	(1)	(2)	(3)
<i>AM</i>	-0.011 ** (-2.205)	-0.022 ** (-2.011)	-0.217 * (-1.887)
<i>SIZE</i>	0.019 *** (7.067)	0.344 *** (42.080)	-0.663 *** (-7.052)
<i>LEV</i>	-0.054 *** (-3.221)	0.014 (0.380)	3.363 *** (7.838)
<i>ROA</i>	-0.761 *** (-15.633)	-0.468 *** (-5.334)	-9.700 *** (-9.638)
<i>AINVEST</i>		0.093 *** (6.290)	3.247 *** (7.328)
<i>INV</i>	-0.046 (-1.643)	-0.033 (-0.642)	-3.512 *** (-4.771)
<i>REC</i>	0.073 ** (2.399)	0.126 ** (2.042)	-1.636 ** (-1.963)
<i>ACH</i>	0.017 *** (3.082)		0.467 *** (3.440)

续前表

变量	多元线性回归		Logistic 回归
	<i>AINVEST</i>	<i>AFEE</i>	<i>AOP</i>
	(1)	(2)	(3)
<i>BIG10</i>	0.004 (0.899)	0.109 *** (9.700)	-0.046 (-0.379)
<i>GROWTH</i>	-0.023 *** (-5.461)	0.015 ** (2.377)	0.066 (0.637)
<i>CUR</i>	0.053 *** (2.802)		1.357 *** (2.756)
<i>OCF</i>	-0.129 *** (-4.198)	0.227 *** (4.196)	
<i>AGE</i>	-0.006 (-1.556)		0.559 *** (4.934)
<i>TOP1</i>	-0.034 * (-1.738)		-1.016 ** (-2.170)
<i>ABSACC</i>	0.082 ** (2.291)		0.826 (0.977)
<i>AFEE</i>			0.708 *** (4.474)
<i>AOP</i>		0.156 *** (7.106)	
<i>EHOLD</i>		-0.030 (-0.776)	
<i>COMP</i>		0.061 *** (5.834)	-0.339 *** (-2.811)
<i>SEO</i>		0.022 *** (2.780)	
<i>MA</i>		0.055 *** (7.179)	
<i>MI</i>		0.045 *** (11.053)	
<i>INDEP</i>			-0.869 (-0.830)
<i>Z</i>			0.012 (1.294)
<i>Constant</i>	4.096 *** (72.815)	4.278 *** (20.873)	-9.420 *** (-3.225)
<i>IND/YEAR</i>	YES	YES	YES
<i>N</i>	26 404	24 915	23 355
Adj/Pseudo R <sup>2</sup>	0.098	0.616	0.286

注：\*\*\*、\*\*、\*分别代表1%、5%、10%的显著性水平；括号内为t值和z值；标准误经过公司水平的cluster处理。下同。



### 五、稳健性检验

#### (一) 内生性检验

本文研究高管纵向兼任与审计决策之间的因果关系，这一关系可能会受到内生性问题的干扰。例如，存在高管纵向兼任的公司更倾向于聘请特定类型的审计师，这类审计师更愿意减少审计投入、降低审计收费、出具更多标准无保留意见，即存在自选择问题。因此，本文采用一系列检验来缓解潜在的内生性问题。

#### 1. 倾向得分匹配模型 (PSM)。

为保证控制变量在处理组与控制组之间不存在显著差异，本文选取模型 (1)~模型 (3) 的所有控制变量作为匹配的协变量，采用一对一有放回的半径匹配方法 ( $r=0.01$ )，生成配对样本。处理组和控制组样本不存在显著差异，配对结果满足平衡性检验的要求。采用 PSM 配对样本的回归结果如表 4 所示。可以看出，在采用倾向评分匹配样本进行回归时，AM 在各模型中依然显著为负，与表 3 的结果较为一致。

表 4 PSM 检验

变量	多元线性回归		Logistic 回归
	AINVEST	AFEE	AINVEST
	(1)	(2)	(3)
AM	-0.010** (-1.983)	-0.027** (-2.324)	-0.220* (-1.851)
Constant	4.028*** (36.984)	4.248*** (20.069)	-10.361*** (-3.452)
Controls	YES	YES	YES
IND/YEAR	YES	YES	YES
N	23 612	22 391	22 318
Adj/Pseudo R <sup>2</sup>	0.090	0.619	0.285

#### 2. Heckman 两阶段模型。

本文采用 Heckman 两阶段模型控制自选择问题。采用同年度、同省份、同行业的高管纵向兼任的平均值 (IV) 作为工具变量。同行业同年度的公司在经营环境和公司治理特征上较为相似，其他公司的高管纵向兼任情况会对本公司高管聘任决策产生影响，但是行业内其他公司的高管纵向兼任情况不会影响到本公司的审计决策，符合工具变量的相关性和外生性要求。在第二阶段中，将逆米尔斯比率 (IMR) 作为一个控

制变量加入模型 (1)~模型 (3) 中重新进行回归。

Heckman 两阶段回归结果如表 5 所示。在控制了潜在的自选择问题以后，AM 的估计系数依然全部显著为负，说明审计师对存在高管纵向兼任的公司减少了审计投入、收取了更低的审计费用、出具了更多的标准无保留审计意见。上述研究结果有力地支持了假设 H<sub>1a</sub>，表明本文的研究结论没有受到潜在的自选择问题的影响。

表 5 Heckman 两阶段回归

变量	(1)	(2)	(3)	(4)
	AM	AINVEST	AFEE	AOP
AM		-0.010** (-2.038)	-0.024** (-2.088)	-0.229* (-1.945)
IV	0.587*** (8.321)			
IMR		0.061* (1.933)	0.010 (0.295)	-0.057 (-0.085)
Controls	YES	YES	YES	YES
IND/YEAR	YES	YES	YES	YES
N	27 678	25 299	23 829	22 479
Adj/PseudoR <sup>2</sup>	0.066	0.098	0.615	0.284

#### 3. 工具变量法。

为了解决潜在的遗漏变量和反向因果问题，本文采用工具变量法进行回归。在利用工具变量控制潜在的内生性问题后，审计师对存在纵向高管兼任的公司依旧减少了审计投入，收取了更低的审计费用，但是对审计意见的回归结果不显著。上述结果总体说明高管纵向兼任能够影响审计决策。限于篇幅，该结果未予列示。

#### (二) 其他稳健性检验

为进一步保证回归结果的可靠性，本文尝试改变审计决策的度量方法以及改变模型的估计方法。首先，参考翟胜宝等 (2017)<sup>[16]</sup> 的做法，采用未经对数化处理的审计时滞、异常审计费用和审计意见类型序数<sup>①</sup>重新对主模型进行了回归；其次，参考 Petersen (2009)<sup>[25]</sup> 的研究，在公司和年份层面对标准误进行了双向聚类调整；再次，将模型 (3) 由 Logit 回归调整为 Probit 回归。上述结果显示 AM 的估计系数

① 审计意见类型序数取值规则：标准无保留意见取值为 0，带强调事项段的无保留意见取值为 1，保留意见、保留意见加事项段或强调段取值为 2，无法发表意见取值为 3。采用定序 Logit 回归方法。

依然全部显著为负，本文研究结论依旧稳健。限于篇幅，其他稳健性检验结果未予列示。

### 六、机制检验和拓展性分析

#### (一) 影响机制

本文的研究结果表明，从审计师的视角看，高管纵向兼任更多地发挥了“监督效应”，因此，审计师减少了审计投入、降低了审计收费、更有可能出具标准无保留意见的审计报告。那么，高管纵向兼任是通过何种渠道对审计决策产生影响的？下面，本文分别从高管纵向兼任对内部控制和代理成本的影响两个角度，对其影响机制进行了检验。

1. 内部控制在高管纵向兼任与审计决策关系中的作用分析。

已有文献表明，健全的内部控制在减少企业的控制风险，缓解公司与审计师之间的信息不对称。审计师对内部控制健全的公司降低了审计收费（牟韶红等，2014<sup>[26]</sup>）、更有可能出具标准无保留审计意见（徐玉霞和王冲，2012<sup>[27]</sup>）。高管纵向兼任是大股东深入公司内部约束管理层的一种手段，有助于提升公司的内部治理水平，最终影响到了审计决策。为了验证这一影响路径，借鉴 Baron 和 Kenny（1986）<sup>[28]</sup>、蒋德权等（2018）<sup>[29]</sup>的中介效应检验方法，本文的具体模型如下（以审计投入为例）：

$$AINVEST_{i,t} = \beta_0 + \beta_1 AM_{i,t} + \sum CONTROLS_{i,t}$$

$$+ \sum IND + \sum YEAR + \varepsilon \quad (\text{Path A})$$

$$IC_{i,t} = \alpha_0 + \alpha_1 AM_{i,t} + \alpha_2 SIZE_{i,t} + \alpha_3 LEV_{i,t} + \alpha_4 ROA_{i,t}$$

$$+ \alpha_5 OCF_{i,t} + \alpha_6 AGE_{i,t} + \alpha_7 DUAL_{i,t}$$

$$+ \alpha_8 INDEP_{i,t} + \sum IND + \sum YEAR + \varepsilon$$

(Path B)

$$AINVEST_{i,t} = \mu_0 + \mu_1 AM_{i,t} + \mu_2 IC_{i,t} + \sum CONTROLS_{i,t}$$

$$+ \sum IND + \sum YEAR + \varepsilon \quad (\text{Path C})$$

其中，*IC* 表示内部控制水平，采取内部控制指数度量，数据来源于迪博数据库。路径模型 Path B 的设定及控制变量选取参考 Stanley（2011）<sup>[30]</sup> 的研究，具体包括公司规模（*SIZE*）、负债水平（*LEV*）、盈利能力（*ROA*）、经营现金流占比（*OCF*）、公司年龄（*AGE*）、两职合一（*DUAL*）和董事会独立性（*INDEP*）。Path A 和 Path C 的控制变量与前文一致。

Path A 的检验结果已经报告在表 3 中，表 6 报告了 Path B 和 Path C 的检验结果。可以看出，在 Path B（中介因子检验）中，*AM* 的系数显著为正，表明高管纵向兼任显著提升了内部控制水平，从而影响到审计决策；在 Path C（包含中介因子）中，控制内部控制水平（*IC*）以后，*AM* 的系数在审计投入、审计费用、审计意见中均显著为负，同时 Sobel *Z* 检验均在 1% 的水平上显著。中介因子内部控制水平均显著为负。以上结果说明内部控制水平在高管纵向兼任和审计决策的关系中发挥了部分中介作用。

表 6 内部控制中介效应检验结果

变量	(1)	(2)	(3)	(4)
	Path B	Path C		
	<i>IC</i>	<i>AINVEST</i>	<i>AFEE</i>	<i>AOP</i>
<i>AM</i>	0.005*** (2.632)	-0.011** (-2.185)	-0.023** (-2.059)	-0.203** (-2.009)
<i>IC</i>		-0.086*** (-5.514)	-0.068* (-1.916)	-5.212*** (-24.710)
<i>Controls</i>	YES	YES	YES	YES
<i>IND/YEAR</i>	YES	YES	YES	YES
<i>N</i>	27 193	26 326	24 572	23 300
Adj/Pseudo <i>R</i> <sup>2</sup>	0.261	0.099	0.615	0.382
Sobel <i>Z</i>	—	-7.453***	-4.57***	-7.001***

2. 代理成本在高管纵向兼任与审计决策关系中的作用分析。

高管纵向兼任影响审计决策的第二个可能路径是

影响了公司的代理成本。高管纵向兼任能够强化控股股东对上市公司的控制力，有效缓解第一类代理问题；但是，控股股东对上市公司控制力的增加，也可

能加剧其对中小股东的利益侵占行为，从而增加大小股东之间的第二类代理问题。为此，本文检验了代理成本在高管纵向兼任与审计决策中的作用。借鉴戴亦一等（2016）<sup>[31]</sup>的研究，本文用经营费用率，即管理费用与销售费用之和除以营业收入来衡量股东与管理

者之间的代理成本（第一类代理成本，*AGCOST*）；借鉴胡茜茜等（2018）<sup>[32]</sup>的研究，采用其他应收款占总资产的比重来衡量大股东的利益侵占（第二类代理成本，*TUNNEL*）。

表 7 代理成本中介效应检验结果

第一组：第一类代理成本				
变量	(1)	(2)	(3)	(4)
	Path B	Path C		
	<i>AGCOST</i>	<i>AINVEST</i>	<i>AFEE</i>	<i>AOP</i>
<i>AM</i>	-0.012 *** (-3.859)	-0.010 ** (-2.019)	-0.020 * (-1.770)	-0.211 * (-1.806)
<i>AGCOST</i>		0.082 *** (3.791)	0.342 *** (7.106)	1.781 *** (3.878)
<i>Controls</i>	YES	YES	YES	YES
<i>IND/YEAR</i>	YES	YES	YES	YES
<i>N</i>	29 057	26 345	24 685	23 302
Adj/Pseudo <i>R</i> <sup>2</sup>	0.241	0.099	0.620	0.293
Sobel <i>Z</i>	—	-5.909 ***	-6.423 ***	-5.953 ***
第二组：第二类代理成本				
变量	(1)	(2)	(3)	(4)
	Path B	Path C		
	<i>TUNNEL</i>	<i>AINVEST</i>	<i>AFEE</i>	<i>AOP</i>
<i>AM</i>	-0.000 (-0.256)	-0.011 ** (-2.305)	-0.022 ** (-2.014)	-0.239 ** (-2.045)
<i>TUNNEL</i>		0.642 *** (7.946)	1.314 *** (8.736)	4.979 *** (3.793)
<i>Controls</i>	YES	YES	YES	YES
<i>IND/YEAR</i>	YES	YES	YES	YES
<i>N</i>	29 366	26 397	24 909	23 348
Adj/Pseudo <i>R</i> <sup>2</sup>	0.162	0.103	0.620	0.291

表 7 的第一组和第二组分别报告第一类和第二类代理成本中介效应检验结果。在第一组的 Path B 中，高管纵向兼任显著降低了经营费用率，表明高管纵向兼任通过降低第一类代理成本，影响了审计决策；在第一组的 Path C 中，控制了经营费用率以后，*AM* 的估计系数在审计投入、审计费用和审计意见中均显著为负，同时 Sobel *Z* 值均在 1% 的显著性水平上显著。这说明第一类代理成本在高管纵向兼任和审计决策的关系中发挥了部分中介作用。相反，在第二组的 Path B 中，*AM* 的估计系数在统计上不显著，中介效

应检验终止，说明高管纵向兼任并未显著影响第二类代理成本。上述对两类代理成本的测试说明，高管纵向兼任发挥“监督作用”的路径是降低了第一类代理成本，但是并没有影响第二类代理成本。

(二) 拓展性分析：大股东控制力的影响

上文研究结果表明，从审计师的视角来看，高管纵向兼任发挥了监督作用，降低了审计风险。那么，不同类型公司的高管纵向兼任效果是否存在差异？本部分探讨了大股东控制力对于高管纵向兼任和审计决策关系的影响。

如果高管纵向兼任能够发挥有效的治理作用，那么，大股东对上市公司的控制力越强，通过高管纵向兼任来监督上市公司的必要性就越小。这是因为，如果大股东对上市公司的控制能力较强，能够通过管理层聘任和薪酬决策等来与管理层建立更加紧密的联系，达到降低第一类代理成本的目的。由于大股东在上市公司的董事会席位与持股比例相关，即持股比例越高，董事会席位越多，对上市公司的控制力越强。因此，我们预期，高管纵向兼任对审计决策的影响在

大股东持股比例较低的公司更加显著。

本文根据大股东持股比例的行业-年度中位数将样本公司分为大股东控制权较强和较弱两组重新回归，表8报告了实证结果。在大股东控制权较弱的子样本中，AM的估计系数分别在1%、5%和5%的显著性水平上显著为负；但是，在大股东控制权较强的样本中，AM均不显著。这说明，从审计师视角来看，高管纵向兼任的“监督作用”只存在于大股东控制能力较弱的样本中。

表8 高管纵向兼任、大股东控制力与审计决策

变量	控制权弱	控制权强	控制权弱	控制权强	控制权弱	控制权强
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>AINVEST</i>	<i>AINVEST</i>	<i>AFEE</i>	<i>AFEE</i>	<i>AOP</i>	<i>AOP</i>
<i>AM</i>	-0.025*** (-3.650)	0.003 (0.517)	-0.031** (-1.971)	-0.011 (-0.756)	-0.298** (-2.057)	-0.091 (-0.454)
<i>Controls</i>	YES	YES	YES	YES	YES	YES
<i>IND/YEAR</i>	YES	YES	YES	YES	YES	YES
<i>N</i>	13 552	12 852	12 918	11 997	12 036	11 248
Adj/Pseudo <i>R</i> <sup>2</sup>	0.101	0.107	0.597	0.636	0.308	0.237

(三) 拓展性分析：区分纵向兼任高管的职位特征

参考前人研究，本文将上市公司董事长或总经理在控股股东或其实际控制人处任职界定为纵向兼任高管。上市公司的董事长与总经理在职责方面存在较大差异，董事长是董事会的领导者，负责召集董事会会议、提名高级管理人员、审查经理层的发展计划；总经理是管理层的最高代表，负责制定和执行上市公司的日常经营计划及投资决策。那么，董事长兼任和总经理兼任对审计决策的影响是否存在差异呢？

为探究这一问题，本文将纵向兼任高管细分为董事长纵向兼任 (*AM\_Chairman*) 与总经理纵向兼任 (*AM\_CEO*) 两种情形，并分别考察其对审计决策的影响。从表9结果可以看出，当上市公司存在董事长兼任时，审计师减少了审计投入，降低了审计收费，并且出具了更多的标准无保留审计意见；但是，总经理纵向兼任对审计决策并无显著影响。上述结果表明，董事长纵向兼任与总经理纵向兼任的效果并不一致，在中国上市公司的实际情况下，纵向兼任董事长的权力更大，更能有效发挥监督职能。

表9 董事长纵向兼任 vs 总经理纵向兼任

变量	(1)	(2)	(3)	(4)	(5)	(6)
	<i>AINVEST</i>	<i>AFEE</i>	<i>AOP</i>	<i>AINVEST</i>	<i>AFEE</i>	<i>AOP</i>
<i>AM_Chairman</i>	-0.012** (-2.328)	-0.024** (-2.173)	-0.292** (-2.469)			
<i>AM_CEO</i>				0.005 (0.776)	0.019 (1.485)	0.082 (0.618)
<i>Controls</i>	YES	YES	YES	YES	YES	YES
<i>IND/YEAR</i>	YES	YES	YES	YES	YES	YES
<i>N</i>	26 522	24 915	23 355	26 522	24 915	23 355
Adj/Pseudo <i>R</i> <sup>2</sup>	0.090	0.616	0.287	0.090	0.616	0.286

(四) 拓展性分析：区分产权性质

进一步地，本文探讨产权性质在高管纵向兼任对审计决策的影响中的调节效应。国有企业和民营企业由于经营目标和管理体制的差异，其代理问题的表现形式并不相同。相比于民营企业，由于上市公司财富转移并不会增加国有企业的控股股东（包括国资委、地方政府等）的私人收益，因此国有企业的大股东通过违规担保、关联交易等行为侵占中小股东利益的掏空动机更弱，第二类代理问题并不严重；相反，国有企业的经理人权限较大，由于缺乏有效的监督和制约机制，由内部人控制引发的第一类代理问题更加严重。高管纵向兼任作为大股东强化监督的一种手段，能够有效缓解第一类代理问题，并降低审计风险。因

此，本文预期高管纵向兼任对审计决策的影响在国有企业中更加显著。

本文根据产权性质将样本分为国有企业和民营企业两组，重新回归，结果报告在表 10 中。回归结果表明，审计师对存在高管纵向兼任的国有企业，降低了审计投入，并收取了更低的审计费用；相反，此效应应在民营企业中并不显著。在对审计意见的回归中，AM 的估计系数在国有企业和民营企业两个子样本中均不显著。上述结果表明，对于存在高管纵向兼任的公司，审计决策在国有企业与非国有企业存在差异，审计师对国有企业降低了审计投入，收取了更低的审计费用，但在出具审计意见方面未表现出显著的差异。

表 10 高管纵向兼任、产权性质与审计决策

变量	国有企业	民营企业	国有企业	民营企业	国有企业	民营企业
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>AINVEST</i>	<i>AINVEST</i>	<i>AFEE</i>	<i>AFEE</i>	<i>AOP</i>	<i>AOP</i>
<i>AM</i>	-0.021 *** (-2.926)	-0.002 (-0.313)	-0.031 * (-1.733)	-0.005 (-0.335)	-0.246 (-1.346)	-0.230 (-1.549)
<i>Controls</i>	YES	YES	YES	YES	YES	YES
<i>IND/YEAR</i>	YES	YES	YES	YES	YES	YES
<i>N</i>	11 096	15 308	9 862	15 053	9 305	13 700
Adj/Pseudo <i>R</i> <sup>2</sup>	0.088	0.105	0.661	0.587	0.309	0.290

七、研究结论与讨论

高管纵向兼任是我国资本市场上的一个普遍现象，高管纵向兼任的经济后果在理论界和实务界都备受关注，但目前并未得出一致的结论。本文基于 2004—2018 年间中国 A 股上市公司数据，实证检验高管纵向兼任对审计决策的影响。研究结果发现：从审计决策的角度来看，高管纵向兼任更多地发挥了“监督效应”。对于存在高管纵向兼任的上市公司，审计师会减少审计投入、降低审计收费，更有可能出具标准无保留意见的审计报告。机制检验表明，高管纵向兼任能够提高上市公司内部控制水平，降低股东与管理层之间的代理成本。此外，高管纵向兼任对审计决策的影响只在大股东控制力较弱的样本和国有企业样本中显著。

本文从上市公司专业信息鉴证者的视角，验证了

高管纵向兼任的积极作用。这说明，从监管的角度来讲，可能没有必要限制高管的纵向兼任。特别是对于大股东控制力较弱的公司，更应该通过高管纵向兼任来有效监督管理层，在不增加第二类代理成本的基础上，降低第一类代理成本。

本文也具有一定的局限性。一方面，本文仅从代理成本的角度思考高管纵向兼任对审计决策的影响，高管纵向兼任通过其他渠道对审计决策的影响也不容忽视；另一方面，与其他手段相比，大股东通过高管纵向兼任来监督上市公司的有效性本文也未探讨。未来，针对高管纵向兼任的经济后果，还有诸多有趣的话题有待深入探讨。例如，高管纵向兼任会给上市公司带来更多的资源，减少公司的融资约束吗？存在高管纵向兼任的公司能够提高公司业绩承诺的可靠性吗？这些都是未来值得研究的方向。

## 参考文献

- [1] 潘红波, 韩芳芳. 纵向兼任高管, 产权性质与会计信息质量 [J]. 会计研究, 2016 (7): 19-26.
- [2] 闫珍丽, 梁上坤, 袁淳. 高管纵向兼任、制度环境与企业创新 [J]. 经济管理, 2019 (10): 90-107.
- [3] 韩金红, 余珍. 纵向兼任高管与企业投资效率——基于“监督效应”和“掏空效应”分析 [J]. 审计与经济研究, 2019 (4): 66-74.
- [4] 佟爱琴, 李孟洁. 产权性质、纵向兼任高管与企业风险承担 [J]. 科学学与科学技术管理, 2018 (1): 118-126.
- [5] Arnoldi J, Chen X, Na C. Vertical Interlocks of Executives and Performance of Chinese State Owned Firms [J]. Academy of Management Annual Meeting Proceedings, 2013 (1): 10967-10967.
- [6] Chen X, Yang C. Vertical Interlock and the Value of Cash Holdings [J]. Accounting & Finance, 2021, 61 (1): 561-593.
- [7] 郑志刚, 孙娟娟. 我国上市公司治理发展历史与现状评估 [J]. 金融研究, 2009 (10): 118-132.
- [8] Jensen M C, Meckling W H. Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure [J]. Journal of Financial Economics, 1976, 3 (4): 305-360.
- [9] La Porta R F, Lopez-de-Silanes, Shleifer A. Corporate Ownership around the World [J]. The Journal of Finance, 1999, 54 (2): 471-517.
- [10] 姜付秀, 郑晓佳, 蔡文婧. 控股家族的“垂帘听政”与公司财务决策 [J]. 管理世界, 2017 (3): 125-145.
- [11] 杨肃昌, 马亚红. 对外直接投资、控制风险与上市公司审计费用 [J]. 审计研究, 2020 (2): 76-86.
- [12] Simunic D A. The Pricing of Audit Services: Theory and Evidence [J]. Journal of Accounting Research, 1980, 18 (1): 161-190.
- [13] 郑杲娉, 薛健, 陈晓. 兼任高管与公司价值: 来自中国的经验数据 [J]. 会计研究, 2014 (11): 24-29.
- [14] 冉秋红, 周宁慧. 纵向兼任高管, 机构投资者持股与智力资本价值创造 [J]. 软科学, 2018 (12): 50-54.
- [15] Knechel W R, Payne J L. Additional Evidence on Audit Report Lag [J]. Auditing: A Journal of Practice & Theory, 2001, 20 (1): 137-146.
- [16] 翟胜宝, 许浩然, 刘耀淞, 唐玮. 控股股东股权质押与审计师风险应对 [J]. 管理世界, 2017 (10): 51-65.
- [17] 粟立钟, 张润达, 鲁睿. 股权激励与外部审计师风险应对 [J]. 审计研究, 2019 (3): 120-128.
- [18] Mao M Q, Yu Y. Analysts' Cash Flow Forecasts, Audit Effort, and Audit Opinions on Internal Control [J]. Journal of Business Finance & Accounting, 2015, 42 (5/6): 635-664.
- [19] Hay D C, Knechel W R, Wong N. Audit Fees: A Meta-analysis of the Effect of Supply and Demand Attributes [J]. Contemporary Accounting Research, 2006, 23 (1): 141-191.
- [20] 陈婧, 张金丹, 方军雄. 公司债务违约风险影响审计收费吗 [J]. 财贸经济, 2018 (5): 71-87.
- [21] 伍利娜. 盈余管理对审计费用影响分析——来自中国上市公司首次审计费用披露的证据 [J]. 会计研究, 2003 (12): 39-44.
- [22] O'Keefe T B, Simunic D A, Stein M T. The Production of Audit Services: Evidence from a Major Public Accounting Firm [J]. Journal of Accounting Research, 1994, 32 (2): 241-261.
- [23] 王娟, 潘秀丽. 慈善捐赠对审计意见的影响——基于慈善捐赠利己动机视角的实证分析 [J]. 审计研究, 2018 (3): 14.
- [24] 樊纲, 王小鲁, 余静文. 中国分省份市场化指数报告 (2016) [M]. 北京: 社会科学文献出版社, 2017.
- [25] Petersen M A. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches [J]. The Review of Financial Studies, 2009, 22 (1): 435-480.
- [26] 牟韶红, 李启航, 于林平. 内部控制、高管权力与审计费用——基于2009—2012年非金融上市公司数据的经验研究 [J]. 审计与经济研究, 2014 (4): 40-49.
- [27] 徐玉霞, 王冲. 风险导向审计、内部控制与审计师行为——基于我国上市公司的实证检验 [J]. 经济评论, 2012 (5): 123-133.
- [28] Baron R M, Kenny D A. The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations [J]. Journal of Personality and Social Psychology, 1986, 51 (6): 1173-1182.
- [29] 蒋德权, 姚振晔, 陈冬华. 财务总监地位与企业股价崩盘风险 [J]. 管理世界, 2018 (3): 153-166.
- [30] Stanley J D. Is the Audit Fee Disclosure a Leading Indicator of Clients' Business Risk? [J]. Auditing: A Journal of Practice & Theory, 2011, 30 (3): 157-179.
- [31] 戴亦一, 肖金利, 潘越. “乡音”能否降低公司代理成本? ——基于方言视角的研究 [J]. 经济研究, 2016 (12): 147-160.
- [32] 胡茜茜, 朱永祥, 杜勇. 网络环境下中小股东的治理效应研究——基于代理成本视角 [J]. 财经研究, 2018 (5): 109-120.

(学术顾问: 梁上坤, 责任编辑: 张安平)

# 信息披露违规处罚、连锁董事公司分析师预测与 分析师声誉：基于董事网络的溢出效应

Regulatory Penalty for Information Disclosure, Analysts' Forecast of  
Interlocked Director Firms and Analyst Reputation:  
Spillover Effect Based on Director Network

丁鑫 陆阳 杨忠海

DIN Xin LU Yang YANG Zhong-hai

**[摘要]** 本文基于中国证监会对上市公司监管处罚公告的数据统计, 研究信息披露违规处罚公告对连锁董事公司分析师预测的影响以及分析师声誉的调节效应。实证研究发现, 公司发生信息披露违规处罚后, 分析师对其连锁董事公司的盈余预测准确性显著降低。明星分析师发挥负面的调节效应, 与非明星分析师相比, 明星分析师对连锁董事公司预测准确性更低。进一步研究发现, 信息披露违规处罚引发资本市场对连锁董事公司的负面反应; 分析师对连锁董事公司盈余预测的正向偏差显著增大, 与非明星分析师相比, 明星分析师对连锁董事公司盈余预测的正向偏差更大; 分析师包括明星分析师与非明星分析师对连锁董事公司盈余预测的负向偏差无显著变化。本文不仅丰富监管处罚的溢出效应研究, 对于防范资本市场中潜在的金融风险, 提高分析师声誉激励机制的有效性具有重要启示。

**[关键词]** 信息披露违规处罚 连锁董事 溢出效应 分析师预测 明星分析师

**[中图分类号]** F832.5 **[文献标识码]** A **[文章编号]** 1000-1549 (2022) 11-0065-12

**Abstract:** Based on China Securities Regulatory Commission statistics on penalties for Chinese listed companies, this paper studies the impact of information disclosure violation penalties on analysts' forecast accuracy of interlocked director firms, and the moderating effect of analysts' reputation. The empirical study finds that the accuracy of analysts' earnings forecast of the interlocked director firms decreases significantly after the company is punished for information disclosure violation penalty. Moreover, star analysts have a negative moderating effect on the relationship between the information disclosure violation penalty and the accuracy of analysts' forecasts. Compared with non-star analysts, star analysts have lower prediction accuracy for interlocked director firms. Further study found that the information disclosure violation penalty caused the negative reaction of capital market to the interlocked director firms. The positive bias of analysts' earnings forecast for interlocked director firms increases significantly, and the positive bias of star analysts' earnings forecast for interlocked director firms is greater than that of non-star analysts. Analysts, including star analysts, have no significant influence on the negative bias of earnings forecast of interlocked director firms. This paper not only enriches the research on the spillover effect of regulatory penalties, but also has important implications for preventing potential financial risks in the capital market and improving the effectiveness of analysts' reputation incentive mechanism.

**Key words:** Information disclosure violation penalty Interlocked director Spillover effect Analysts' forecast Star analyst

**[收稿日期]** 2021-04-20

**[作者简介]** 丁鑫, 女, 1984年3月生, 东北林业大学经济管理学院副教授, 研究方向为公司财务决策、公司治理与财务信息披露研究; 陆阳, 女, 1991年12月生, 哈尔滨商业大学会计学院博士研究生, 研究方向为资本市场与会计信息、高管社会网络; 杨忠海, 男, 1972年5月生, 哈尔滨工程大学经济管理学院教授, 博士生导师, 研究方向为资本市场与会计信息、高管社会网络。本文通讯作者为杨忠海, 联系方式为 yzhcdk@126.com。

**[基金项目]** 中央高校基本科研业务费项目“宏观经济政策与会计信息质量及披露违规处罚溢出效应研究”(项目编号: 307CFJ1900905)。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

监管处罚不仅会对受罚者产生影响,也会对受罚者之外的与受罚者相关联的组织或个人产生影响,即惩罚存在溢出效应(Trevino, 1992<sup>[1]</sup>; Beatty等, 2013<sup>[2]</sup>)。本文关注的是监管处罚的溢出效应。对于监管处罚溢出效应的作用主体,已有研究大多是关注与违规处罚公司关联的上市公司(Brown和Zhou, 2015<sup>[3]</sup>; Brown等, 2018<sup>[4]</sup>)。本文拓展了监管处罚溢出效应的作用主体,将研究视角聚焦于监管处罚对连锁董事公司分析师预测行为的影响。

公司受到违规被处罚后,因投资者的归因偏差与归因搜索(Ross, 1977<sup>[5]</sup>; Weiner, 1985<sup>[6]</sup>),导致连锁董事公司股价与声誉受到负面影响(Kang, 2008<sup>[7]</sup>)。连锁董事公司股价的下跌会影响其经营与融资成本,从而使连锁董事公司盈余的不确定性增加。同时,基于董事网络间“趋同效应”(Brown和Higgins, 2005<sup>[8]</sup>; 陈运森和郑登津, 2017<sup>[9]</sup>),处罚公司的连锁董事公司也很可能存在信息披露违规行为,连锁董事公司的信息不确信性程度进一步加剧。信息不确定性影响分析师预测偏差(Zhang, 2006<sup>[10]</sup>)。信息不确定程度越大,分析师对信息的反应越不完全,分析师预测误差越大(Gu和Wu, 2003<sup>[11]</sup>; Zhang, 2006<sup>[10]</sup>)。因此,信息披露违规处罚发生后,由于连锁董事公司的信息不确信性程度加剧,分析师对连锁董事公司盈余预测准确性可能受到影响,本文予以实证检验。

进一步地,根据分析师的声誉有效性假说,明星分析师具有优越的人力资本,能够提供更多公司层面的特质信息,盈余预测的准确性更高(Xu等, 2013<sup>[12]</sup>; 张然等, 2017<sup>[13]</sup>)。然而,根据迎合理论和利益结盟假说,经验研究发现,声誉越高的分析师,越倾向于采取“跟风”策略(蔡庆丰等, 2011<sup>[14]</sup>; 董大勇等, 2012<sup>[15]</sup>),盈余预测的准确性越低。在中国资本市场中,分析师的声誉效应并没有战胜利益结盟动机,分析师与上市公司利益结盟,盈余预测的乐观度提升,准确性进一步下降(唐松莲和陈伟, 2017<sup>[16]</sup>)。因此,本文试图探索信息披露违规处罚后,明星分析师是否能够发挥积极的声誉效应,实证考察明星分析师对信息披露违规处罚与连锁董事公司分析师盈余预测准确性关系的调节效应。

因此,本文基于2010—2017年证券分析师对沪

深非金融A股上市公司盈余预测样本,首先,检验信息披露违规处罚对连锁董事公司分析师盈余预测产生的溢出效应;其次,考察明星分析师对上述溢出效应的调节作用;再次,进一步分析信息披露违规处罚引发资本市场对连锁董事公司的负面反应;最后,实证考察信息披露违规处罚对连锁董事公司分析师盈余预测正向偏差和负向偏差的不同影响。本文得到以下结论:第一,信息披露违规处罚发生后,分析师对其连锁董事公司盈余预测准确性显著降低。第二,信息披露违规处罚发生后,明星分析师连锁董事公司盈余预测准确性发挥显著的负面调节效应。第三,信息披露违规处罚在短期内引发连锁董事公司超额累计收益显著下降。第四,信息披露违规处罚显著增大分析师对连锁董事公司盈余预测的正向偏差,对分析师盈余预测负向偏差无显著影响,并且相比于非明星分析师,明星分析师对连锁董事公司盈余预测的正向偏差更大。

本文的研究贡献在于以下几个方面:

第一,本文拓展监管处罚溢出效应研究的作用主体,丰富监管处罚溢出效应的研究内容。对于监管处罚溢出效应的作用主体,已有研究大多是关注与违规处罚公司关联的上市公司(即连锁董事公司)(Kang, 2008<sup>[7]</sup>; Brown和Zhou, 2015<sup>[3]</sup>; Brown等, 2018<sup>[4]</sup>)。不同的是,本文基于董事网络的路径,重点关注的是连锁董事公司的分析师盈余预测,拓展监管处罚溢出效应所涉及的研究主体,揭示监管处罚通过董事网络对分析师盈余预测产生溢出效应的经验证据。该结果揭示资本市场中潜在的金融风险,为防范金融风险,加强监管科技建设,提供新的研究视角。

第二,从信息披露违规处罚这一外生冲击事件的视角,丰富分析师盈余预测相关领域的文献。现有文献从公司特征、信息披露质量、公司治理水平、制度特征以及非理性因素等方面研究其对分析师预测偏差的影响。本文提供信息披露违规处罚公告事件影响分析师对连锁公司盈余预测偏差的经验证据。

第三,在Gu等(2013)<sup>[17]</sup>以及唐松莲和陈伟(2017)<sup>[16]</sup>研究的基础上,本文研究发现公司发生信息披露违规处罚后,与非明星分析师相比,明星分析师对连锁董事公司发布更为乐观的盈余预测,预测准确性更低,说明明星分析师与连锁董事公司或证券公司之间存在利益交换的可能性,进一步验证了利益结盟假说,该结果为中国资本市场中亟需建立分析师声誉激励的完备机制提供经验证据参考。



## 二、文献综述与研究假设

### (一) 信息披露违规处罚与分析师预测准确性

Beatty 等 (2013)<sup>[2]</sup> 认为监管处罚决定不仅给被处罚对象带来显著的负面影响, 而且波及被处罚对象以外的个人或组织, 即惩罚存在溢出效应。监管处罚可能通过各种社会网络关系对与违规公司相关联的其他主体产生溢出效应。

监管处罚在董事网络关系中存在溢出效应。根据归因理论 (Ross, 1977<sup>[5]</sup>), 未预期的负面结果 (如财务舞弊) 很可能引发投资者的“归因搜索” (Weiner, 1985<sup>[6]</sup>)。公司发生财务舞弊处罚后, 投资者不仅将注意力转向被控公司的管理层, 同时对董事履行其监管职能的有效性提出了质疑 (Hillman 和 Dalziel, 2003<sup>[18]</sup>), 该公司外部董事的名誉将受损害, 更有可能在董事职业市场上受到惩罚的形式承担个人责任 (Brochet 和 Srinivasan, 2014<sup>[19]</sup>)。Fich 和 Shivdasani (2007)<sup>[20]</sup> 研究发现公司受到监管处罚时, 其连锁董事公司会有显著为负的超额收益。Kang (2008)<sup>[7]</sup> 运用事件研究法, 实证研究发现公司因财务舞弊被监管处罚后, 投资者由于怀疑欺诈性的财务活动也会发生在连锁公司, 导致连锁董事公司市值下降。进一步发现, 当连锁董事在公司担任审计委员会主席或者公司治理委员会主席, 连锁董事公司市值影响很可能更大。周泽将和刘中燕 (2015)<sup>[21]</sup> 研究发现投资者对受罚独立董事兼任公司呈现出短暂的、微弱的负面市场反应。可见, 监管处罚的效应可以通过董事网络从被处罚的公司蔓延到连锁公司, 在短期内引发连锁董事公司市值下降, 不仅声誉受损, 而且会影响连锁董事公司的融资成本从而使盈余的不确定性增加。信息的不确定性影响分析师预测偏差 (Zhang, 2006<sup>[10]</sup>)。信息不确定程度越大, 分析师对信息的反应越不完全, 分析师预测偏差越大 (Gu 和 Wu, 2003<sup>[11]</sup>; Zhang, 2006<sup>[10]</sup>)。众多研究验证证券分析师对信息反应不足假说。“反应不足”通常归因在不确定性环境下分析师的判断与偏见, 如在信息的识别与传递过程中分析师保守主义或者过度自信 (Zhang, 2006<sup>[10]</sup>)。因此, 信息披露违规处罚后, 连锁董事公司信息不确定性程度的增加, 可能会降低分析师对连锁董事公司盈余预测的准确性。

同时, Jennings 等 (2011)<sup>[22]</sup> 发现在同行公司因操纵收益而被 SEC 调查后, 公司减少了可自由支配

的应计利润。SEC 对公司的行业领导者、紧密竞争对手和众多业内同行的风险因素定性披露的审查, 而没有收到 SEC 披露质量意见书的公司会在很大程度上修正随后年度公司的披露质量 (Brown 等, 2018<sup>[4]</sup>)。因此, 监管处罚在同行业关系网络中产生溢出效应, 提高同行业公司的信息披露质量。基于董事网络关系, Zhong 等 (2017)<sup>[23]</sup> 研究发现监管处罚后, 连锁董事发挥积极治理效应, 连锁公司的财务报告透明度也得到显著提高。信息披露水平是影响分析师预测准确性的关键因素。因此, 信息披露违规处罚后, 连锁董事公司财务报告透明度的提高, 意味着分析师可以获得更高质量的公司特质信息, 有助于提高盈余预测的准确性。

根据以上理论分析, 本文提出竞争性假设:

**H1a:** 公司发生信息披露违规处罚, 分析师对其连锁董事公司盈余预测准确性降低。

**H1b:** 公司发生信息披露违规处罚, 分析师对其连锁董事公司盈余预测准确性提高。

### (二) 信息披露违规处罚与分析师盈余预测准确性: 分析师声誉效应

自 2003 年开始, 《新财富》杂志主办的“明星分析师”评选成为中国分析师行业中最具影响力的市场评价活动。获奖的分析师在中国通常称为“明星分析师”。“明星”荣誉称号对于分析师是一项无形资产, 代表资本市场对分析师能力的认可, 能够显著提高分析师的社会地位与经济利益。然而, 胡奕明和金洪飞 (2006)<sup>[24]</sup> 认为, 由于中国的资本市场中并没有完备的声誉机制, 分析师在积累自身声誉时对投资者利益保护等道德约束重视不够。一方面, 基于迎合理论, 声誉越高的分析师, 其试错成本越高, “羊群行为”动机越强烈, 越倾向于采取“跟风”策略 (蔡庆丰等, 2011<sup>[14]</sup>; 董大勇等, 2012<sup>[15]</sup>), 盈余预测偏差越大。经验研究发现公司受到信息披露违规处罚, 连锁董事公司市值在短期内受到负面影响 (Kang, 2008<sup>[7]</sup>), 资本市场对连锁董事公司悲观预期。因此, 基于迎合动机, 明星分析师可能迎合投资者对连锁董事公司负面预期的动机更强烈, 发布更为悲观的盈余预测报告, 导致盈余预测偏差进一步升高, 准确性进一步降低。

同时, 根据利益结盟假说, 唐松莲和陈伟 (2017)<sup>[16]</sup> 研究发现, 在中国资本市场中分析师的声誉效应并没有战胜利益结盟动机, 分析师与上市公司利益结盟,

盈余预测的乐观度提升,准确性进一步下降。分析师的佣金压力越大,其推荐股票的态度就越乐观,预测的偏差增加,准确性越低(Gu等,2013<sup>[17]</sup>;官峰等,2015<sup>[25]</sup>)。违规处罚引发连锁董事公司出现负面市场反应(Kang,2008<sup>[7]</sup>),因此,证券公司以及连锁董事公司的管理层存在向上调整市场预期的动机。基于利益结盟假说,证券公司可能会借助交易佣金对分析师施加压力。与非明星分析师相比,明星分析师的薪酬更高,因而面对佣金压力更大,将导致明星分析师发布更为乐观的盈余预测报告,盈余预测偏差进一步增大。因此,无论是基于迎合动机还是利益结盟动机,信息披露违规处罚后,与非明星分析师相比,明星分析师盈余预测偏差进一步增大,准确性更低。

因此,根据迎合理论和利益结盟假说分析,我们提出假设2:

H2:公司受到信息披露违规处罚,与非明星分析师相比,明星分析师对其连锁董事公司盈余预测准确性更低。

### 三、研究设计

#### (一) 样本选择与数据来源

本文选择2010—2017年沪深两市的A股上市公司为研究样本。首先,剔除了ST类公司、金融公司、资不抵债的公司以及当年IPO公司。其次,由于本文考察公司违规受处罚对其连锁董事公司的溢出效应,因此删除所有在样本期间内连锁董事公司受处罚的观测值以及违规的公告时间与违规时间不在同一年度内的观测值。再次,数据统计结果显示,受到信息披露违规处罚的公司样本占约28%,与没有受到违规处罚的公司样本量差距较大。为了避免样本选择偏差影响回归结果的准确性,本文采用同年度、同规模、同杠杆率、同盈利能力、同股权集中度、同独董比例、同产权性质、同机构投资者比例、同市账比例、分析师所属券商公司规模相同、分析师工作经验相同以及分析师学历相同为标准,按照1:1匹配的倾向评分匹配法(PSM)对样本进行筛选,最后得到1542家公司,3635个观测值。最后,行业分类依据中国2012年证监会颁布的《上市公司行业分类指引》。本文的数据主要来源是国泰安金融研究数据库(CSMAR)和Wind资讯金融数据库。

图1列示处理组和控制组匹配前密度图。图2列示了处理组和控制组分别按照1:1匹配后的密度图。

如图1所示,在匹配前,处理组的p-score先低于控制组,后又高于控制组。在匹配后,如图2所示,两组的p-score基本一致,说明处理组和控制组在倾向得分匹配后达到了共同支撑标准。

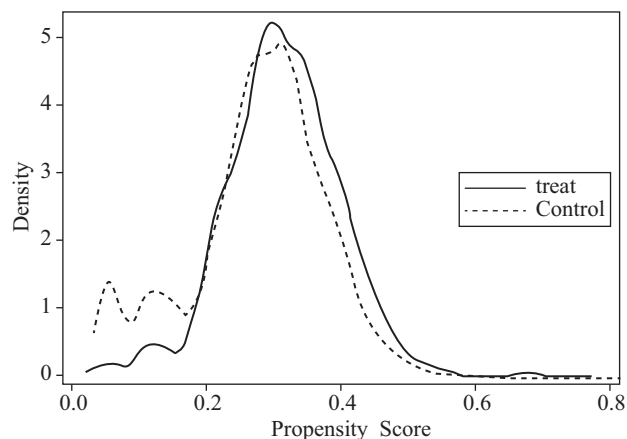


图1 倾向评分匹配前的密度函数图

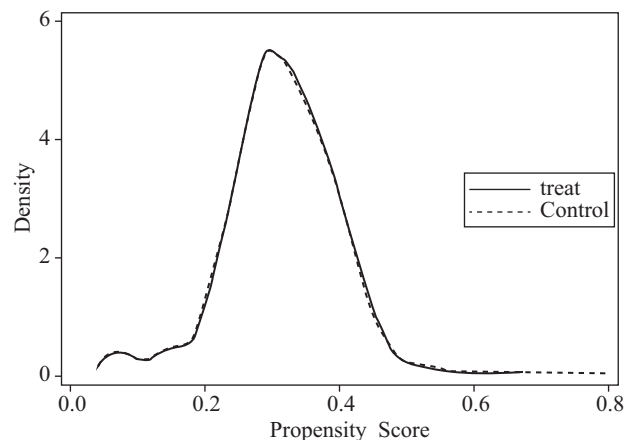


图2 1:1倾向评分匹配后的密度函数图

#### (二) 变量定义

##### 1. 被解释变量。

分析师盈余预测准确性(*Ananerror*)。本研究检验了违规公司的信息披露违规处罚对其连锁董事公司的分析师预测准确性产生的影响,因此分析师盈余预测准确性(*Ananerror*)是本研究的被解释变量。考虑到每个分析师在一年内会发布多次盈余预测分析报告以及连锁董事公司违规处罚公告时间点的差异性,在借鉴Lang和Lundholm(1996)<sup>[26]</sup>、白晓宇(2009)<sup>[27]</sup>对分析师预测准确性界定的基础上,本文选择一年内每个分析师对其分析公司最后一次的分析报告数据,计算分析师预测年净利润和公司实际的年净利润差值的绝对值,并且用公司实际净利润的绝对值进行标准化来表示分析师预测偏差,计算公式如下:

$$Ananerror = \frac{abs(Fnetpro - Netpro)}{abs(Netpro)} \quad (1)$$

其中, *Fnetpro* 表示分析师预测本公司的年净利润。*Netpro* 表示公司实际的年净利润, 然后据此计算出每个公司所有分析师预测偏差的均值。*Ananerror* 的值越大, 分析师盈余预测的准确性越低。

2. 解释变量。

(1) 信息披露违规处罚 (*Shock*)。

信息披露违规处罚是本文的解释变量之一。

第一步, 本文以违规处罚公告年作为目标公司受到监管处罚冲击的年份。只要公司在当年受到至少一次行政性处罚, 本文均且仅保留一条年度-公司处罚样本, 认为该公司在当年受到了违规处罚。

第二步, 筛选出连锁董事公司。如果一家公司的董事会至少有一名董事同时担任另一家公司的董事, 则该公司在同一年被视为连锁董事公司。

第三步, 将连锁董事公司与违规公司进行配对, 与受处罚的数据进行合并, 我们得到连锁董事公司违规处罚数据库。由于本文关注的是监管处罚的溢出效应, 所以本文首先筛选受到证监会违规处罚的公司, 然后匹配与之有共同连锁董事且本公司未受到处罚的公司, 获得违规公司的连锁董事公司, 作为我们研究的目标公司。

第四步, 按处罚类型进行分类, 获得连锁董事公司信息披露违规处罚数据库。根据中国证监会违规处罚公告的原因, 参照滕飞等 (2016)<sup>[28]</sup> 的研究, 将虚构利润、虚列资产、虚假记载、推迟披露、重大遗漏、批露不实和一般会计处理不当等违规类型界定为

信息披露违规处罚, 其他经营违规处罚。

第五步, 信息披露违规处罚变量, *Shock*。若目标公司至少有一个连锁董事公司在当年内受到了信息披露违规处罚, 那么对于目标公司 (即研究与违规处罚公司关联的连锁董事公司) 而言, *Shock* 赋值为 1; 若目标公司没有连锁董事公司在当年内受到监管处罚, 那么对于目标公司 (即研究与违规处罚公司关联的连锁董事公司) 而言, *Shock* 赋值为 0。

(2) 信息披露违规处罚与明星分析师的交互项 (*StarShock*)。

为进一步验证中国证券市场中分析师声誉对上述溢出效应的影响, 构建信息披露违规处罚与明星分析师的交互项, *Starshock*。其中, 明星分析师设为虚拟变量, 用 *Star* 表示。参照黄俊等 (2018)<sup>[29]</sup> 按照新财富分析师排名, 位列前五时, 将 *Star* 赋值为 1, 表示明星分析师, 否则, 将 *Star* 赋值为 0, 表示非明星分析师。

3. 控制变量。

在控制变量的界定上, 本文参考了王攀娜和罗宏 (2017)<sup>[30]</sup>、罗棻心等 (2018)<sup>[31]</sup>、王永妍等 (2019)<sup>[32]</sup> 的研究, 选择公司规模 (*Size*)、资产负债率 (*Lev*)、股票波动率 (*Volat*)、独立董事比例 (*Indirector*)、第一大股东持股比例 (*TOP*)、营业收入增长率 (*Growth*)、市账比 (*MB*)、产权性质 (*State*)、机构投资者持股比例 (*Inshare*)、分析师所属公司的规模 (*Reportnum*)、分析师工作经验 (*Exp*)、分析师学历 (*Edu*)、行业 (*Indcd*) 和年份 (*Year*) 因素作为控制变量。控制变量的定义与计算方法见表 1。

表 1 控制变量的定义和计算方法

变量名	含义	计算方法
<i>Size</i>	公司规模	期末总资产的自然对数
<i>Lev</i>	资产负债率	期末总负债/期末总资产
<i>Volat</i>	股票波动率	股票日平均收益的标准差
<i>Indirector</i>	独立董事比例	独立董事数量/所有董事数量
<i>TOP</i>	股权集中度	第一大股东持股比例
<i>Growth</i>	营业收入增长率	(当年营业收入-前一年营业收入)/前一年营业收入
<i>MB</i>	市账比	总市值/期末总资产
<i>State</i>	产权性质	国有企业=1; 民营企业=0
<i>Inshare</i>	机构投资者持股比例	机构投资者所持股份占总股份的比例

续前表

变量名	含义	计算方法
<i>Reportnum</i>	分析师所属公司的规模	分析师所属公司当年发布分析报告的数量
<i>Exp</i>	分析师工作经验	分析师出具第一个预测报告后发布预测报告的季度数,取自然对数
<i>Edu</i>	分析师学历	低学历(本科及以下)=0;高学历(硕士研究生和博士研究生)=1
<i>Indcd</i>	依据中国证监会颁布的《上市公司行业分类指引》	
<i>Year</i>	2010—2017年度数据	

### (三) 研究模型

为研究信息披露违规处罚对连锁董事公司分析师盈余预测准确性的影响,验证研究假设 H1,构建模型(2):

$$Ananerror_{i,t} = \alpha + \beta_1 Shock_{i,t} + \beta_2 ControlVariables_{i,t} + \beta_3 Indcd + \beta_4 Year + \varepsilon \quad (2)$$

其中, *Ananerror* 为被解释变量,表示连锁董事公司的分析师预测准确性。*Shock* 为解释变量,表示信息披露违规处罚。若 *Shock* 的系数显著为正,说明公司发生信息披露违规处罚,分析师对其连锁董事公司盈余预测的准确性越低,研究假设 H1a 得到验证;相反, *Shock* 的系数显著为负,研究假设 H1b

得到验证。

## 四、实证结果

### (一) 描述性统计

表2列示了在1:1倾向评分匹配样本下变量的描述性统计结果。分析师预测准确性(*Ananerror*)的均值为0.860,标准差为2.288,说明中国分析师盈余预测的准确性可能存在较大的差异。信息披露违规处罚(*Shock*)的均值为0.686,说明对于目标公司而言,有68.6%的连锁董事公司受到信息披露违规处罚。其他控制变量的描述性统计结果详见表2。

表2 描述性统计

变量	样本量	均值	中位数	标准差	最大值	最小值
<i>Ananerror</i>	3 635	0.860	0.216	2.288	18.200	0.002
<i>Shock</i>	3 635	0.686	1.000	0.464	1.000	0.000
<i>Star</i>	3 635	0.105	0.000	0.306	1.000	0.000
<i>Starshock</i>	3 635	0.080	0.000	0.271	1.000	0.000
<i>Size</i>	3 635	22.537	22.300	1.365	27.000	20.200
<i>Lev</i>	3 635	0.434	0.430	0.202	0.863	0.056
<i>TOP</i>	3 635	0.361	0.343	0.158	0.794	0.088
<i>Inshare</i>	3 635	0.448	0.468	0.240	0.903	0.007
<i>Volat</i>	3 635	0.464	0.406	0.222	1.250	0.163
<i>MB</i>	3 635	2.208	1.690	1.924	9.850	0.156
<i>Indirector</i>	3 635	0.374	0.333	0.054	0.571	0.333
<i>State</i>	3 635	0.392	0.000	0.488	1.000	0.000
<i>Growth</i>	3 635	0.185	0.126	0.331	1.850	-0.430
<i>Reportnum</i>	3 635	1 398.120	1 284.000	858.107	3 261.000	37.000
<i>Exp</i>	3 635	2.431	2.480	0.734	3.830	0.588
<i>Edu</i>	3 635	0.929	1.000	0.257	1.000	0.000

(二) 多元回归分析

表3报告了违规处罚、连锁董事公司分析师预测与分析师声誉的实证检验的结果。列(1)报告了信息披露违规处罚与连锁董事公司分析师预测准确性的回归结果。信息披露违规处罚(*Shock*)与连锁董事公司分析师预测准确性(*Ananerror*)的系数为0.217,在1%的水平上显著,说明信息披露违规处罚显著降低了分析师对连锁董事公司盈余预测的准确性,支持研究假设H1a。

表3列(2)报告了明星分析师调节效应的检验结果。列(3)和列(4)分别报告了明星分析师样本组和非明星分析师样本组中,信息披露违规处罚(*Shock*)与连锁董事公司分析师预测准确性(*Ananerror*)的回归结果。列(2)全样本中,信息披露违规处罚(*Shock*)与明星分析师(*Star*)交互项

(*Starshock*)的系数为1.219,显著为正,说明信息披露违规处罚后,与非明星分析师相比,明星分析师对连锁董事公司预测准确性更低,支持研究假设H2。进一步按*Star*进行分组,列(3)明星分析师组中,*Shock*对*Ananerror*回归系数为0.999,高于列(4)非明星分析师组中*Shock*的系数0.109,并且列(3)和列(4)组间系数差异卡方值为5.28,显著为正,说明信息披露违规处罚后,明星分析师和非明星分析师对连锁董事公司的预测准确性存在显著差别,与非明星分析师相比,明星分析师对连锁董事公司预测的准确性更低,进一步验证了研究假设H2。

表3的检验结果说明信息披露违规处罚公告发生后,分析师对连锁董事公司预测的准确性降低,并且明星分析师发挥显著的负面调节效应。

表3 违规处罚、连锁董事公司分析师预测与分析师声誉的检验结果

解释变量	被解释变量: <i>Ananerror</i>			
	全样本	全样本	明星分析师	非明星分析师
	(1)	(2)	(3)	(4)
<i>Shock</i>	0.217*** (3.142)	0.082 (1.178)	0.999** (2.480)	0.109 (1.622)
<i>Star</i>		-0.162 (-0.960)		
<i>Starshock</i>		1.219*** (3.850)		
<i>Size</i>	-0.055 (-0.965)	-0.051 (-0.906)	0.919*** (3.190)	-0.178*** (-3.761)
<i>Lev</i>	1.051*** (3.592)	1.046*** (3.605)	-0.096 (-0.072)	1.541*** (5.864)
<i>TOP</i>	-0.766*** (-3.002)	-0.902*** (-3.445)	-2.056 (-1.550)	-0.514** (-2.195)
<i>Inshare</i>	-0.133 (-0.736)	-0.106 (-0.581)	-0.350 (-0.357)	-0.375** (-2.218)
<i>Volat</i>	0.962*** (2.721)	0.889*** (2.588)	-1.402* (-1.740)	0.281 (0.849)
<i>MB</i>	0.042 (1.179)	0.044 (1.239)	0.215 (1.329)	0.023 (0.638)
<i>Indirector</i>	0.509 (0.873)	0.430 (0.728)	1.444 (0.456)	1.105* (1.832)

续前表

解释变量	被解释变量: <i>Ananerror</i>			
	全样本	全样本	明星分析师	非明星分析师
	(1)	(2)	(3)	(4)
<i>State</i>	0.065 (0.738)	0.047 (0.530)	-0.316 (-0.792)	0.040 (0.452)
<i>Growth</i>	-0.091 (-0.379)	-0.067 (-0.291)	2.231*** (3.260)	-0.765*** (-5.985)
<i>Reportnum</i>	-0.000 (-0.734)	-0.000** (-2.102)	-0.001*** (-4.048)	-0.000 (-1.111)
<i>Exp</i>	0.093* (1.937)	0.082* (1.770)	1.035*** (3.765)	-0.047 (-1.113)
<i>Edu</i>	-0.080 (-0.529)	-0.149 (-0.993)	0.294 (0.294)	-0.135 (-0.907)
<i>Constant</i>	2.825** (2.093)	3.008** (2.298)	-13.454 (-1.551)	5.525*** (5.112)
<i>Indcd/Year</i>	已控制	已控制	已控制	已控制
<i>N</i>	3 635	3 635	381	3 254
<i>R-squared</i>	0.050	0.063	0.559	0.072
<i>adj_R<sup>2</sup></i>	0.039	0.053	0.511	0.061
<i>F</i>	8.113	6.890	11.430	6.429
<i>Chi2</i>			5.28**	

注:\*\*\*、\*\*、\*分别表示变量在1%、5%、10%的水平上显著,括号中数字为稳健的*t*值。下同。

## 五、进一步分析与检验

### (一) 信息披露违规处罚后连锁董事公司的市场反应

经验研究发现监管处罚在短期内引发连锁董事公司市值下降,声誉受损(Kang, 2008<sup>[7]</sup>),但是,在中国资本市场中,信息披露违规处罚在短期内是否也对连锁董事公司产生负面效应,本文在此进行检验,从而进一步验证本文研究假设H1a的推理与论证的严谨性。

参考Kang(2008)<sup>[7]</sup>的做法,采用普林斯顿大学

方法检验累计超额收益率(CAAR)的显著性,事件窗口期定为信息披露违规处罚事件日的前后5天。违规处罚公告发生,连锁董事公司的累计超额收益在时间窗口期内的变化如表4所示。表4的最后一行列示了事件窗口期内,连锁董事公司累计超额收益(CAAR)的下滑幅度,该数值为-0.0069,且在1%水平上显著,这说明信息披露违规处罚公告发生后,连锁董事公司超额累计收益显著下降,股价显著下跌,说明信息披露违规处罚引发资本市场对连锁董事公司的负面反应。

表4 连锁董事公司累计超额收益率时序表[-5, 5]

变量	时期	系数	标准差	<i>T</i> 值
<i>AAR<sub>t</sub></i>	-5	-0.0031		-2.17*
	-4	-0.0043		-2.73**
	-3	-0.0028		-1.68

续前表

变量	时期	系数	标准差	T 值
AAR <sub>t</sub>	-2	-0.003 0		-1.71
	-1	-0.004 0		-2.19*
	0	-0.004 6		-2.38*
	1	-0.004 9		-2.40*
	2	-0.007 2		-3.36***
	3	-0.008 7		-3.89***
	4	-0.007 7		-3.32***
	5	-0.008 0		-3.30***
CAAR	[-5, 5]	-0.006 9	0.002 0	-3.50***

(二) 信息披露违规处罚与连锁董事公司分析师预测的正向偏差和负向偏差

在主检验中, 根据公式 (1) 本文用分析师预测年净利润和公司实际的年净利润差值的绝对值指标 (*Ananerror*), 度量分析师盈余预测的准确性。为进一步分析信息披露违规处罚对连锁董事公司分析师预测正向偏差和负向偏差的不同效应, 并对研究假设 H2 的理论推理进一步检验, 我们在公式 (1) 基础上去掉绝对值, 创建分析师预测偏差变量 *Error*, 如公式

(3) 所示:

$$Error = \frac{Fnetpro - Netpro}{Netpro} \quad (3)$$

*Error* 的值越大, 表明分析师预测偏差越大。当 *Error* 大于 0 时, 表示分析师预测存在正向偏差, 用变量 *Error+* 表示; 当 *Error* 小于 0 时, 表示分析师预测存在负向偏差, 用变量 *Error-* 表示。将主检模型中的因变量 (*Ananerror*) 分别替换成 *Error*、*Error+* 和 *Error-*, 进行回归检验, 得到结果见表 5。

表 5 信息披露违规处罚与连锁董事公司分析师预测偏差的检验结果

解释变量	被解释变量: <i>Error</i>		被解释变量: <i>Error+</i>		被解释变量: <i>Error-</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Shock</i>	0.251*** (3.626)	0.106 (1.518)	0.305*** (3.258)	0.116 (1.187)	-0.017 (-1.388)	-0.022* (-1.726)
<i>Star</i>		-0.206 (-1.270)		-0.431* (-1.814)		-0.067 (-1.488)
<i>Starshock</i>		1.323*** (4.190)		1.550*** (3.974)		0.063 (1.179)
其他变量	已控制	已控制	已控制	已控制	已控制	已控制
<i>N</i>	3 635	3 635	2 554	2 554	1 081	1 081
<i>R-squared</i>	0.054	0.069	0.060	0.073	0.149	0.152
<i>adj_R<sup>2</sup></i>	0.044	0.058	0.045	0.058	0.118	0.119
<i>F</i>	6.360	6.685	9.739	6.221	4.801	4.645

注: \*\*、\*、\* 分别表示变量在 1%、5%、10% 的水平上显著, 括号中数字为稳健的 *t* 值。

表 5 报告了信息披露违规处罚与连锁董事公司分析师预测偏差实证检验的结果。如表 5 所示, 列 (1) 信息披露违规处罚 (*Shock*) 与连锁董事公司分析师预测偏差 (*Error*) 的系数为 0.251, 在 1% 的水平上显著, 说明信息披露违规处罚显著增大了分析师

对连锁董事公司盈余预测的偏差, 降低了分析师预测准确性, 再次验证了研究假设 H1a。列 (2) 报告了明星分析师对信息披露违规处罚与连锁董事公司分析师预测偏差的调节效应的回归结果。列 (2) 信息披露违规处罚 (*Shock*) 与明星分析师 (*Star*) 交互项

(*Starshock*) 的系数为 1.323, 在 1% 的水平上显著为正, 说明信息披露违规处罚后, 与非明星分析师相比, 明星分析师对连锁董事公司预测偏差更大, 预测准确性更低, 再次验证研究假设 H2。

表 5 的列 (3) 报告了信息披露违规处罚与连锁董事公司分析师预测正向偏差的回归结果。信息披露违规处罚 (*Shock*) 与连锁董事公司分析师预测正向偏差 (*Error+*) 的系数为 0.305, 在 1% 的水平上显著, 说明信息披露违规处罚显著增大了分析师对连锁董事公司盈余预测的正向偏差。列 (4) 报告了明星分析师对信息披露违规处罚与连锁董事公司分析师预测正向偏差的调节效应的回归结果。列 (4) 信息披露违规处罚 (*Shock*) 与明星分析师 (*Star*) 交互项 (*Starshock*) 的系数为 1.550, 在 1% 的水平上显著为正, 说明信息披露违规处罚后, 与非明星分析师相比, 明星分析师加剧了连锁董事公

司预测正向偏差。

表 5 的列 (5) 报告了信息披露违规处罚与连锁董事公司分析师预测负向偏差的回归结果。信息披露违规处罚 (*Shock*) 与连锁董事公司分析师预测负向偏差 (*Error-*) 的系数为 -0.017, 统计上并不显著, 说明信息披露违规处罚对连锁董事公司分析师盈余预测负向偏差无显著影响。列 (6) 报告了明星分析师对信息披露违规处罚与连锁董事公司分析师预测负向偏差调节效应的回归结果。信息披露违规处罚 (*Shock*) 与明星分析师 (*Star*) 交互项 (*Starshock*) 的系数也不显著, 说明信息披露违规处罚后, 明星分析师对连锁董事公司盈余预测的负向偏差无显著调节效应。

进一步地, 使用 *Error*、*Error+* 和 *Error-* 按明星分析师和非明星分析师对模型 (2) 进行分组检验, 考察组间差异的显著性, 回归结果见表 6。

表 6 明星分析师与非明星分析师的盈余预测偏差分组差异检验结果

解释变量	被解释变量: <i>Error</i>		被解释变量: <i>Error+</i>		被解释变量: <i>Error-</i>	
	明星分析师	非明星分析师	明星分析师	非明星分析师	明星分析师	非明星分析师
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Shock</i>	1.193*** (2.961)	0.131* (1.952)	1.196** (2.303)	0.170* (1.828)	0.089 (1.226)	-0.021 (-1.632)
其他变量	已控制	已控制	已控制	已控制	已控制	已控制
<i>N</i>	381	3 254	308	2 246	73	1 008
<i>R-squared</i>	0.561	0.080	0.666	0.086	0.769	0.139
<i>adj_R<sup>2</sup></i>	0.512	0.069	0.620	0.070	0.604	0.105
<i>F</i>	11.484	7.164	14.563	5.344	4.662	4.122
<i>Chi2</i>	7.50***		4.29**		3.77*	

如表 6 所示, 列 (1) 明星分析师组中, 信息披露违规处罚 (*Shock*) 与连锁董事公司分析师预测偏差 (*Error*) 的系数为 1.193, 在 1% 的水平上显著; 列 (2) 非明星分析师组中, 信息披露违规处罚 (*Shock*) 与连锁董事公司分析师预测偏差 (*Error*) 的系数为 0.131, 在 10% 的水平上显著。列 (1) 与列 (2) 两组组间系数差异卡方值为 7.50, 且在 1% 的水平上显著。该结果表明信息披露违规处罚后, 与非明星分析师相比, 明星分析师对连锁董事公司预测偏差更大, 预测准确性更低, 进一步支持研究假设 H2。

表 6 的列 (3) 明星分析师组中, 信息披露违规处罚 (*Shock*) 与连锁董事公司分析师预测正向偏差

(*Error+*) 的系数为 1.196, 在 5% 的水平上显著; 列 (4) 非明星分析师组中, 信息披露违规处罚 (*Shock*) 与连锁董事公司分析师预测正向偏差 (*Error+*) 的系数为 0.170, 在 10% 的水平上显著。列 (3) 与列 (4) 两组组间系数差异卡方值为 4.29, 且在 5% 的水平上显著。该结果表明信息披露违规处罚后, 与非明星分析师相比, 明星分析师对连锁董事公司预测正向偏差更大, 进而导致预测准确性更低。

表 6 的列 (5) 和列 (6) 中信息披露违规处罚 (*Shock*) 与连锁董事公司分析师预测负向偏差 (*Error-*) 均不显著。该结果表明信息披露违规处罚后, 明星分析师和非明星分析师并没有对连锁董事公



司预测负向偏差产生显著影响。

综上所述,信息披露违规处罚显著增大分析师对连锁董事公司盈余预测的正向偏差,对分析师盈余预测负向偏差无显著影响,并且明星分析师对连锁董事公司盈余预测正向偏差发挥显著的负面调节效应,即相比于非明星分析师,明星分析师显著增大连锁董事公司分析师预测的正向偏差,对盈余预测的负向偏差无显著影响。

根据前文研究假设 H2 的理论推理:基于迎合动机或者利益结盟动机,信息披露违规处罚后,与非明星分析师相比,明星分析师盈余预测准确性更低,表 5 和表 6 这一结果意味着 H2 中明星分析师对其连锁董事公司盈余预测准确性更低,不是由于明星分析师迎合动机导致,更可能是与分析师和连锁董事公司或证券公司之间存在利益结盟有关。根据 Gu 等(2013)<sup>[17]</sup>的研究,证券公司与分析师之间存在利益交换的可能性,证券公司可能会借助交易佣金对分析师施加压力,导致分析师发布更为乐观的预测,准确性降低。唐松莲和陈伟(2017)<sup>[16]</sup>研究也发现分析师与上市公司利益结盟,导致盈余预测的准确性进一步下降,分析师的声誉效应并没有战胜利益结盟动机。在上一部分分析中,我们发现信息披露违规处罚引发资本市场对连锁董事公司的负面反应,连锁董事公司的股价发生显著下降,投资者对连锁董事公司持有负面预期。因此,连锁董事公司的管理层或证券公司很可能存在向上调整市场预期的动机,借助交易佣金对分析师施压。与非明星分析师相比,明星分析师的薪酬更高,因而面对佣金压力更大,导致分析师(包括明星分析师)对连锁董事公司发布更为乐观的盈余预测,盈余预测偏差进一步增大。

## 六、结论与启示

本文基于 2010—2017 年证券分析师对沪深非金融 A 股上市公司盈余预测样本,研究公司发生信息披露违规处罚这一外生事件,对其连锁董事公司分析师盈余预测产生的溢出效应以及明星分析师的调节效应。通过实证研究,本文得出以下结论:

第一,信息披露违规处罚发生后,分析师对其连锁董事公司盈余预测准确性显著降低。

第二,信息披露违规处罚发生后,明星分析师和非明星分析师对连锁董事公司的预测准确性存显著差别。明星分析师对连锁董事公司盈余预测准确性发挥显著的负面调节效应,与非明星分析师相比,明星分析师对连锁董事公司预测准确性更低。

第三,信息披露违规处罚引发资本市场对连锁董事公司的负面反应。连锁董事公司超额累计收益显著下降,股价显著下跌。

第四,信息披露违规处罚显著增大分析师对连锁董事公司盈余预测的正向偏差,对分析师盈余预测负向偏差无显著影响,并且相比于非明星分析师,明星分析师对连锁董事公司盈余预测的正向偏差更大。

本文的研究结论对证监会具有重要的启示:

第一,防控监管处罚因董事网络溢出的金融风险。基于董事网络,信息披露违规处罚后,分析师对连锁董事公司预测准确性降低,将会损害信息传递和资本市场的效率,随之产生的市场风险首先是向不具备风险承受能力的普通投资者扩散,更有可能通过行业、流动性、产品联结等其他渠道引发跨市场、跨机构的风险联动,演化成区域性、系统性金融风险。证监会需要加强防控因董事网络产生的金融风险的扩散与传染。

第二,目前中国资本市场中,亟待建立分析师声誉激励的完备机制,包括声誉回报机制,奖惩分明,优胜劣汰机制。分析师声誉激励机制有效性存在问题。信息披露违规处罚后,“明星”名誉对分析师预测准确性发挥负向调节作用,明星分析师对连锁董事公司盈余预测的乐观偏差更大。该结果说明在中国资本市场中,分析师利益结盟动机超过声誉效应,导致现有的声誉机制在一定程度上被扭曲。建议进一步完善明星分析师评选的机制,建立更客观公正的量化评价标准。证监会可以应用大数据、云计算等监管科技手段,发挥分析师“明星”荣誉的积极治理效应,提升主动监管效率,对整个资本市场运行状态进行实时的风险监测与预警。

## 参考文献

- [1] Trevino L K. The Social Effects of Punishment in Organizations: A Justice Perspective [J]. The Academy of Management Review, 1992, 17 (4): 647-676.

- [2] Beatty A, Liao S, Yu J. The Spillover Effect of Fraudulent Financial Reporting on Peer Firms' Investment [J]. *Journal of Accounting & Economics*, 2013, 55 (2/3): 183-205.
- [3] Brown L D, Zhou L. Interactions between Analysts' and Managers' Earnings Forecasts [J]. *International Journal of Forecasting*, 2015, 31 (2): 501-514.
- [4] Brown S V, Tian X, Tucker J W. The Spillover Effect of SEC Comment Letters on Qualitative Corporate Disclosure: Evidence from the Risk Factor Disclosure [J]. *Contemporary Accounting Research*, 2018, 35 (2): 622-656.
- [5] Ross L. The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process [J]. *Advances in Experimental Social Psychology*, 1977, 10: 173-220.
- [6] Weiner B. An Attributional Theory of Achievement Motivation and Emotion [J]. *Psychological Review*, 1985, 92 (4): 548-573.
- [7] Kang E. Director Interlocks and Spillover Effects of Reputational Penalties from Financial Reporting Fraud [J]. *Academy of Management Journal*, 2008, 51 (3): 537-555.
- [8] Brown L D, Higgins H N. Managers' Forecast Guidance of Analysts: International Evidence [J]. *Journal of Accounting & Public Policy*, 2005, 24 (4): 280-299.
- [9] 陈运森, 郑登津. 董事网络关系、信息桥与投资趋同 [J]. *南开管理评论*, 2017 (3): 159-171.
- [10] Zhang X F. Information Uncertainty and Analyst Forecast Behavior [J]. *Contemporary Accounting Research*, 2006, 23 (2): 565-590.
- [11] Gu Z, Wu J S. Earnings Skewness and Analyst Forecast Bias [J]. *Journal of Accounting and Economics*, 2003, 35 (1): 5-29.
- [12] Xu N H, Chan K C, Jiang X, et al. Do Star Analysts Know More Firm-specific Information? Evidence from China [J]. *Journal of Banking & Finance*, 2013, 37 (1): 89-102.
- [13] 张然, 汪荣飞, 王胜华. 分析师修正信息、基本面分析与未来股票收益 [J]. *金融研究*, 2017 (7): 156-174.
- [14] 蔡庆丰, 杨侃, 林剑波. 羊群行为的叠加及其市场影响——基于证券分析师与机构投资者行为的实证研究 [J]. *中国工业经济*, 2011, (12): 111-121.
- [15] 董大勇, 张尉, 赖晓东, 刘海斌. 谁领先发布: 中国证券分析师领先—跟随影响因素的实证研究 [J]. *南开管理评论*, 2012 (5): 56-63, 104.
- [16] 唐松莲, 陈伟. 声誉提升抑或利益结盟: 关联证券分析师实地调研动因研究 [J]. *管理世界*, 2017 (9): 178-179.
- [17] Gu Z, Li Z, Yang Y G. Monitors or Predators: The Influence of Institutional Investors on Sell Side Analysts [J]. *The Accounting Review*, 2013, 88 (1): 137-169.
- [18] Hillman A J, Dalziel T. Boards of Directors and Firm Performance: Integrating Agency and Resource Dependence Perspectives [J]. *Academy of Management Review*, 2003, 28 (3): 383-396.
- [19] Brochet F, Srinivasan S. Accountability of Independent Directors: Evidence from Firm's Subject to Securities Litigation [J]. *Journal of Financial Economics*, 2014, 111 (2): 430-449.
- [20] Fich E M, Shivdasani A. Financial Fraud, Director Reputation, and Shareholder Wealth [J]. *Journal of Financial Economics*, 2007, 86 (2): 306-336.
- [21] 周泽将, 刘中燕. 中国独立董事声誉机制的有效性研究——基于违规处罚市场反应视角的经验证据 [J]. *中央财经大学学报*, 2015 (8): 102-112.
- [22] Jennings J N, Kedia S, Rajgopal S. The Deterrent Effects of SEC Enforcement and Class Action Litigation [J]. *SSRN Electronic Journal*, 2011.
- [23] Zhong Q, Liu Y, Yuan C. Director Interlocks and Spillover Effects of Board Monitoring: Evidence from Regulatory Sanctions [J]. *Accounting and Finance*. 2017, 57 (5): 1605-1633.
- [24] 胡奕明, 金洪飞. 证券分析师关注自己的声誉吗? [J]. *世界经济*, 2006 (2): 71-81, 96.
- [25] 官峰, 靳庆鲁, 张佩佩. 机构投资者与分析师行为——基于定向增发解禁背景 [J]. *财经研究*, 2015 (6): 132-144.
- [26] Lang M H, Lundholm R J. Corporate Disclosure Policy and Analyst Behavior [J]. *The Accounting Review*, 1996, 71 (4): 467-492.
- [27] 白晓宇. 上市公司信息披露政策对分析师预测的多重影响研究 [J]. *金融研究*, 2009 (4): 92-112.
- [28] 滕飞, 辛宇, 顾小龙. 产品市场竞争与上市公司违规 [J]. *会计研究*, 2016 (9): 32-40.
- [29] 黄俊, 黄超, 位豪强, 王敏. 卖空机制提高了分析师盈余预测质量吗——基于融资融券制度的经验证据 [J]. *南开管理评论*, 2018 (2): 135-148.
- [30] 王攀娜, 罗宏. 放松卖空管制对分析师预测行为的影响——来自中国准自然实验的证据 [J]. *金融研究*, 2017 (11): 191-206.
- [31] 罗核心, 麻志明, 王亚平. 券商跟踪海外上市公司对国内分析师盈余预测准确性的影响 [J]. *金融研究*, 2018 (8): 190-206.
- [32] 王永妍, 耿卉, 王国臣. 股权激励影响分析师预测行为吗? [J]. *中央财经大学学报*, 2019 (4): 76-91.

(学术顾问: 梁上坤, 责任编辑: 张安平)

# 第三支柱个人养老金对养老金替代率的影响研究

## The Impact of the Introduction of the Third Pillar Individual Pension on Pension Replacement Rate

罗浩 周延

LUO Hao ZHOU Yan

**[摘要]** 在第一支柱养老金降费和养老金替代率逐年走低的大背景下, 维持经济增长与减缓养老金替代率下降成为现阶段提高社会福利的重要课题。本文构建世代交叠一般均衡模型, 考察了第三支柱个人养老金对总和养老金替代率等变量的影响。通过数值模拟得到结论: 高水平的第一支柱缴费率将制约经济发展水平; 保持现有第一支柱缴费比例 24%, 同时引入第三支柱个人养老金缴费率在 4%~26% 时, 能使总和养老金替代率达到 75%~82%。并且, 该模型还表明提高第三支柱个人养老金缴费率将提高经济发展水平。因此, 逐步降低第一支柱养老金缴费率并加快发展第三支柱个人养老金, 有助于提升总和养老金替代率。

**[关键词]** 第三支柱个人养老金 养老金替代率 养老金缴费率 世代交叠模型

**[中图分类号]** F840.67 **[文献标识码]** A **[文章编号]** 1000-1549 (2022) 11-0077-12

**Abstract:** With the reduction of the first pillar pension fee and the decline of pension replacement rate over year, maintaining both economic growth and stable pension replacement rate has become an important subject for improving social welfare. This paper constructs a generational overlap general equilibrium model and examines the effects of third-pillar personal pensions on pension replacement rates and other variables. High first-pillar contribution rate was revealed to restrain the level of economic development by numerical simulation. Maintaining the current first-pillar contribution rate of 24%, third-pillar personal pension contribution rate of 4%~26% can lead to pension replacement rate of 75%~82%. Moreover, increasing the third pillar personal pension contribution rate has positive impact on economic development as demonstrated by our model. Therefore, gradually reducing the first-pillar pension contribution rate and accelerating the development of the third-pillar personal pension will improve the total pension replacement rate.

**Key words:** Third pillar individual pension Pension substitution rate Pension contribution rate OLG

**[收稿日期]** 2022-04-25

**[作者简介]** 罗浩, 女, 1991年6月生, 华东师范大学经济学院博士研究生, 研究方向为养老金; 周延, 女, 1965年8月生, 华东师范大学经济学院教授, 博士生导师, 研究方向为养老金与养老保险。本文通讯作者为罗浩, 联系方式为 ecnuluhao@163.com。

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

2021年12月17日,中央全面深化改革委员会第二十三次会议审议通过了《关于推动个人养老金发展的意见》,标志着我国第三支柱个人养老金制度进入新时代。当今中国老龄化程度不断加深,与此同时,逐年下降的出生率导致少子化趋势明显,家庭养老负担不断加重,养老需求将加速升级。这一政策的出台是积极应对人口老龄化、实现养老金体系可持续发展的重要举措。

1994年世界银行首次提出养老金三支柱<sup>①</sup>概念。随着第三支柱个人养老金制度的落地<sup>②</sup>,中国养老金三支柱体系建设步入全面发展完善的新阶段。第一支柱包含基本养老保险和全国社会保障基金,截至2021年年末第一支柱参保人数已覆盖近10.3亿人;第二支柱是企业或机关事业单位在基本养老保险的基础上额外为职工建立的职业年金,覆盖近0.7亿人;第三支柱个人养老金以个人名义自愿参与,政府给予部分税收优惠,尚处于起步阶段。但当前养老金体系对由基本养老保险和社保基金补充构成的第一支柱依赖度较高,导致养老金体系结构性问题突出,呈现出第一支柱独大,第二支柱覆盖面有限,第三支柱几乎缺位的局面。虽然第一支柱总量和金额不断提升,但第一支柱养老金替代率仍然逐年走低,目前已不足50%,低于国际劳工组织规定的最低标准<sup>③</sup>。与此同时,第一支柱所面临的支付压力也在不断增加。根据中国社科院世界社保研究中心发布的《中国养老金精算报告2019—2050》,到2035年养老金结余将耗尽。另外,为了减轻企业负担,保持经济活力,自2019年5月,国家通过实施降低社保费率综合政策,形成企业发展与养老金制度发展的良性循环,养老金单位缴费降至16%。在此情形下,需要利用第二、第三支柱进行补充。第二支柱方面,参保人群多为机

关事业单位以及国有企业职工,覆盖范围有限,并且短期难以实现快速增长。相较于第一、第二支柱,第三支柱是以个人养老为目的、自愿参加并主导、国家给予税收优惠支持、投资运营市场化的积累型养老金制度。同时,市场化自主的养老选择也能加强居民自我保障意识和能力,引导居民通过长期投资和规划增加养老储蓄。所以,第三支柱个人养老金将对第一、第二支柱形成有力补充。

鉴于此,在养老需求加速升级的背景下,现阶段养老金三支柱体系还存在发展不平衡不充分、养老保障水平下降、养老保险基金不可持续的问题。为了维持退休后老年人的福利水平并解决养老金财务可持续性问题,第三支柱个人养老金亟待快速发展。但发展第三支柱个人养老金是否能提升总和养老金替代率<sup>④</sup>?第三支柱个人养老金投资是否会挤占当期消费从而影响经济增长?基于这种设想,本文构建一个世代交叠模型来探讨第三支柱个人养老金对养老金替代率、经济增长等变量的影响;测算出合理养老金替代率水平下最优第三支柱个人养老金缴费率的范围;测算了不同第一支柱缴费水平与不同第三支柱个人养老金缴费率水平的组合下,所能达到的总和养老金替代率的范围,以期对政策制定提供一定的参考。

## 二、文献综述

养老金替代率一般分为第一支柱基础养老金替代率<sup>⑤</sup>和第三支柱个人养老金替代率(贾洪波和温源,2005<sup>[1]</sup>)。有学者发现第一支柱基本养老保险<sup>⑥</sup>与第三支柱养老金产品<sup>⑦</sup>之间存在挤占效应,第一支柱养老金替代率越高导致个人在成年期购买第三支柱养老金产品的比例越低(蹇滨徽等,2021<sup>[2]</sup>)。当前我国第一支柱养老金中个人缴费率为8%,下降空间比较小。即使第一支柱缴费率已进行两次降费,但16%

- ① 第一支柱养老金由政府统一组织,实施普遍保障,具有强制性;第二支柱职业养老保险计划由雇主发起,包括职业年金和企业年金;第三支柱个人养老金计划,完全由个人自愿参与。
- ② 2022年4月21日,国务院办公厅发布《关于推动个人养老金发展的意见》,个人养老金制度正式落地。
- ③ 国际劳工组织《社会保障最低标准公约》规定,养老金的最低替代率为55%。
- ④ 总和养老金替代率指的是在多层次养老保险体制下,各层次养老保险待遇替代率加总形成的总替代率。由于第二支柱覆盖率较低,本文暂未考虑第二支柱,总和养老金替代率即第一支柱养老金替代率和第三支柱养老金替代率的加总。
- ⑤ 本文定义的第一支柱养老金替代率等同于参考文献中的第一支柱基础养老金替代率。
- ⑥ 第一支柱养老金也即第一支柱基本养老保险,实行社会统筹与个人账户相结合,费用由企业和职工个人共同承担。
- ⑦ 第三支柱养老保险产品也即商业养老保险,商业养老保险是第三支柱养老金的重要组成部分。2018年5月,中国银保监会印发《个人税收递延型商业养老保险业务管理暂行办法》的通知。

的缴费率仍然处于较高水平。第一支柱缴费率有一定的下调空间,降低第一支柱缴费率将优化企业、个人和政府三方的负担,也可以维持第一支柱养老金替代率在合理的水平(边恕和孙雅娜,2018<sup>[3]</sup>)。

第一支柱缴费率的下降可以减轻政府与企业的负担,许多国家在探索如何在降低第一支柱缴费率的情况下维持经济增长与第一支柱养老金替代率不变以保障个体福利最优(彭浩然等,2018<sup>[4]</sup>)。有研究者认为这是政策目标上的“不可能三角”,当第一支柱缴费率水平降低时,第一支柱缴费率的下降将难以维持经济和第一支柱养老金替代率同时增长(曾益等,2016<sup>[5]</sup>;景鹏等,2020<sup>[6]</sup>)。第一支柱养老金降费与公共政策目标存在冲突,政府需要在最大化经济增长率和养老金替代率中进行权衡。(汪玲等,2021<sup>[7]</sup>)。第一支柱养老金缴费率上升,经济增长率下降,第一支柱养老金替代率呈现出先上升后下降的倒U型关系(Cipriani和Pascucci,2020<sup>[8]</sup>;彭浩然等,2018<sup>[4]</sup>;景鹏和郑伟,2019<sup>[9]</sup>)。以往研究表明,降低第一支柱单位缴费率将必然导致第一支柱养老金替代率下降,但降费并不一定降低整体养老保障水平。多数研究者通过引入其他工具配合养老金降费以维持公共政策目标的平衡,例如公共教育支出(郭庆旺和贾俊雪,2009<sup>[10]</sup>;Kaganovich和Zilcha,2012<sup>[11]</sup>;彭浩然等,2018<sup>[4]</sup>)、国有资本划拨(杨俊和龚六堂,2008<sup>[12]</sup>;景鹏等,2020<sup>[13]</sup>;汪玲等,2021<sup>[7]</sup>)、财政补贴(Kaganovich和Zilcha,1999<sup>[14]</sup>;严成樑,2017<sup>[15]</sup>;景鹏和朱文佩,2021<sup>[16]</sup>)、延迟退休(张熠,2011<sup>[17]</sup>;袁中美,2013<sup>[18]</sup>;周延和杨筱靛,2015<sup>[19]</sup>)。

有研究者探讨了第三支柱个人养老金对整体养老金待遇水平的影响,蹇滨徽等(2021)<sup>[2]</sup>发现下调第一支柱养老金替代率对第三支柱养老金产品的发展有更大的杠杆效应,第三支柱养老金产品的增长不仅弥补了第一支柱基本养老金保障水平下降的空缺,还在一定程度上提升了整体养老金待遇水平。国内研究者对第三支柱个人养老金与总和养老金替代率的关系研究尚处于定性研究阶段。现收现付制在许多国家都成为公共财政的一大负担,在我国,随着人口老龄化和生育率下降,这一养老金体系将不堪重负(封进,2004<sup>[20]</sup>)。安婧(2017)<sup>[21]</sup>认为适当降低第一支柱养老金替代率,通过提升第三支柱个人养老金替代率进行补充可以有效保障养老金待遇水平的稳定。郑秉文

(2022)<sup>[22]</sup>认为在第一支柱养老金替代率逐年走低的形势下,我国养老金体系应从负债型向资产型转变,从现收现付制向基金累积制过渡,扩大第三支柱个人养老金比例从而扩大养老金资产规模,实施市场化投资策略。目前对养老金替代率的研究主要运用精算模型或者局部均衡方法,也有少数文献使用一般均衡方法进行研究。例如,杨再贵(2011)<sup>[23]</sup>使用OLG模型在社会福利最大化条件下,求解最优第一支柱养老金替代率。贾洪波(2021)<sup>[24]</sup>使用一般均衡模型衡量了缴费率的降低对第一支柱养老金替代率的影响。Gonand和Legros(2009)<sup>[25]</sup>、Fehr和Uhde(2013)<sup>[26]</sup>都对现收现付制第一支柱养老金替代率运用一般均衡方法进行研究。

在已有研究基础上,本文使用两期OLG模型进一步探索了第三支柱个人养老金和总和养老金替代率之间的定量关系。现有研究专注于讨论第一支柱养老金替代率,本文探讨的是总和养老金替代率。本文的理论结果主要是基于第一支柱养老金降费和养老金替代率持续下降的现实背景。养老金替代率下降可能会降低老年退休人员的福利水平。由此,不同于现有研究,首先,本文将恒定的养老金替代率作为一个优化目标,测算出了在维持总和养老金替代率不变的前提下,不同第一支柱养老金缴费率对应的最优第三支柱个人养老金缴费率。其次,本文也定量分析了各种不同政策和个人变量下第三支柱个人养老金的发展对未来我国养老金替代率的影响。最后,本文还考察了在经济发展处于非稳态的情况下,不同第一支柱缴费率和第三支柱个人养老金缴费率组合对经济增长和养老金替代率的影响。这样的非稳态分析可以更真实地反映在经济增长尚未达到稳态的经济体中第三支柱和第一支柱个人养老金的影响。

### 三、理论模型

本文构建一个简单的跨期迭代模型,将第三支柱个人养老金和第一支柱养老金纳入分析框架,考察第三支柱个人养老金与工资、消费、第一支柱缴费率、养老金替代率等变量的关系。具体模型设定如下:

#### (一) 个人

假定一个经济中个体只存活两期,其中,第一期为成年期 $t$ ,第二期为老年期 $t+1$ 。个体在成年期进行工作,通过工作获得收入并进行经济决策,消

费、缴纳养老金或者存入个人养老金账户。在老年期个体已退休，不工作但是领取退休金，退休金分为基本养老保险退休金和第三支柱个人养老金。成年期个体人数为  $L_t$ ，由于处于老年期的个体来自上一期的成年期人口，可记为  $L_{t-1}$ ，令第  $t$  期人口增长率为  $n$ ，则有：

$$L_t = L_{t-1}(1+n) \quad (1)$$

在成年期，个人无弹性供给 1 单位有效劳动获得工资收入  $w_t$ ，按一定比例缴纳第一支柱养老金和第三支柱个人养老金，第一支柱养老金缴费率为  $\tau$ ，第三支柱个人养老金缴费率为  $\rho$ 。由于个人养老金回报率高于储蓄利率，本文假设理性人将用于储蓄的可支配收入全部投入个人养老金账户以获得个人效用最大化，个人养老金账户本金为  $S_{t+1}$ ，剩余可支配收入用于成年期消费  $C_t^y$ 。在老年期，个人不再供给劳动，依靠第三支柱个人养老金本金  $I_{t+1}$ ，以及领取的第一支柱养老金  $P_{t+1}$  来满足老年期消费  $C_{t+1}^o$ ， $r_{t+1}$  表示第三支柱个人养老金回报率。由于老年期领取的第一支柱养老金为同期年轻人缴纳，令第  $t+1$  期的成年时期的个体人数为  $L_{t+1}$ ，老年时期的个体人数为  $L_t$ ，第一支柱养老金的约束为：

$$\tau w_{t+1} L_{t+1} = P_{t+1} L_t \quad (2)$$

假设第  $t+1$  期人口增长率保持不变，结合式 (1) 和式 (2)，老年期可以领取的第一支柱养老金为：

$$P_{t+1} = \tau w_{t+1} (1+n) \quad (3)$$

老年期领取的第三支柱个人养老金为成年期第三支柱个人养老金本金和投资回报率：

$$I_{t+1} = S_{t+1} (1+r_{t+1}) \quad (4)$$

综合以上分析，个人在成年期和老年期的消费预算约束分别为：

$$C_t^y = w_t - S_{t+1} - \tau w_t \quad (5)$$

$$C_{t+1}^o = P_{t+1} + S_{t+1} (1+r_{t+1}) = \tau w_{t+1} (1+n) + S_{t+1} (1+r_{t+1}) \quad (6)$$

假设个人一生效用取决于成年期和老年期消费，采用对数线性形式表示为：

$$U_t = \ln C_t^y + \beta \ln C_{t+1}^o \quad (7)$$

其中， $\beta$  表示时间偏好贴现因子。为最大化个体一生的效用，在第  $t$  期，个体将进行选择分配消费水平和第三支柱个人养老金缴费水平。构建拉格朗日函数，求得一阶导数，得到效应最大化的一阶条件：

$$C_{t+1}^o = \beta (1+r_{t+1}) C_t^y \quad (8)$$

由于个人在成年期将所有储蓄投入到第三支柱个人养老金中，在老年期领取的养老金为第一支柱与第三支柱个人养老金的总和。我们将养老金替代率定义为老年期领取的养老金总和与成年期收入的比例，即：

$$\eta = \frac{P_{t+1} + I_{t+1}}{w_t} \quad (9)$$

假设个人在老年期消费掉所有养老金，养老金收入即等于老年期消费，根据养老金替代率的定义，结合式 (5)、(6) 和 (8)，可以得到养老金替代率关于利率和工资的表达式：

$$\frac{C_{t+1}^o}{w_t} = \frac{\beta}{1+\beta} \left[ (1-\tau)(1+r_{t+1}) + \tau \frac{w_{t+1}}{w_t} (1+n) \right] \quad (10)$$

将式 (6) 代入式 (10)，消去  $C_{t+1}^o$  并整理，我们可以得到最优第三支柱个人养老金为：

$$S_{t+1} = \frac{\beta}{1+\beta} (1-\tau) w_t - \frac{\tau w_{t+1} (1+n)}{(1+r_{t+1})(1+\beta)} \quad (11)$$

式 (11) 说明最优第三支柱个人养老金与第一支柱养老金成反比，与时间偏好贴现因子成正比。

### (二) 企业

在完全竞争市场上，假设企业的生产函数为柯布-道格拉斯形式且满足稻田条件，企业使用两种生产要素，资本要素与劳动要素。这样，企业生产函数的具体形式可以表示为：

$$Y_t = K_t^\alpha (A_t L_t)^{1-\alpha} \quad (0 < \alpha < 1) \quad (12)$$

其中， $K$  为资本， $L$  为劳动， $A$  表示劳动效率， $t$  表示时间， $\alpha$  为资本对总产量的贡献， $1-\alpha$  为有效劳动对总产量的贡献。假设劳动效率按照外生速率  $g$  增长，则有  $A_{t+1}/A_t = 1+g$ 。令单位有效劳动的资本为  $k_t = K_t/A_t L_t$ ，以及单位有效劳动的产出为  $y_t = Y_t/A_t L_t$ ，式 (12) 中生产函数的密集形式可以表示为：

$$y_t = k_t^\alpha \quad (13)$$

利润函数为总产出减去资本投入的利息和总工

资，可以表示为：

$$\pi_t = Y_t - q_t K_t - w_t L_t \quad (14)$$

其中， $\pi_t$  为  $t$  时期的利润， $q_t$  为  $t$  时期的利率， $w_t$  为  $t$  时期的工资。利润最大化时的实际工资和实际利率分别为  $Y_t$  对于资本和劳动的一阶导数，可以表示为：

$$q_t = \alpha k_t^{\alpha-1} \quad (15)$$

$$w_t = (1-\alpha) k_t^\alpha A_t \quad (16)$$

### (三) 市场出清条件

资本市场出清条件为厂商使用的资本存量等于上一期年轻人的第三支柱个人养老金储蓄，即：

$$K_{t+1} = S_{t+1} L_t \quad (17)$$

用单位有效劳动资本的形式表示资本出清条件，式 (17) 可以写成：

$$k_{t+1} = \frac{S_{t+1}}{(1+n)A_{t+1}} \quad (18)$$

商品市场出清条件为当期产出和资本存量等于当期消费和下一期的资本存量之和，即：

$$Y_t + (1-\delta)K_t = C_t^y L_t + C_t^o L_{t-1} + K_{t+1} \quad (19)$$

其中  $\delta$  为资本折旧率。使用单位有效劳动产出和单位有效劳动资本表示该条件，式 (19) 可以写成：

$$y_t + (1-\delta)k_t = \frac{C_t^y}{A_t} + \frac{C_t^o}{(1+n)A_t} + (1+n)k_{t+1} \frac{A_{t+1}}{A_t} \quad (20)$$

将式 (5)、(6)、(13)、(16) 和 (18) 代入式 (20) 并化简，可以得到实际利率在满足市场出清条件下的表达式：

$$r_t = \alpha k_t^{\alpha-1} - \delta \quad (21)$$

将式 (16) 和 (21) 代入式 (10) 右边，可以得到养老金替代率关于单位有效劳动资本的表达式：

$$\frac{C_{t+1}^o}{w_t} = \frac{\beta}{1+\beta} \left[ (1-\tau)(1+\alpha k_{t+1}^{\alpha-1} - \delta) + \tau \left( \frac{k_{t+1}}{k_t} \right)^\alpha \left( \frac{A_{t+1}}{A_t} \right) (1+n) \right] \quad (22)$$

式 (22) 中括号内第一项可以看成是第三支柱个人养老金收益率对替代率的影响，第二项可以看成是第一支柱缴费率对替代率的影响。式 (22) 说明养老金替代率主要取决于时间偏好贴现因子  $\beta$  和第一支柱缴费率  $\tau$ 。第一支柱缴费率对养老金替代率有两

种相互对抗的效用：一方面第一支柱缴费率会对第三支柱个人养老金的替代率产生负面影响；另一方面，提高第一支柱缴费率可以直接提高第一支柱养老金的替代率。当经济增长率、劳动效率增长率和人口增长率之积大于实际利率时，第一支柱缴费率对替代率有积极贡献。时间偏好贴现因子对养老金替代率则主要表现为促进作用。除此之外，这两个变量与养老金替代率还有更为复杂的关系。在系统动态变化过程中，这两个参数还可以通过影响经济增长率  $k_{t+1}/k_t$  和单位有效劳动资本存量  $k_t$  来影响养老金替代率。

### (四) 一般均衡求解

将式 (16) 和 (21) 代入式 (11)，再将其代入式 (18)，我们可以得到描述整个经济系统的单位有效劳动的资本动态积累方程：

$$k_{t+1} = \left[ \beta \frac{(1-\tau)k_t^\alpha A_t}{1+n} - \frac{\tau k_{t+1}^\alpha}{1+\alpha k_{t+1}^{\alpha-1} - \delta} \right] \frac{1-\alpha}{1+\beta} \quad (23)$$

对于给定初始单位有效劳动资本，我们可以由式 (23) 得到下一期单位有效劳动资本。将两期单位有效劳动资本代入式 (22)，我们就可以得到经济动态变化过程中的养老金替代率。

当系统处于稳态时，单位有效劳动资本保持不变，即  $k_{t+1} = k_t$ ，将该条件代入式 (23)，我们可以算得稳态时的单位有效劳动资本为：

$$k^* = \left( \frac{-(\beta'(1-\delta) - \tau - \alpha'\alpha) + \sqrt{(\beta'(1-\delta) - \tau - \alpha'\alpha)^2 + 4\beta'\alpha\alpha'(1-\delta)}}{2\alpha\beta'} \right)^{\frac{1}{\alpha-1}} \quad (24)$$

其中， $\beta' = \beta \frac{1-\tau}{1+n} \frac{A_t}{A_{t+1}}$ ， $\alpha' = \frac{1+\beta}{1-\alpha}$ 。

稳态时的总和养老金替代率为：

$$\frac{C_{t+1}^o}{w_t} = \frac{\beta}{1+\beta} \left[ (1-\tau)(1-\delta) + \left( \frac{-(\beta'(1-\delta) - \tau - \alpha'\alpha) + \sqrt{(\beta'(1-\delta) - \tau - \alpha'\alpha)^2 + 4\beta'\alpha\alpha'(1-\delta)}}{2\beta} + \tau \right) (1+n) \frac{A_{t+1}}{A_t} \right] \quad (25)$$

第一支柱养老金替代率为：

$$\frac{P_{t+1}}{w_t} = \frac{\tau w_{t+1} (1+n)}{w_t} = \tau (1+n) \left( \frac{k_{t+1}}{k_t} \right)^\alpha \left( \frac{A_{t+1}}{A_t} \right) \quad (26)$$

#### 四、数值模拟与结果分析

##### (一) 参数选取

本文假设个体在30岁时进入成年期工作,因为每一期的间隔为30年,个体在60岁时进入老年期(郭凯明和龚六堂,2012<sup>[27]</sup>;景鹏和郑伟,2019<sup>[9]</sup>;彭浩然等,2018<sup>[4]</sup>)。在已有文献的基础上,对理论模型中的参数进行取值,通过数值模拟考察第三支柱个人养老金对养老金替代率及其经济增长率的影响。

时间偏好贴现因子 $\beta$ 。时间偏好因子越大说明,相较于成年期,个人越重视老年期的消费。既有文献对时间偏好因子的取值在0.98~0.99之间,假设每期间隔为30年,则 $\beta$ 的取值在0.55~0.74之间(景鹏和郑伟,2019<sup>[9]</sup>)。鉴于此,本文设定时间偏好因子为0.7。同时,本文还分析了时间偏好贴现因子在不同水平下其他变量如何变化,分别为 $\beta=0.6$ 和 $\beta=0.8$ 。

资本产出弹性 $\alpha$ 。已有文献大多将其取值在0.3~0.5之间,由于物质资本不断增长,资本产出弹性将不断下降,将资本产出弹性设定为0.31(景鹏和郑伟,2019<sup>[9]</sup>)。鉴于此,本文参考现有研究的做法,取 $\alpha$ 基准值为0.31。

人口增长率 $n$ 。根据国家统计局发布数据,全国年平均人口增长率为0.3%,假定每一期由30年构成,则设定 $n=0.1$ <sup>①</sup>。

资本折旧率 $\delta$ 。现有文献对资本折旧率的测算大部分设定为5%(王小鲁,2000<sup>[28]</sup>;郭庆旺和贾俊雪,2004<sup>[29]</sup>;孙辉等,2010<sup>[30]</sup>;陈昌兵,2020<sup>[31]</sup>)。考虑到未来中国企业进入成熟发展期,存续时间延长,折旧率将下降。据此本文将年度资本折旧率设定为4%,即设定 $\delta$ 的基准值为0.7。

劳动效率增长率 $g$ 。我们假设劳动效率恒定不变,即 $g=1$ 。

##### (二) 数值模拟

根据公式和相关参数设定,通过Python编程求解得到最优第三支柱个人养老金缴费率水平下养老金替代率变化的一般均衡效应。我国第一支柱养老金

费率自1997年建立以来一直都处于较高水平,总体缴费率达到了28%甚至更高。过高的第一支柱缴费率对第二、第三支柱在资源分配上有一定的挤占效应,也从一定程度上降低了第二、第三支柱的参保意愿。近年来,中国政府加大了减税降费的力度,到2019年,国务院办公厅引发《降低社会保险费率综合方案的通知》,第一支柱缴费率总体降至24%,其中社会统筹账户缴费率降至16%,个人账户缴费率保持8%不变。尽管如此,我国第一支柱养老金缴费率仍处于较高水平,有一定的下降空间。

通过数值模拟,假定养老金替代率处于75%的水平<sup>②</sup>,第一支柱养老金缴费率从24%开始以0.5%的间隔下降,直到下降至16%。当养老金替代率水平恒定不变时,在稳态条件下,对于不同的第一支柱缴费率,我们可以通过式(22)和式(24)计算出来相应的时间偏好因子。我们将保持养老金替代率恒定,不同第一支柱缴费率情况下,均衡系统中其余各变量如单位有效劳动资本、单位有效劳动工资、账户收益率以及第三支柱个人养老金缴费率对应的值分别在表1中给出。从表中可以看到,保持养老金替代率恒定的条件下,减少第一支柱缴费率会增加时间偏好贴现因子。第一支柱缴费率从24%降至16%会影响时间偏好贴现因子从0.11增加至0.89。由于第一支柱缴费率的下降导致个人预期第一支柱养老金替代率下降,从预防动机来看,个人将减少成年期消费支出并进行储蓄来应对老年期消费。同时,稳态时单位有效劳动资本量从0.70%增长到11.32%,增加了10.62%,单位有效劳动领取的工资随着第一支柱缴费率的降低提升了20%。一方面,缴费率的降低提升了单位有效劳动资本量从而增加单位有效劳动产出,最终提高稳态时经济的发展水平。另一方面,缴费率的降低增加了成年期的可支配收入,促使个人同时增加成年期消费和个人养老金投资,从而促进了经济的长期增长。总之,降低第一支柱养老金缴费率可以通过增加第三支柱个人养老金缴费率从而将养老金转化为长期资本,实现养老金与资本市场的良性互动,促进经济增长并维持养老金替代率不变。

① 人口增长率为0.3%,假定30年为一期,则 $n=(1+0.3\%)^{30}=0.1$ 。

② 在第三支柱发达的国家,养老金替代率通常在75%左右,故本文选取总和养老金替代率在75%水平进行模拟。



表1 养老金替代率在75%时第一支柱缴费率 $\tau$ 变化的一般均衡效应

第一支柱缴费率 $\tau$	时间偏好贴现因子 $\beta$	单位有效劳动资本	单位有效劳动工资	第三支柱个人养老金收益率	第三支柱个人养老金缴费率
0.240	0.105 4	0.007 0	0.16	9.01	0.04
0.235	0.146 2	0.011 0	0.18	6.35	0.06
0.230	0.187 8	0.015 6	0.20	4.83	0.07
0.225	0.230 2	0.020 5	0.22	3.85	0.09
0.220	0.273 5	0.025 9	0.23	3.17	0.11
0.215	0.317 7	0.031 7	0.25	2.66	0.13
0.210	0.362 9	0.037 8	0.26	2.27	0.15
0.205	0.409 3	0.044 1	0.27	1.96	0.17
0.200	0.456 8	0.050 8	0.29	1.71	0.19
0.195	0.505 8	0.057 8	0.30	1.51	0.21
0.190	0.556 1	0.065 0	0.31	1.33	0.23
0.185	0.608 1	0.072 5	0.32	1.18	0.25
0.180	0.661 7	0.080 2	0.33	1.06	0.27
0.175	0.717 3	0.088 1	0.34	0.94	0.28
0.170	0.774 9	0.096 2	0.35	0.84	0.30
0.165	0.834 7	0.104 6	0.36	0.76	0.32
0.160	0.896 9	0.113 2	0.36	0.68	0.34

表2描绘了不同时间偏好贴现因子设定下,稳态时系统中各变量的变化情况。单位有效劳动资本量随第一支柱缴费率的增加而下降,由于稳态单位有效劳动产出由稳态单位有效劳动资本量决定,单位有效劳动资本量越低,单位有效劳动产出也越低,导致经济增长率下降。数值模拟结果说明高水平的第一支柱缴费率将制约经济发展水平。表2也列示了在时间偏好贴现因子恒定时,不同的第一支柱缴费率和第三支柱个人养老金缴费率组合下养老金替代率的变化。 $\beta=0.6$ 时,以目前第一支柱缴费率水平 $\tau=0.24$ ,第三支柱个人养老金缴费率为0.21时,养老金替代率即可达到80%。当第一支柱缴费率继续下调,从24%下调至2%,下调幅度为22%。同时,第三支柱个人养

老金缴费率从21%增长至36%,增长幅度为15%,第一支柱缴费率的下调幅度大于第三支柱个人养老金缴费率的增长幅度,但养老金替代率最低也能达到60%。由于第三支柱个人养老金进入资本市场享有收益率,当收益率高于人口增长率时,其对养老金替代率的贡献将高于第一支柱养老金。如表2所示,在我们的参数设定下,第三支柱个人养老金收益率均高于人口增长率。进一步研究,我们假设时间偏好因子增加,具体数值如表2所示。可以看到,相较于 $\beta=0.6$ 时,当 $\beta=0.7$ 和 $\beta=0.8$ 时,在同一第一支柱缴费率水平下,单位有效劳动资本、单位有效劳动工资、第三支柱个人养老金缴费率和养老金替代率都增加了。所以,更偏好第二期消费时,能达到更高的养老金替代率。

表2 最优第三支柱个人养老金缴费率水平下养老金替代率变化的一般均衡效应

时间偏好贴现因子	第一支柱缴费率	单位有效劳动资本	单位有效劳动工资	第三支柱个人养老金收益率	第三支柱个人养老金缴费率	养老金替代率
$\beta=0.6$	0.24	0.059 5	0.30	1.46	0.21	0.80
	0.22	0.063 6	0.31	1.36	0.22	0.78
	0.20	0.068 0	0.31	1.27	0.23	0.76
	0.18	0.072 7	0.32	1.18	0.25	0.74

续前表

时间偏好 贴现因子	第一支柱 缴费率	单位有效 劳动资本	单位有效 劳动工资	第三支柱个人 养老金收益率	第三支柱个人 养老金缴费率	养老金替代率
$\beta=0.6$	0.16	0.077 6	0.33	1.10	0.26	0.73
	0.14	0.082 8	0.33	1.02	0.27	0.71
	0.12	0.088 3	0.34	0.94	0.29	0.69
	0.10	0.094 2	0.34	0.87	0.30	0.67
	0.08	0.100 4	0.35	0.80	0.31	0.65
	0.06	0.107 0	0.36	0.73	0.33	0.64
	0.04	0.114 0	0.36	0.67	0.34	0.62
	0.02	0.121 3	0.37	0.61	0.36	0.60
$\beta=0.7$	0.24	0.069 7	0.31	1.24	0.24	0.81
	0.22	0.074 4	0.32	1.15	0.25	0.79
	0.20	0.079 4	0.33	1.07	0.26	0.77
	0.18	0.084 7	0.33	0.99	0.28	0.75
	0.16	0.090 3	0.34	0.92	0.29	0.73
	0.14	0.096 1	0.35	0.85	0.30	0.72
	0.12	0.102 4	0.35	0.78	0.32	0.70
	0.10	0.108 9	0.36	0.72	0.33	0.68
	0.08	0.115 9	0.37	0.66	0.35	0.66
	0.06	0.123 2	0.37	0.60	0.36	0.65
	0.04	0.130 9	0.38	0.55	0.38	0.63
	0.02	0.139 0	0.39	0.49	0.39	0.61
$\beta=0.8$	0.24	0.079 4	0.33	1.07	0.26	0.82
	0.22	0.084 7	0.33	0.99	0.28	0.80
	0.20	0.090 2	0.34	0.92	0.29	0.78
	0.18	0.096 1	0.35	0.85	0.30	0.76
	0.16	0.102 2	0.35	0.78	0.32	0.74
	0.14	0.108 7	0.36	0.72	0.33	0.73
	0.12	0.115 5	0.37	0.66	0.35	0.71
	0.10	0.122 7	0.37	0.60	0.36	0.69
	0.08	0.130 3	0.38	0.55	0.38	0.67
	0.06	0.138 2	0.39	0.50	0.39	0.66
	0.04	0.146 6	0.39	0.45	0.41	0.64
	0.02	0.155 4	0.40	0.40	0.43	0.62

表3描绘了在不同第一支柱缴费率水平下，系统中各个变量的变化情况。我们设定第一支柱养老金缴费率为24%、20%以及16%，测算了在保持现有第一支柱缴费率的基础上引入第三支柱个人养老金将如何影响总和养老金替代率和经济增长率。并且预测了未来继续下调第一支柱缴费率，各变量如何变化。在各第一支柱缴费率水平下，时间偏好贴现因子均正比于第三支柱个人养老金缴费率，说明第三支柱个人养老金缴费率的高低取决于个人对第二期消费的重视程度。在第一支柱缴费率24%的水平下，随着时间偏好因子的变化，第三支柱个人养老金缴费率在4%~26%时，养老金替代率可达75%~82%；在第一支柱缴费率20%的水平下，第三支柱个人养老金缴费率在4%~29%时，养老金替代率则达到了71%~78%；在第一支柱缴费率16%的水平下，第三支柱个人养

老金缴费率在5%~32%时，养老金替代率则达到了66%~74%。可以看到，不论第一支柱缴费率在何种水平下，养老金替代率都随第三支柱个人养老金缴费率增加而增加。并且，稳态单位有效劳动资本量随着第三支柱个人养老金缴费率的增加而增长，说明第三支柱个人养老金缴费率的增长能促进单位有效劳动产出，从而提高经济发展水平。第三支柱个人养老金体系的建立使个人自愿减少当期消费并将资金存至个人养老金账户中进行购买能实现养老保障功能的金融产品，通过市场化运作也可以实现个人养老金的保值和增值。另外，第三支柱个人养老金作为长期资产对于资本市场而言，可以扩大资本市场规模、提升市场流动性以及平抑市场波动，促进资本市场结构完善并长期健康发展。而健康的资本市场也可以助力个人养老金的长期稳定增长。

表3 不同第一支柱缴费率水平下养老金替代率变化的一般均衡效应

第一支柱缴费率	时间偏好贴现因子	单位有效劳动资本	单位有效劳动工资	第三支柱个人养老金收益率	第三支柱个人养老金缴费率	养老金替代率
$\tau = 0.24$	0.8	0.079 4	0.33	1.07	0.26	0.82
	0.7	0.069 7	0.31	1.24	0.24	0.81
	0.6	0.059 5	0.30	1.46	0.21	0.80
	0.5	0.048 9	0.28	1.78	0.18	0.79
	0.4	0.038 0	0.26	2.26	0.15	0.78
	0.3	0.027 0	0.24	3.06	0.11	0.77
	0.2	0.016 2	0.20	4.67	0.08	0.76
	0.1	0.006 5	0.15	9.50	0.04	0.75
$\tau = 0.20$	0.8	0.090 2	0.34	0.92	0.29	0.78
	0.7	0.079 4	0.33	1.07	0.26	0.77
	0.6	0.068 0	0.31	1.27	0.23	0.76
	0.5	0.056 1	0.29	1.55	0.20	0.75
	0.4	0.043 8	0.27	1.98	0.17	0.74
	0.3	0.031 2	0.25	2.70	0.13	0.73
	0.2	0.018 9	0.21	4.13	0.09	0.72
	0.1	0.007 6	0.16	8.44	0.04	0.71

续前表

第一支柱 缴费率	时间偏好 贴现因子	单位有效 劳动资本	单位有效 劳动工资	第三支柱个人 养老金收益率	第三支柱个人 养老金缴费率	养老金替代率
$\tau=0.16$	0.8	0.102 2	0.35	0.78	0.32	0.74
	0.7	0.090 3	0.34	0.92	0.29	0.73
	0.6	0.077 6	0.33	1.10	0.26	0.73
	0.5	0.064 3	0.31	1.35	0.23	0.72
	0.4	0.050 4	0.29	1.73	0.19	0.71
	0.3	0.036 1	0.26	2.37	0.15	0.69
	0.2	0.021 9	0.22	3.65	0.10	0.68
	0.1	0.008 9	0.17	7.49	0.05	0.66

### (三) 非稳态分析

前文主要进行了系统的稳态分析，稳态分析是对政策产生的最终长期效应进行研究，而非稳态分析可以研究系统对政策变动的及时响应。当系统处于稳态时，单位有效劳动资本增长率保持不变。如图1和图2所示，当单位有效劳动资本增长率等于1时均衡系统达到稳态。但在实际中，经济均衡系统可能处于非稳态，所以进一步分析非稳态条件能提供更实际的资本变化情况。给定初始资本，我们可以通过式(23)求得下一期单位有效劳动资本，从而得到非稳态情况的单位有效劳动资本增长率。图1描绘了不同时间偏好贴现因子 $\beta$ 下单位有效劳动资本增长率随不同初始单位有效劳动资本的变化。可以看到，在任意相同初始单位有效劳动资本下，单位有效劳动资本增长率随 $\beta$ 增大。这是因为个人对第二期消费的重视程度随时间偏好贴现因子

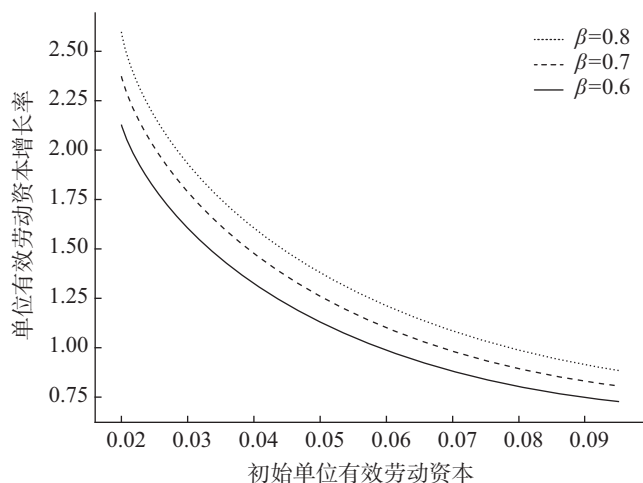


图1 不同 $\beta$ 下单位有效劳动资本增长率随初始单位有效劳动资本的变化

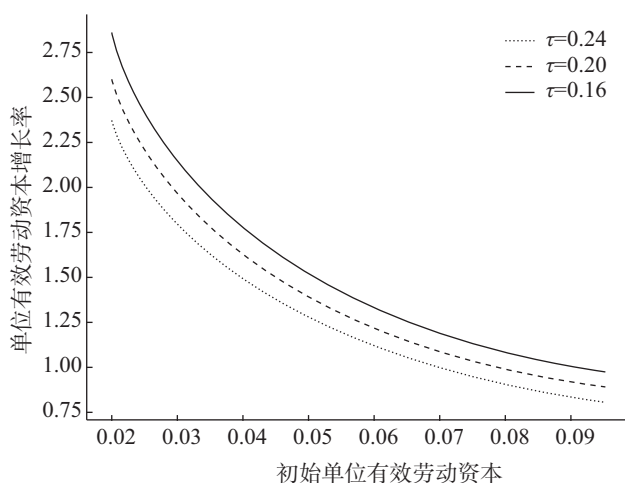


图2 不同第一支柱缴费率水平下单位有效劳动资本增长率随初始单位有效劳动资本的变化

增大而增大，初始单位有效劳动资本相同时，个人工资水平相等，时间偏好贴现因子增大导致个人储蓄率增高，最终用于投资的资本越多，从而促进单位有效劳动资本增长。

相反地，第一支柱缴费率水平与单位有效劳动资本增长率成负相关关系。图2描绘了第一支柱缴费率在24%、20%以及16%时对应的单位有效劳动资本增长率的变化。在图2中，点线表示的为高第一支柱缴费率，短横线表示的为中第一支柱缴费率，实线表示的为低第一支柱缴费率。可以看到，在同一初始单位有效劳动资本下，实线对应的单位有效劳动增长率曲线在虚线对应的单位有效劳动增长率曲线之上，单位有效劳动资本增长率随着第一支柱缴费率的降低而提升。这说明降低第一支柱缴费率可以提高单位有效劳动资本增长率。

图3描绘了第一支柱缴费率在24%和16%以及

时间偏好因子在 0.6、0.7、0.8 的水平下替代率随初始单位有效劳动资本的变化。可以看到，不同组合下替代率均随单位有效劳动资本的增高而走低。由于单位有效劳动资本走高意味着经济在不断发展，图 3 验证了我国替代率逐年走低这一事实。图 3 表明在多种时间偏好贴现因子设定下，降低第一支柱缴费率的同时增加时间偏好贴现因子可以提高总和养老金替代率。由此可见，大力发展第三支柱个人养老金制度，减少第一支柱缴费率，可以在不降低总和养老金替代率的情况下促进经济发展，适合我国国情。

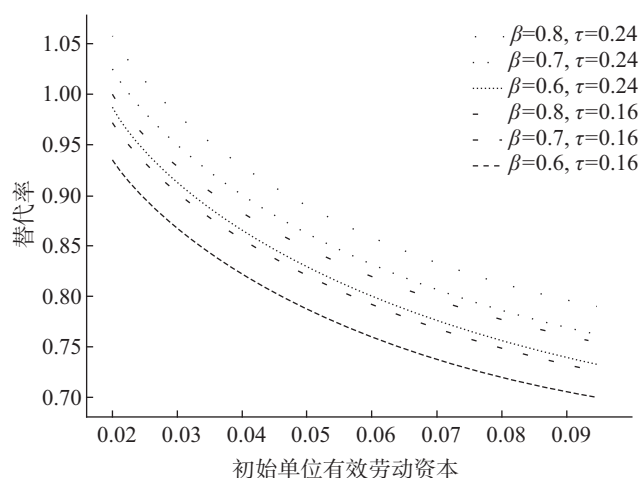


图 3 不同第一支柱缴费率和不同时间偏好因子下替代率随初始单位有效劳动资本的变化

### 五、结论

本文构建一个世代交叠一般均衡模型，在第一支柱养老金降费 and 养老金替代率逐年走低的大背景下，将第三支柱个人养老金纳入分析框架，考察第三支柱个人养老金缴费率与工资、消费、第一支柱缴费率、养老金替代率等变量的关系，并结合中国的现实经济参数进行了数值模拟，得到如下结论：

第一，高水平的第一支柱缴费率将制约经济发展水平。降低第一支柱缴费率可以提升单位有效劳动资本量从而增加单位有效劳动产出，并且，缴费率的下降增加了成年期的可支配收入，由此带来的收入效应促使个人同时增加成年期消费和个人养老金投资，从而促进了经济的长期增长。

第二，第三支柱个人养老金能大幅提高养老金替代率。当  $\beta=0.6$  时，第一支柱缴费率维持 24% 的比例，相应的第三支柱个人养老金缴费率为 21%，养老金替代率可达到 80%；当第一支柱缴费率降至 2%

时，第三支柱个人养老金缴费率增加至 36% 时也可达到 60% 的养老金替代率水平。若时间偏好贴现因子更高，加入第三支柱个人养老金作为补充对养老金替代率的提升效果更好。

第三，发展第三支柱个人养老金并降低第一支柱缴费率能实现经济增长率与养老金替代率的同步提升。以目前第一支柱缴费水平  $\tau=0.24$  测算，第三支柱个人养老金缴费率在 4%~26% 时，养老金替代率可达 75%~82%；若降费至 20%，第三支柱个人养老金缴费率在 4%~29% 时，养老金替代率在 71%~78% 的范围；若降费至 16%，第三支柱个人养老金缴费率在 5%~32% 时，养老金替代率也可达到 66%~74%。并且，稳态单位有效劳动资本量随着第三支柱个人养老金缴费率的增加而增长，说明第三支柱个人养老金缴费率的增长能提高经济发展水平。

最后，我们对模型进行了非稳态分析，验证了在多种时间偏好贴现因子设定下，降低第一支柱缴费率对经济增长的影响以及对总和养老金替代率的影响，所得结果与稳态时一致。说明减少第一支柱缴费率的同时增加第三支柱缴费率可以在不降低总和养老金替代率的情况下促进经济发展。

本文从宏观层面考虑了第三支柱个人养老金对养老金替代率的影响。对于在第一支柱养老金降费与第一支柱养老金替代率不断下降的背景下，如何通过建立第三支柱个人养老金制度将总和养老金替代率维持在足够保障参保者基本生活水平的程度的同时促进经济增长有政策指导意义。

本文基于上述结论提出以下政策建议：我们认为，为实现全体人民共同富裕的新时代要求，应对人口老龄化加剧趋势，积聚长期性的养老基金，建设可持续发展的第三支柱个人养老金制度是完善我国养老金体系的迫切选择。首先，降低第一支柱养老金缴费比例的同时加快第三支柱个人养老金制度的顶层设计。由于降低第一支柱养老金缴费比例，可以有效提升经济活力，由此带来的养老金替代率降低问题则可以通过提升居民第三支柱个人养老金的参与度来解决。除此之外，我们还发现发展第三支柱个人养老金并不会减缓经济发展，反而第三支柱个人养老金贡献的长期资本能促进经济的发展，有效增强我国养老金体系的稳定性和可持续性。由于第三支柱个人养老金有维持经济发展和养老金替代率水平的重要作

用,我们建议出台一系列增加第三支柱个人养老金参与度的惠民政策。首先,我们可以通过促进各支柱养老金的灵活衔接和转换贯通来增加第三支柱个人养老金的参与度。其次,我们可以加快构建统一的个人养老金账户制度,三支柱间通过信息共享以

实现各支柱账户联通、互转,便于居民统筹规划和动态安排。同时,也有利于提高第三支柱个人养老金的筹资能力,促进三支柱养老金体系良性发展。第三,我们还可以通过对第三支柱个人养老金收入提供税收优惠来提高其参与度。

## 参考文献

- [1] 贾洪波,温源.基本养老金替代率优化分析[J].中国人口科学,2005(1):83-89.
- [2] 蹇滨徽,杨亮,林义.多层次养老保险制度下家庭商业养老保险需求与养老金替代率研究[J].中国软科学,2021(5):38-48.
- [3] 边恕,孙雅娜.企业职工养老保险“保基本”的内涵回归及调整方案[J].社会政策研究,2018(1):85-96.
- [4] 彭浩然,邱桓沛,朱传奇,李昂.养老保险缴费率、公共教育投资与养老金替代率[J].世界经济,2018(7):148-168.
- [5] 曾益,张心洁,刘凌晨.从“单独二孩”走向“全面二孩”:中国养老金支付危机能破解吗[J].财贸经济,2016(7):133-146.
- [6] 景鹏,周佩,胡秋明.养老保险缴费率、经济增长与养老金替代率——兼论政策缴费率与实际缴费率的关系[J].经济科学,2020(6):124-136.
- [7] 汪玲,邱桓沛,申曙光.国有资本划拨养老保险与公共政策目标治理[J].财贸经济,2021(4):53-66.
- [8] Cipriani G P, Pascucci F. Pension Policies in a Model with Endogenous Fertility [J]. Journal of Pension Economics and Finance, 2020, 19 (1): 109-125.
- [9] 景鹏,郑伟.养老保险缴费率、财政支出结构与经济增长[J].世界经济,2019(12):121-144.
- [10] 郭庆旺,贾俊雪.公共教育政策、经济增长与人力资本溢价[J].经济研究,2009(10):22-35.
- [11] Kaganovich M, Zilcha I. Pay-as-you-go or Funded Social Security? A General Equilibrium Comparison [J]. Journal of Economic Dynamics and Control, 2012, 36 (4): 455-467.
- [12] 杨俊,龚六堂.国有资本收入对养老保险的划拨率研究[J].金融研究,2008(11):46-55.
- [13] 景鹏,王媛媛,胡秋明.国有资本划转养老保险基金能否破解降费“不可能三角”[J].财政研究,2020(2):80-95.
- [14] Kaganovich M, Zilcha I. Education, Social Security, and Growth [J]. Journal of Public Economics, 1999, 71 (2): 289-309.
- [15] 严成樑.延迟退休、财政支出结构调整与养老金替代率[J].金融研究,2017(9):51-66.
- [16] 景鹏,朱文佩.财政补贴能否破解养老保险降费的“不可能三角”[J].财经科学,2021(8):64-77.
- [17] 张熠.延迟退休年龄与养老保险收支余额:作用机制及政策效应[J].财经研究,2011(7):4-16.
- [18] 袁中美.延迟退休与养老金替代率的探讨[J].人口与经济,2013(1):101-106.
- [19] 周延,杨筱靓.合意养老金替代率下退休年龄的测算与优化——以上海市为例[J].现代财经(天津财经大学学报),2015(4):13-21.
- [20] 封进.中国养老保险体系改革的福利经济学分析[J].经济研究,2004(2):55-63.
- [21] 安婧.OECD与中国养老金替代率水平的比较分析[J].劳动保障世界,2017(9):20-21.
- [22] 郑秉文.养老金三支柱理论嬗变与第三支柱模式选择[J].华中科技大学学报(社会科学版),2022(2):20-37.
- [23] 杨再贵.不定寿命条件下城镇公共养老金最优替代率的理论与实证研究[J].管理评论,2011(2):28-32.
- [24] 贾洪波.降低单位缴费率对城镇人口养老金替代率的一般均衡效应[J].数量经济技术经济研究,2021(11):103-121.
- [25] Gonand F, Legros F. Pension Reforms in France: Impact on Growth and Inter-Generational Redistributive Effects [J]. Geneva Papers on Risk and Insurance Issues and Practice, 2009, 34 (4): 639-659.
- [26] Fehr H, Uhde J. On the Optimal Design of Pension Systems [J]. Empirica, 2013, 40 (3): 457-482.
- [27] 郭凯明,龚六堂.社会保障、家庭养老与经济增长[J].金融研究,2012(1):78-90.
- [28] 王小鲁.中国经济增长的可持续性 & 制度变革[J].经济研究,2000(7):3-15.
- [29] 郭庆旺,贾俊雪.中国潜在产出与产出缺口的估算[J].经济研究,2004(5):31-39.
- [30] 孙辉,支大林,李宏瑾.对中国各省资本存量的估计及典型性事实:1978—2008[J].广东金融学院学报,2010(3):103-116.
- [31] 陈昌兵.可变折旧率的另一种估计方法——基于中国各省份资本折旧的极大似然估计[J].经济研究,2020(1):49-64.

(责任编辑:李 晟 张安平)

# 劳动雇佣关系中互惠的触发机制和影响因素探究

## ——基于礼物交换实验结果的分析

Reciprocal Trigger Mechanism of Manager-employee Relationship and Its Influencing Factors: An Analysis Based on the Gift Exchange Experiment's Results

武志伟 殷明

WU Zhi-wei YIN Ming

**[摘要]** 劳动力市场中雇主与雇员之间的隐性互惠偏好能够导致雇佣关系中互惠行为的发生,但现有文献缺乏对这种互惠行为触发条件和影响因素的深入研究。笔者通过双边礼物交换实验检验了劳动力市场雇佣关系中互惠的触发机制以及影响互惠关系的因素,并采用个体固定效应模型对实验数据进行了分析。结果显示:雇主与雇员之间存在着稳健的互惠行为倾向,但在不同条件下双方的互惠程度存在差异;损失风险下雇佣双方互惠的触发需要二重条件,既需要赠礼水平的提高,也需要互惠能够提高自身收益;雇主和雇员交流的效率会影响互惠程度的变化;重新匹配有利于雇主和雇员在新的雇佣环境中采取更为主动的互惠行为;雇员生产力水平的提高对于互惠程度的提升有显著促进作用。本研究拓展了劳动雇佣关系中关于互惠行为的学术探讨边界,研究结论丰富了劳动雇佣关系和互惠偏好理论方面的相关文献,对于通过互惠机制解决我国劳动力市场的摩擦问题也有一定的现实意义。

**[关键词]** 雇佣关系 互惠 触发机制 礼物交换实验

**[中图分类号]** F272.92 **[文献标识码]** A **[文章编号]** 1000-1549 (2022) 11-0089-13

**Abstract:** In the labor market, both managers and employees have implicit preferences for reciprocal behavior, leading to the occurrence of the reciprocal behavior in the manager-employee relationship, but the existing literature is still lack of in-depth research on the triggering conditions and influencing factors. In the laboratory, the triggering of the reciprocity mechanism has been revealed through a gift-exchange experiment, and the relevant factors affecting the employment relationship have also been studied. Experimental datas have been analysed through individual fixed effects model. The experimental results show that there is a robust reciprocal behavior tendency between managers and employees, but under different conditions, the degree of reciprocity between the two parties is different; the triggering of the reciprocal mechanism between managers and employees under the risk of loss requires dual conditions, not only the improvement of gift levels, but also their own benefits; the change in the degree of reciprocity depends on the communication between managers and employees; re-matching is beneficial for both managers and employees to take more proactive reciprocal behaviors in the new employment environment; the improvement of employee productivity has a significant effect on the improvement of reciprocity. Based on the research conclusions of the experiment, this study expands the academic discussion boundary of reciprocal behavior in the employment relationship, enriches the references relevant to manager-employee relationship and reciprocal preference theory. The reciprocity mechanism of gift exchange experiment has certain guiding significance for solving the friction problem of manager-employee relationship in our country's labor market.

**Key words:** Manager-employee relationship Reciprocity Trigger mechanism Gift-exchange experiment

**[收稿日期]** 2022-07-12

**[作者简介]** 武志伟,男,1972年7月生,南京大学商学院经济学系副教授,管理学博士,主要研究方向为实验经济学、神经经济学;殷明,男,2000年2月生,南京大学商学院硕士研究生,研究方向为实验经济学。本文通讯作者为武志伟,联系方式为 wuzhiwei@nju.edu.cn。

**[基金项目]** 国家自然科学基金面上项目“大数据视角下城市群要素流动的三维测度、理论机制与壁垒治理——以长三角地区为例”(项目编号:72173063)。

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

众所周知,劳动力市场中的雇佣关系是雇主和雇员追求自身利益最大化过程中一系列博弈行为产生的结果。新古典经济学派认为:理性经济人假设条件下企业只付给工人市场出清时的工资,工人则提供最小化程度的努力。事实上,真实的个体经济行为会受到各种复杂心理动机的影响。劳动力市场的行为主体双方存在“礼物交换”这一互惠偏好,即雇主倾向于提供比雇员机会成本更高的薪酬(超出市场出清水平的工资)作为礼物,而员工则以超过最小努力值的表现作为回礼(Akerlof, 1982<sup>[1]</sup>)。显然,这种相互赠礼的互惠行为明显偏离了“经济人”假设,可以看作是雇主与雇员之间的合作,而这种合作行为有利于博弈双方取得更大的收益,提高社会的总体福利水平。

现有研究证实,这种由“礼物交换”引发的互惠行为能够为雇佣双方以及社会带来很多积极的影响,一是雇主与雇员之间的互惠有利于营造和谐稳定的雇佣环境;二是这种互惠行为有利于打破博弈僵局,提高社会整体的生产效率(Fehr等, 2009<sup>[2]</sup>; Kube等, 2013<sup>[3]</sup>)。有研究发现,提高工资能大幅提高员工生产率,获得更多产出,而减薪会对员工生产率产生持久的负面影响(Gneezy和List, 2006<sup>[4]</sup>)。当然,劳动力市场中由礼物交换动机引发的互惠行为不可能在任何情况下都能顺利实现,而是取决于社会环境中各种因素的相互作用(Falk, 2007<sup>[5]</sup>)。在双边礼物交换环境下,风险偏好、雇员与雇主间的交流、税收以及奖罚机制等因素都会对互惠行为产生影响;在多边礼物交换环境中,雇员则倾向于向提供更高工资的雇主付出更多的劳动作为回报;外生冲击导致的礼物收益也会影响互惠行为(徐瑶, 2020<sup>[6]</sup>)。

近些年来我国劳动力市场整体的雇佣环境虽然有所改善,但仍然存在不少问题,诸如新型就业关系的不稳定(纪雯雯, 2019<sup>[7]</sup>)以及劳动争议案件数量的持续增长等(王智崑等, 2020<sup>[8]</sup>)。其中,雇主和雇员之间出现的摩擦很大程度上来源于双方在追求自身利益最大化过程中陷入的博弈僵局,如何通过有效的方法和机制化解劳动雇佣关系中的各种争议,促进雇佣双方形成和谐高效的互惠关系对于改善我国劳动雇佣环境、提高社会劳动效率具有重要的理论和现实意义。本文在前人研究的基础上,通过礼物交换

实验探究了劳动力市场雇佣关系中互惠的触发条件及影响因素,希望对上述问题的解决有所裨益。

## 二、文献综述及研究假设

### (一) 劳动雇佣关系中的互惠偏好

与新古典经济学派的理性人假设不同,新制度经济学认为雇佣关系中雇主和雇员的行为是有限理性的(香伶和张炳申, 2006<sup>[9]</sup>)。John和Devasheesh(2011)<sup>[10]</sup>指出,劳动雇佣关系的行为主体是情感人和经济人的结合。行为经济学视角中雇主和雇员双方是具有社会属性的决策者,劳动者群体之间存在着互惠行为(严维石, 2012<sup>[11]</sup>)。事实上,劳动力的确不是简单的商品,提供劳动和雇佣劳动双方的社会情感在一定程度上左右了雇佣关系的和谐程度(江永众和程宏伟, 2012<sup>[12]</sup>)。

社会偏好的概念由Camerer(1997)<sup>[13]</sup>首次提出,该理论试图在维持理性人假设的前提下,将社会情感因素加入效用函数来对经济人假设进行修正(陈叶烽等, 2012<sup>[14]</sup>)。根据社会偏好理论,劳动雇佣关系的恶化往往是因为行为主体的自利偏好,而互惠偏好在一定程度上对缓和雇佣关系的恶化具有积极影响。Tekleab和Taylor(2003)<sup>[15]</sup>研究发现,雇佣关系中存在着心理契约上的互惠义务,互惠的存在对于提升雇佣关系的整体满意度十分重要;贾良定等(2007)<sup>[16]</sup>的实证研究也表明,注重互惠的公司,员工的工作态度也更认真;公司的生产力水平与管理层和员工的互惠行为之间存在着很强的正相关关系(Barr和Sermeels, 2009<sup>[17]</sup>);企业在制度设计和合同执行等方面考虑到互惠机制的作用时会促使雇员减少机会主义行为并增强其对企业的信任程度(王立宏, 2011<sup>[18]</sup>)。

根据上述分析我们可以发现,劳动雇佣双方并不是完全理性的利益关系,而是带有社会属性的情感关系。社会偏好(尤其是互惠偏好)会对劳动雇佣关系产生影响。据此,本文提出假设1:

**H1:** 劳动雇佣关系受到行为主体的互惠偏好影响。

### (二) 互惠的触发条件

研究发现,互惠行为的发生与人的社会性特征有关,具有一定的自我执行力(严维石, 2016<sup>[19]</sup>)。即使在有利于自私行为的实验环境中,雇主和雇员间的互惠行为仍然存在(Pereira等, 2006<sup>[20]</sup>)。但不可



否认的是,劳动雇佣关系仍然是一种基于“为己利他”逻辑的行为过程(朱富强,2011<sup>[21]</sup>),雇主和雇员并不是单方面的施惠者和受惠者关系,而是双向的互惠合作关系。从根本上来说,互惠者的利他行为仍然符合理性人假设(Andreoni和Miller,2002<sup>[22]</sup>),互惠仍是亲社会行为的利己主义(肖凤秋等,2014<sup>[23]</sup>)。

“投我以桃,报之以李”,互惠行为遵循的法则本质还是对价理论(凌文铨等,2019<sup>[24]</sup>),雇员在获得高于某个参考水平的工资回报和付出更高成本努力的心理利益之间面临着权衡(Danthine和Kurmman,2007<sup>[25]</sup>),双方的“礼物”价值在心理价位上应趋于相等。因此,参照陈叶烽等(2011)<sup>[26]</sup>对分配动机和分配结果这两种公平观的考察,以及结合互惠发挥影响的内在作用机制(邹文箴等,2012<sup>[27]</sup>),笔者推测:雇主和雇员达成互惠需要一些触发条件,既要满足个体在情感上对互惠的追求,又要满足对自身物质利益的保证。

据此,本文提出假设2:

**H2:** 互惠行为的触发需要情感和利益的双重条件。

### (三) 劳动力市场中影响互惠的因素

劳动力市场中互惠行为的存在取决于环境中各种因素的相互作用(Falk,2007<sup>[5]</sup>),这种互惠行为不仅仅依赖于工资,还受到非金钱激励因素的影响(Mahmood和Zaman,2010<sup>[28]</sup>)。Riedl和Tyran(2005)<sup>[29]</sup>在双边礼物交换实验中引入税收因素,发现税收是影响互惠行为的外部因素之一;Dennis和Anthony(2006)<sup>[30]</sup>在管理者的收益函数中引入了损失厌恶这一心理因素,发现风险中性的管理者面临损失风险时会更谨慎地选择与雇员进行互惠甚至降低互惠程度,而具有冒险精神的管理者面临同样情景则会选择更高的工资来吸引员工的互惠,以争取高回报;Cooper和Lightle(2012)<sup>[31]</sup>将交流机制引入到标准的双边礼物交换实验中,发现信息沟通能够提高雇主和雇员之间的互惠程度,当员工能与管理者直接交流时互惠程度的增加会更明显;骆欣庆等(2019)<sup>[32]</sup>将奖惩机制引入到礼物交换博弈实验,发现惩罚和奖励机制对提高互惠有着显著的促进作用。不仅如此,还有研究发现雇员间的相互竞争会对互惠行为产生影响(Lin和Meng,2019<sup>[33]</sup>),并有可能导致互惠现象不再显著(Gneezy等,2019<sup>[34]</sup>);同时,雇员之间的社会联系也会对互惠行为产生影响(Kocher和Sutter,2007<sup>[35]</sup>)。

据此,本文提出假设3:

**H3:** 雇佣双方的互惠受到劳动力市场中各种非货币因素的影响。

## 三、实验设计

在礼物交换的实验室研究中,学者们通过设计不同的实验将各种因素进行分离来研究这些因素对互惠的影响。研究的侧重点不同,引入的因素不同,结论也有所差异。笔者发现,许多礼物交换实验没有结合劳动力市场的具体情况来研究这种竞争性环境中发生的互惠行为,对现实的解释力较弱。因此,本文构建一个较为真实的虚拟劳动力市场的实验室环境,创新性地提出了互惠行为触发的二重条件,并探究在竞争性环境条件下各类因素对互惠行为的影响,进而对改进真实世界的劳动力市场雇佣关系中存在的摩擦提供建议。

### (一) 实验原理与流程

#### 1. 实验原理。

实验被试为南京大学经济系本科生,共计36人。首先,对实验被试进行随机编码,并要求所有被试在整个实验期间不得透露本人信息,保证匿名。之后根据编码将被试分为两组(两组均保证人数为偶数以达到两两匹配的目的),分别分配到两个不同的实验室,其中一组记为BASE组,另一组记为LOSS组;每一大组内再分为三个小组,分别记为L组、M组和H组,每个小组两两配对(一名被试模拟雇主,另一名被试模拟雇员),并根据提前准备好的座位表上的编码落座。

两个教室的分组(BASE组和LOSS组)的区别在于雇主收益函数设定的不同,设定主要参照了Dennis和Anthony(2006)<sup>[30]</sup>的实验。BASE组雇主的收益函数形式为 $\alpha \times (v-w) \times e^\beta$ ,LOSS组雇主的收益函数形式为 $\alpha \times v \times e^\beta - w$ ,其中 $\alpha$ 和 $\beta$ 为系数, $w$ 为雇主提供的工资水平, $e$ 为雇员提供的努力程度。为了研究损失厌恶在礼物交换实验中的影响,论文引入一个典型收益函数的替代范式(即LOSS组的收益函数)。在该条件下,扮演雇主的被试将面临潜在损失,但在开始前,本实验先为该类被试提供了一些初始资金,以避免真实的损失。因为收益函数设定的不同,所以BASE组和LOSS组的实验互不干扰。

同时,从利益最大化角度来看,雇主的收益是关于雇员努力程度的函数,而雇员努力工作为雇主带来的回报受到自身生产力水平的影响。对于不同生产力

水平的雇员，互惠机制可能也会发生变化。因此，在实验中，我们将 BASE 组和 LOSS 组内部根据雇员生产力的差别细分为三类：低等生产力 ( $L$ ) 类、中等生产力 ( $M$ ) 类和高等生产力 ( $H$ ) 类。

2. 实验流程。

实验过程中，每个实验室均配备有实验组织人员，由该组织人员来详细介绍实验的流程和操作，保证被试能完全清楚规则，不至于出现无效实验。介绍完实验规则以后，组织人员将表格发放给被试（表格每人一张，注明被试身份是雇主还是雇员，并注明实验轮次及自身的收益函数）。在正式实验开始前，先进行一轮模拟实验和答疑，确保被试理解实验规则。

每一轮实验的博弈机制如下：实验开始，雇主选择给出工资  $w$ （工资  $w$  为整数，且  $w \in \{20, 21, \dots, 120\}$ ，即  $w_{\min} = 20, w_{\max} = 120$ ）；考虑 30 秒后，雇员需要根据雇主提供的工资  $w$  决定是否签订合同。如果签订合同，雇员会给出为该工资付出的努力水平  $e$ （努力水平  $e$  为整数，且  $e \in \{1, 2, \dots, 10\}$ ，即  $e_{\min} = 1, e_{\max} = 10$ ）；如果选择不签订合同，则双方收益均为 0。本轮实验结束，双方根据表格上的收益函数计算得出自己本轮实验的收益。

BASE 组实验共有三个阶段：第一阶段有 8 轮实验，第二阶段有 6 轮实验，第三阶段有 10 轮实验，每一阶段的具体实验规则始终一致。第一阶段实验结束后，被试有 15 分钟休息时间，在此期间，被试间不得进行任何交流。开始第二阶段实验前，前一阶段实验每一类下的雇主与雇员重新匹配，但双方身份保持不变。第二阶段实验结束后，被试在休息时被允许在本小组内围绕努力值、工资水平等实验相关内容进行 10 分钟讨论，其间不得跨组讨论，不得涉及双方自身收益值。然后维持第二阶段的匹配进行第三阶段的实验。LOSS 组实验共有两个阶段，第一阶段有 10 轮实验，第二阶段也有 10 轮实验。第二阶段时，每一类下的雇主与雇员重新匹配，但双方身份保持不变。

表 1 实验设计

实验条件	实验阶段	实验轮数	实验简述
BASE 组	第一阶段	8	基准
	第二阶段	6	讨论
	第三阶段	10	重新匹配
LOSS 组	第一阶段	10	基准
	第二阶段	10	重新匹配

(二) 参数选择与激励机制

1. 实验参数。

雇员收益。本实验中雇员的收益受到两个因素的影响：一是雇主提供的工资  $w$ ，二是该雇员为了付出提供的努力水平所需要的货币成本  $c(e)$ 。实验中，如果雇员选择签订合同，则收益函数  $\pi = w - c(e)$ ；如果选择拒绝签订合同，则雇员的净收益为初始资金 0。每一轮实验时，员工付出的努力水平  $e$  与付出此水平的努力所需要的货币成本  $c(e)$  对应如下。

表 2 雇员努力成本

努力水平： $e$	1	2	3	4	5
货币成本： $c(e)$	20	21	22	24	26
努力水平： $e$	6	7	8	9	10
货币成本： $c(e)$	28	30	32	35	38

雇主收益。本实验中雇主的收益受到三个因素的影响：一是实验设定中两个不同的收益函数形式，BASE 组的雇主收益保证为正，LOSS 组的雇主收益则有在付高薪时面临亏损的风险；二是为了雇佣员工所付出的工资水平；三是雇员反馈的努力程度所带来的收益，但该部分收益不仅受努力程度影响，还受到雇员自身生产力水平的影响：同等努力条件下，雇员生产力水平越高，能提供给雇主的收益越高。

表 3 雇主收益函数总表

实验条件		雇主收益函数	
基准组 (BASE)	低生产力	$y = \begin{cases} \frac{1}{10} (v_B - w) \times e^{0.8} \\ 0 \end{cases}$	签订合同 拒绝合同
	中生产力	$y = \begin{cases} \frac{1}{10} (v_B - w) \times e^{0.9} \\ 0 \end{cases}$	签订合同 拒绝合同
	高生产力	$y = \begin{cases} \frac{1}{10} (v_B - w) \times e \\ 0 \end{cases}$	签订合同 拒绝合同
损失组 (LOSS)	低生产力	$y = \begin{cases} \frac{1}{10} v_L \times e^{0.8} - w + 30 \\ 0 \end{cases}$	签订合同 拒绝合同
	中生产力	$y = \begin{cases} \frac{1}{10} v_L \times e^{0.9} - w + 30 \\ 0 \end{cases}$	签订合同 拒绝合同
	高生产力	$y = \begin{cases} \frac{1}{10} v_L \times e - w + 30 \\ 0 \end{cases}$	签订合同 拒绝合同

注： $v_B = 120, v_L = 100, w \in \{20, 21, \dots, 120\}, e \in \{1, 2, \dots, 10\}$ 。

2. 激励机制。

为了激励被试积极参与实验，并按照实验要

求填写表格，实验以课程加分作为实验参与奖励。每轮实验结束后，雇主和雇员分别按照各自收益表上的公式计算收益。所有实验结束后，按相同身份、相同组别的参与者收益排名进行课程附加分激励（排名最低的被试仍可以获得参与奖励，

以此弱化竞争外部性的影响)。除了作为激励的课程分数外，实验中所有其他数值的单位都为虚拟货币单位。

3. 变量选取。

本文各变量定义及其赋值规则如表4所示。

表4 变量定义及赋值

变量名称	变量	变量定义及赋值
工资水平	<i>wage</i>	雇主提供的工资 $w, w \in \{20, 21, \dots, 120\}$
努力水平	<i>effort</i>	雇员提供的努力水平 $e, e \in \{1, 2, \dots, 10\}$
雇主收益	<i>manager_profit</i>	根据雇主收益函数计算得出的收益
雇员收益	<i>employee_profit</i>	根据雇员收益函数计算得出的收益
交流情况	<i>contract</i>	虚拟变量，雇主与雇员有交流则 $contract=1$ ，否则 $contract=0$
匹配情况	<i>another</i>	虚拟变量，雇主与雇员之间重新匹配则 $another=1$ ，否则 $another=0$

四、实验结果与分析

(一) 描述性分析

1. 互惠行为存在性检验。

为了验证互惠行为的存在性，本文首先分析了雇主提供的工资与雇员提供的努力水平之间的关系(图1)。结果显示，尽管因为实验设置使得不同生产力水平下 *wage* 和 *effort* 的关系有些许差异，但在BASE组和LOSS组的实验中，工资水平和努力水平两者都呈同方向增加，证明了礼物交换实验中互惠行为的真实存在。

2. 互惠的触发机制。

根据利益最大化原则，雇主和雇员回礼的原因可能并不是源于超额工资或超额劳动水平，而是源于自身收益的超额增加。因此，本文在此基础上，以雇主和雇员自身的收益来探讨礼物交换实验中互惠行为的

触发机制。如图2所示，BASE组雇主的收益和提供的工资水平的趋势线比较杂乱，无明显趋势。这说明对于雇主来说，当自身收益受到保障，不会有亏损情况时，提供的工资变化与自身收益的变化没有显著影响；而当自身收益有亏损的可能性时（即在LOSS组中），雇主会更加关心自身收益，只有当自身收益因雇员提供的努力程度增加而提高时，才会提供更高的工资作为回礼。

同时，BASE组和LOSS组雇员收益与雇员提供的努力水平正相关，但LOSS组雇员收益的增加会引发导致其努力水平更大幅度的提高。实际上LOSS组雇主因为有亏损可能，最开始给出的工资水平偏低（图2中LOSS组的散点更集中于左下角），随着后期工资水平的提高，雇员心理上感到更大程度的善意，因而雇员会更大幅度地提升努力水平（图3）。

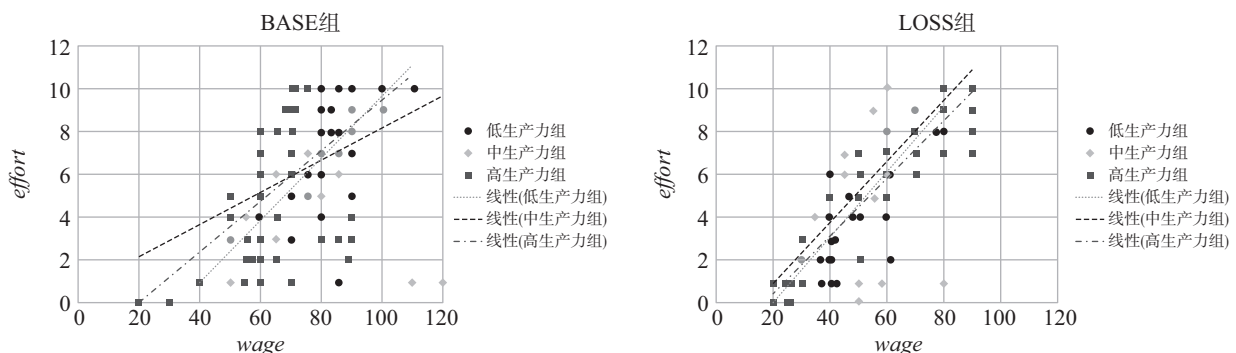


图1 wage-effort 散点图及趋势线

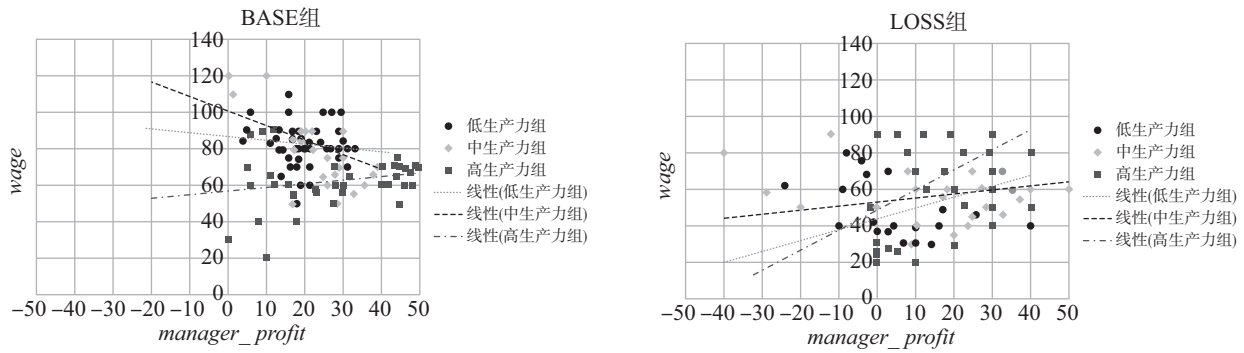


图2 manager\_profit-wage 散点图及趋势线

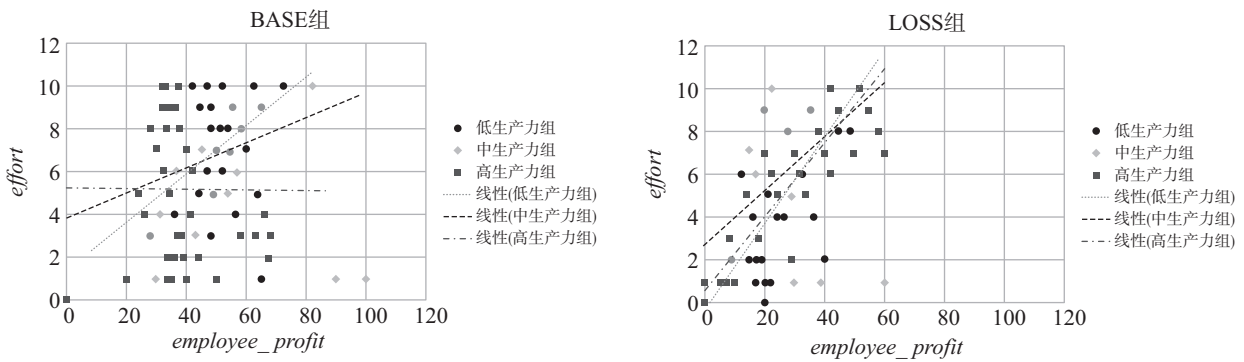


图3 employee\_profit-effort 散点图及趋势线

因此，在雇主利益不会出现损失的情况下，工资和努力程度的超额提供能直接触发雇主和雇员两者的互惠行为。但当雇主利益有损失风险时，雇员提供给雇主以超额的努力程度作为礼物后，雇主并不会直接根据该条件予以回礼，而是会充分考虑自身收益，在确保自身收益有增幅的情况下再选择回礼。于是，在更贴近现实的风险情境下，互惠行为的触发需要二重条件：既需要提供礼物的情感条件，又需要自身收益提高的物质条件。

3. 互惠的影响因素。

(1) 信息交流的影响。

如图4所示，第8轮实验后BASE组两两匹配

的雇主和雇员之间进行了一次10分钟左右的信息交流。以第8轮为节点，经过信息交流，并没有出现互惠程度的提高。相反，大部分雇主保持了原有的工资水平，而部分雇员提供的努力水平出现了较大的波动。

本文认为，该波动的出现与雇主和雇员在劳动力市场上不平等的地位有关。进行信息交流时，雇主处于主导地位，雇员在交流后一旦发现雇主没有按照事先约定提供更高水平的工资，就会选择以极低的努力水平作为报复。因此，信息交流不一定能提高雇主和雇员之间的互惠程度，相反地，因为交流不当，原有的默契和策略被打乱，互惠程度反而会降低。

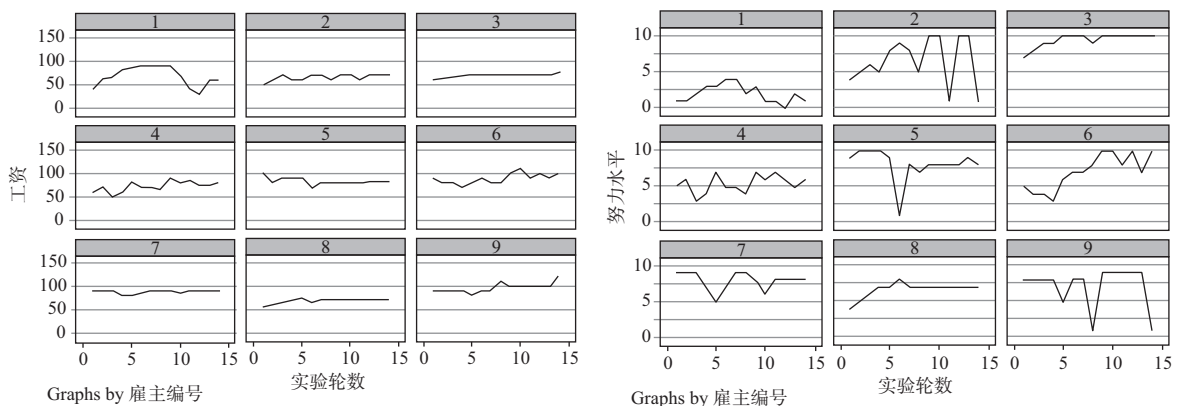


图4 BASE组第一阶段和第二阶段工资水平和努力水平

(2) 重新匹配的影响。

实验结果显示，随着雇主和雇员重新匹配，雇主提供的工资和雇员提供的努力程度有了明显的提高。这说明在有了前一阶段互惠的经验后，雇主和雇员之间都更愿意用更高的工资和努力水平来促进互惠程度的提高，这也较符合现实生活中员工跳槽升薪的真实情况。需要注意的是，观测数据时需要剔除最后一轮实验，因为在有限次博弈中，最后一局理性的博弈方必然会选择背叛对方以寻求个人利

益最大化，仅以最小的努力水平来获取最后收益的行为。

(3) 生产力水平的影响。

图6显示，从整体趋势来看，随着生产力水平的提高，雇主用作回礼的工资水平也在提高。并且对比BASE组和LOSS组来看，在没有损失风险时，雇主与雇员的互惠程度受生产力水平影响更高。在有亏损风险时，不论处于何种情况，雇主作出的决策（回礼）都会更加谨慎。

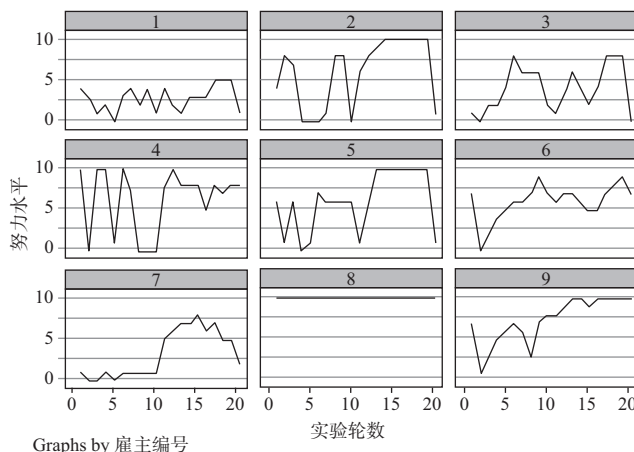
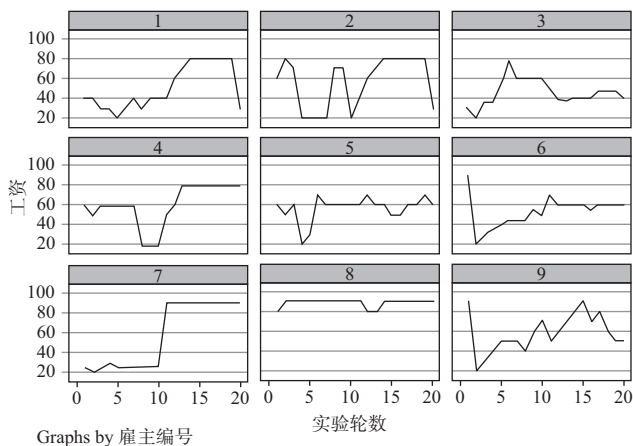


图5 LOSS组第一阶段和第二阶段工资水平和努力水平

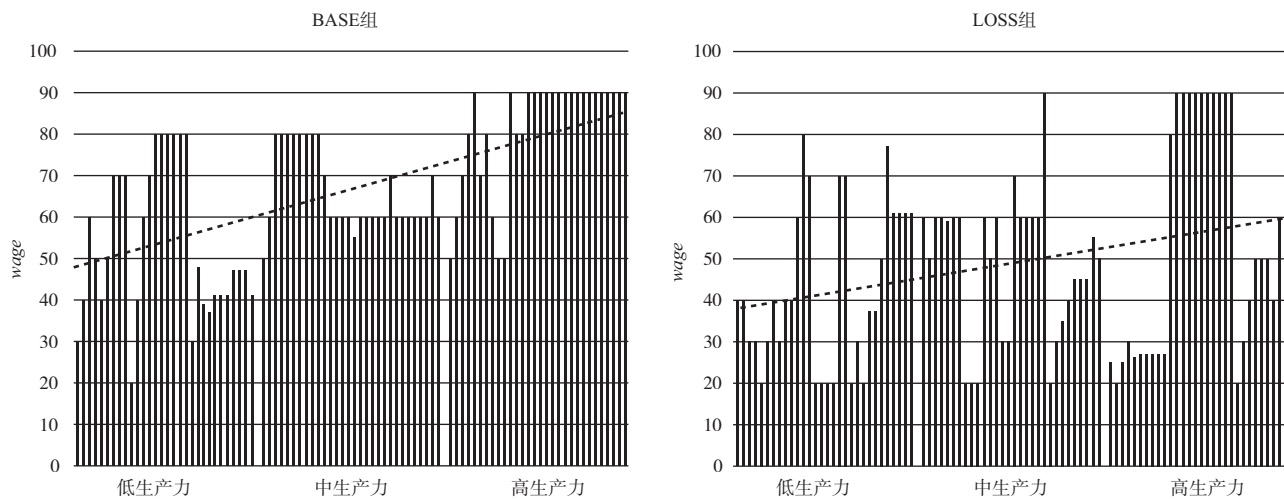


图6 生产力分组雇员工资水平柱状图及趋势线

(二) 计量检验

1. 计量模型。

本文的实验数据分为雇主和雇员的短面板数据集两部分。因此，我们基于个体固定效应模型设计了如下回归模型：

对于雇主：

$$wage\_1_{it} = \alpha_0 + \alpha_1 \times effort_{it} + \alpha_2 \times manager\_profit_{it} + \alpha_3 \times contract + \alpha_4 \times another \quad (1)$$

对于雇员：

$$effort\_1_{it} = \beta_0 + \beta_1 \times wage_{it} + \beta_2 \times employee\_profit_{it} + \beta_3 \times contract + \beta_4 \times another \quad (2)$$

其中， $wage\_1$  和  $effort\_1$  分别表示工资  $wage$  和努力程度  $effort$  的一阶滞后项； $i$  表示被试个体， $t$  表示实验轮数。工资水平与努力水平采用一阶滞后项的原因在于，互惠行为的产生具有时滞效应，即雇主根据上

一轮雇员的努力水平作出本轮的回馈反应，雇员也是根据上一轮雇主提供的工资水平作出本轮的回馈反应。

## 2. BASE 组互惠行为的影响因素。

对于雇主而言，雇员每一轮提供的努力程度和双方的信息交流均对其下一轮提供给雇员的工资水平有一定的影响，雇员提供的努力程度的增加会提高下一轮雇主提供给雇员的工资水平。

回归结果显示，雇员提供的努力水平对下一轮雇主提供的工资有显著的提升作用，并通过了 1% 水平

的显著性检验；雇主的收益对于下一轮提供的工资的回归系数为 -0.025，结果为负且不显著，可能是由于在 BASE 组中雇主的收益没有亏损的风险，因此雇主在回礼给雇员时并不会仔细考虑自身收益，而是受到更直观的努力水平的影响。是否有信息交流对提供工资的影响也在 10% 的水平上显著为正，而是否重新匹配则对互惠程度的影响不显著，这可能是由于 BASE 组第三阶段的重新匹配实验受到了前一阶段信息交流实验的影响。因此，不能通过 BASE 组的实验数据来分析重新匹配对互惠行为的影响。

表 5 BASE 组雇主数据回归结果

变量	(1)	(2)	(3)	(4)
<i>effort</i>	1.119 *** (0.283)		1.021 *** (0.283)	1.037 *** (0.285)
<i>manager_profit</i>		-0.025 (0.071)		
<i>contract</i>			3.896 * (1.615)	
<i>another</i>				-2.559 (1.358)
<i>cons</i>	68.561 *** (1.944)	76.403 *** (1.940)	68.292 *** (1.924)	70.202 *** (2.119)
被试个体 FE	控制	控制	控制	控制
实验轮数 FE	控制	控制	控制	控制
观测值	207	207	207	207
$R^2$	0.073	0.001	0.100	0.090

注：\*\*\*、\*\*、\* 分别表示在 1%、5% 和 10% 的显著性水平；括号内为标准误。下同。

由表 6 可知，对于雇员而言，雇主每一轮提供的工资水平和雇员的自身收益均对其下一轮的努力程度有一定的影响。雇主提供的工资水平的提高对下一轮雇员的努力水平有显著的提升作用，并通过了 1% 的显著性检验。雇员的收益也对下一轮的努力水平有显

著的正向影响，这一结果说明相比雇主，雇员更在乎礼物给自己带来的实际收益，而是否交流和是否重新匹配在 BASE 组的实验中对雇员“回礼”的程度并没有显著的影响。

表 6 BASE 组雇员数据回归结果

变量	(1)	(2)	(3)	(4)
<i>wage</i>	0.090 *** (0.014)		0.088 *** (0.014)	0.089 *** (0.014)
<i>employee_profit</i>		0.056 *** (0.016)		
<i>contract</i>			0.266 (0.356)	

续前表

变量	(1)	(2)	(3)	(4)
<i>another</i>				-0.061 (0.302)
<i>cons</i>	-0.299 (1.033)	5.081*** (0.431)	-0.196 (1.043)	-0.215 (1.115)
被试个体 FE	控制	控制	控制	控制
实验轮数 FE	控制	控制	控制	控制
观测值	207	207	207	207
$R^2$	0.184	0.060	0.186	0.184

3. LOSS 组互惠行为的影响因素。

加入了损失风险后，雇员提供的工作努力程度、雇主当轮的收益和是否重新匹配均对雇主下一轮提供给雇员的工资水平有显著的正向影响。

其中，雇主面临损失风险时，将不再只观察雇员的努力水平，而是观察自身的收益函数，考虑雇员给自己的“礼物”是否能够真正提高自身的收益。与

BASE 组情况相悖，在 LOSS 组中是否重新匹配对互惠行为产生了显著影响，这可能是因为 LOSS 组的实验设置排除了信息交流的干扰影响（表 7）。

类似地，雇主当轮实验提供的工资、雇员自身的收益以及是否重新匹配均对下一轮雇员的“返礼”有显著的正向影响（表 8）。

表 7 LOSS 组雇主数据回归结果

变量	(1)	(2)	(3)
<i>effort</i>	3.462*** (0.395)		2.655*** (0.436)
<i>manager_profit</i>		0.331*** (0.111)	
<i>another</i>			10.475*** (2.797)
<i>cons</i>	37.871*** (2.691)	53.159*** (2.394)	37.219*** (2.594)
被试个体 FE	控制	控制	控制
实验轮数 FE	控制	控制	控制
观测值	171	171	171
$R^2$	0.324	0.052	0.378

表 8 LOSS 组雇员数据回归结果

变量	(1)	(2)	(3)
<i>wage</i>	0.094*** (0.010)		0.079*** (0.012)
<i>employee_profit</i>		0.122*** (0.014)	

续前表

变量	(1)	(2)	(3)
<i>another</i>			0.973 *** (0.424)
<i>cons</i>	0.415 (0.622)	2.340 *** (0.431)	0.793 (0.636)
被试个体 FE	控制	控制	控制
实验轮数 FE	控制	控制	控制
观测值	171	171	171
$R^2$	0.341	0.332	0.362

## 4. 生产力水平的异质性分析。

本文根据雇员生产力水平的差异,将其分为低、中、高生产力组,并对分组数据进行了异质性分析。为控制其他因素的影响,异质性分析只使用了 BASE 组和 LOSS 组第一阶段的实验数据。

由表 9 可知, BASE 组的三个不同生产力样本组中,雇员努力水平对于下一轮雇主提供工资的回归系数分别为-0.150、0.837 和 3.670。虽然由于

样本量问题导致显著性水平不高,但也在一定程度上说明随着生产力水平的提升,雇主更愿意对雇员实施互惠行为,且互惠的程度也会随之提高。但在 LOSS 组中,结论恰恰相反,生产力水平越高,雇主的互惠程度反而越低。这可能是因为在雇主的利益有了损失风险后,高生产力水平的雇员对雇主收益变动的的影响程度加深,雇主的不信任导致无法互惠。

表 9 分样本雇主数据回归结果

变量	低生产力		中生产力		高生产力	
	BASE 组	LOSS 组	BASE 组	LOSS 组	BASE 组	LOSS 组
<i>effort</i>	-0.150 (0.938)	2.335 * (1.300)	0.837 (0.642)	0.664 (1.022)	3.670 * (2.105)	-2.111 (1.733)
<i>cons</i>	78.532 *** (6.055)	35.501 *** (5.520)	74.339 *** (4.659)	46.011 (5.808)	45.090 *** (13.261)	65.327 *** (9.387)
被试个体 FE	控制	控制	控制	控制	控制	控制
实验轮数 FE	控制	控制	控制	控制	控制	控制
观测值	21	27	21	27	21	27
$R^2$	0.079	0.167	0.047	0.013	0.002	0.692

BASE 组中,生产力水平越高的雇员给雇主提供“礼物”的努力程度也越高,高生产力水平的雇员更愿意用高努力水平来换取雇主的高工资水平,更愿意与雇主之间达成互惠。而在 LOSS 组中,高生产力水平的雇员对雇主工资的提高反而没有互惠的互动,只有低生产力组雇主工资的提高对下一轮雇员努力水平

存在显著的正向影响。这也符合前文对 LOSS 组实验的论断:当雇主利益存在损失风险时,雇主和雇员对于双方互惠会更加的谨慎,也就表现为表格中 LOSS 组随着生产力水平提高,工资对于下一轮努力水平的回归系数减少并有可能为负(表 10)。



表 10 分样本雇员数据回归结果

变量	低生产力		中生产力		高生产力	
	BASE 组	LOSS 组	BASE 组	LOSS 组	BASE 组	LOSS 组
<i>wage</i>	0.053 (0.068)	0.077*** (0.027)	0.059 (0.049)	0.042 (0.050)	0.116*** (0.036)	-0.007 (0.038)
<i>cons</i>	2.046*** (5.217)	-0.143 (1.248)	2.416*** (4.034)	2.963 (2.419)	-2.268 (2.552)	5.503** (2.041)
被试个体 FE	控制	控制	控制	控制	控制	控制
实验轮数 FE	控制	控制	控制	控制	控制	控制
观测值	21	27	21	27	21	27
$R^2$	0.144	0.324	0.254	0.031	0.384	0.810

综上所述，根据实验结果，雇主和雇员之间存在明确的互惠倾向，其触发来源于双方给予的直观的“礼物”，而出于双方对损失风险的厌恶，在一定程度上抑制雇主和雇员之间的互惠倾向，导致同等条件下互惠程度的降低，此时双方互惠的触发需要考量自身的收益水平。工资水平和努力水平、信息交流和雇主雇员间的重新匹配均会在一定程度上增强两者间的互惠程度，雇员自身生产力水平的提升也能够促进雇主与其达成更高层次的互惠。

### 五、研究结论与展望

#### (一) 研究结论

本文使用实验经济学方法探究了劳动力市场雇佣关系中互惠的触发条件和影响因素，试图对我国日益加剧的劳动关系摩擦的解决有所裨益。通过对实验结果的整理、统计和计量分析，得出以下主要研究结论。

#### 1. 雇主和雇员双方存在着互惠行为倾向。

本文的实验验证了在双边礼物交换环境中劳动雇佣双方存在着显著的互惠关系。一般来说，雇主提供更高的工资会带来员工更大的努力水平。同样，员工努力水平的提升也会进一步激励雇主提供更高的工资。在实验过程中我们明显地观察到雇主提供工资水平的变化会给员工的努力水平带来显著的正向影响，反之亦然。这一结果也符合我们在日常生活中对雇佣关系的认知：雇主愿意给工作更努力的员工支付更高的工资，员工同样会为支付更高工资的雇主付出更多的工作努力。

#### 2. 互惠行为的触发需要情感和利益的双重条件。

劳动关系中的雇佣双方通常会面临信息不对称等

问题带来的损失风险，实验发现损失风险的情况下，雇主往往会表现出强烈的风险厌恶，对雇员努力水平的变化也会更加敏感。此时互惠的触发条件不仅取决于双方直观的“礼物赠予”，也更会受到自身收益的影响。只有当对方给予的“礼物”能真实地提高自身收益，才会进一步促进互惠。这一结果表明互惠的产生不仅仅是基于心理上的善意，也是由相关的利益因素触发的。本文将此总结为损失厌恶条件下雇佣双方互惠行为的触发二重条件，这一结论增加了我们对互惠触发条件的认知。

#### 3. 互惠行为受到信息交流和人事变动等因素的影响。

雇主与雇员间的信息交流一方面为双方达成互惠提供了面对面的平台，另一方面也可能干扰原有的互惠行为，交流不当或未能达成协议反而会恶化原有雇佣关系。本文在实验中引入交流机制后，部分雇主和雇员在交流后几轮选择背叛，进而导致后期互惠的不稳定，这一结果表明合理有效的交流才能促进雇佣关系的优化。同样，我们在实验中通过对被试进行重新匹配模拟劳动市场中的人事调动，结果表明重新匹配后，新搭档的雇主与雇员之间往往能更快进入互惠阶段，并且互惠程度较重新匹配前均有所提高，这一结果说明有过人事调动的雇主和雇员更会通过行动促成双方互惠的雇佣关系。

#### 4. 雇员生产力水平差异会影响互惠程度。

雇员自身的生产力水平差异并不会影响互惠行为的发生，但会影响互惠的程度。在雇主能保证自身收益（即无损失风险）时，更愿意对高生产力水平的雇员提供更高的工资，高生产力水平的雇员也愿意用更高的努力水平回报雇主的善意；当雇主不

能保证自身收益（即有损失风险）时，雇主反而更容易与低生产力水平的雇员达成互惠，这可以看作是雇主规避高风险的一种行为。因此，在风险情况下如何促进雇主与高生产力水平雇员之间的信任，降低雇主的风险厌恶水平对于提升双方的互惠程度至关重要。

## （二）管理启示

上述实验及研究结论赋予我们的管理启示，对于通过促进雇佣关系中的互惠来解决我国劳动力市场的摩擦问题具有一定的指导意义。

1. 雇主要建立有助于与员工形成互惠的正向激励制度。

雇主在制定员工工资时，其重心应该放在工资变化的幅度上，比如明确的升职路径、清楚的奖金分成等，这些都对雇员的未来预期产生正面影响，从而对其在实际工作中的努力带来很好的正向激励效果；在与雇员交流协商时，要制定合理的用工合同以及提前做好惩罚措施，这样能有效避免员工的违约行为；同时，要明确向雇员表达自己的风险偏好和预期。公司还要注意搭建平等沟通的企业文化和信息通畅的工作环境，促进管理层和基层员工的交流，建立彼此间的信任，形成良好的雇佣关系氛围。

2. 雇员要采取各种有效行动促进稳定互惠关系的形成。

雇员要积极响应互惠动机雇主发出的信号，努力提高自身的生产力水平，认识到雇主对努力水平变化的感受是非常敏感的，一旦懈怠就可能面临薪资水平的降低。在与雇主进行交流时也要注意具体协议的签订，以保障自身的努力得到回报。对于待就业的雇员来说，选择成立时间较长的企业是比较正确的选择，

因为有雇佣经验的雇主往往更愿意和雇员之间达成互惠，并不断提高互惠程度以达成双赢。同时，雇员也需要关注雇主的风险和收益预期，对于风险厌恶程度高的雇主应采取更大的努力程度来达成双方互惠。

3. 劳动保障部门应建立长效制度保障雇主与雇员之间互惠的实现。

对于劳动保障部门（特别是一些负责制定劳动保护法的部门以及工会），既要通过宣传手段加强对和谐雇佣环境的建设，也要通过法律法规等手段加强对雇佣双方利益的共同保障。从互惠角度，劳动保障部门应建立企业与劳动者（尤其是高生产力劳动者）的诚信体系，使得雇主和员工在签订劳动合同前能清楚了解对方的用工和工作记录，降低损失风险，促进互惠的实现。尤其要注意的是在保障雇员利益的同时，要充分考虑到雇主的利益。如果一味保障雇员利益，而对雇主的利益造成损害（或是增加了雇主面临亏损的风险），反而会导致雇主不愿意与雇员互惠共赢，恶化了雇佣关系，不利于劳动力市场的和谐。

## （三）局限与展望

本文研究了劳动雇佣关系中互惠的触发条件和影响因素等，对于本问题的研究，未来可以从以下两个方面进行更加深入的探究与解答。一是从互惠行为的内在机理角度进行研究，例如对于雇佣双方而言，其互惠行为的内在驱动力是基于公平性原则还是基于差异厌恶等心理因素。二是从双边雇佣关系延伸到多边雇佣关系的研究，本文的实验设定在一个双边的雇佣关系环境中，未来研究可以继续设定多边的雇佣关系环境，检验在多变雇佣关系中互惠机制的触发条件及其影响因素是否依旧稳健。

## 参考文献

- [1] Akerlof G A. Labor Contracts as Partial Gift Exchange [J]. The Quarterly Journal of Economics, 1982, 97 (4): 543-569.
- [2] Fehr E, Goette L, Zehnder C. A Behavioral Account of the Labor Market: The Role of Fairness Concerns [J]. Annual Review of Economics, 2009, 1 (1): 355-384.
- [3] Kube S, Maréchal M A, Puppe C. Do Wage Cuts Damage Work Morale? Evidence from a Natural Field Experiment [J]. Journal of the European Economic Association, 2013, 11 (4): 853-870.
- [4] Gneezy U, List J A. Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments [J]. Econometrica, 2006, 74 (5): 1365-1384.
- [5] Falk A. Gift Exchange in the Field [J]. Econometrica, 2007, 75 (5): 1501-1511.
- [6] 徐瑶. 多边礼物交换中的互惠行为及其负外部性: 理论与实验证据 [D]. 杭州: 浙江大学, 2020.
- [7] 纪雯雯. 可持续社会的就业与未来雇佣关系的研究动态 [J]. 中国人力资源开发, 2019 (6): 98-112.
- [8] 王智崑, 赵继伦, 于桂兰. 我国劳动争议案件数量增长状况与原因——基于1991—2016年相关数据的实证分析 [J]. 山东大学学报

- (哲学社会科学版), 2020 (5): 64-73.
- [9] 香伶, 张炳申. 新制度经济学对劳动力市场理论的影响 [J]. 财贸经济, 2006 (1): 82-85.
- [10] 约翰·巴德, 迪瓦希什·海沃, 孟泉. 雇佣关系: 人力资源管理的基础 [J]. 中国人力资源开发, 2011 (9): 85-95.
- [11] 严维石. 内部劳动力市场中的互惠行为与技能外溢效应——基于经济社会学视角 [J]. 浙江大学学报 (人文社会科学版), 2012 (3): 191-197.
- [12] 江永众, 程宏伟. 劳动关系研究的多学科比较——基于劳动经济学和人力资源管理学的视角 [J]. 学术研究, 2012 (5): 91-98, 159-160.
- [13] Camerer C F. Progress in Behavioral Game Theory [J]. Journal of Economic Perspectives, 1997, 11 (4): 167-188.
- [14] 陈叶烽, 叶航, 汪丁丁. 超越经济人的社会偏好理论: 一个基于实验经济学的综述 [J]. 南开经济研究, 2012 (1): 63-100.
- [15] Tekleab A G, Taylor M S. Aren't There Two Parties in an Employment Relationship? Antecedents and Consequences of Organization-employee Agreement on Contract Obligations and Violations [J]. Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 2003, 24 (5): 585-608.
- [16] 贾良定, 宋继文, 李超平, 崔荣军, 陈永霞. 领导风格与员工工作态度——互惠和信任的中介作用的实证研究 [J]. 中大管理研究, 2007 (1): 13-45.
- [17] Barr A, Serneels P. Reciprocity in the Workplace [J]. Experimental Economics, 2009, 12 (1): 99-112.
- [18] 王立宏. 基于不完全合约的企业内部互惠机制研究 [J]. 辽宁大学学报 (哲学社会科学版), 2011 (4): 93-99.
- [19] 严维石. 小群体中互惠行为机理研究: 基于行为经济学视角 [J]. 中央财经大学学报, 2016 (3): 82-87.
- [20] Pereira P T, Silva N, e Silva J A. Positive and Negative Reciprocity in the Labor Market [J]. Journal of Economic Behavior & Organization, 2006, 59 (3): 406-422.
- [21] 朱富强. “为己利他”行为机理的行为特性: 互惠合作 [J]. 改革与战略, 2011 (1): 11-18.
- [22] Andreoni J, Miller J. Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism [J]. Econometrica, 2002, 70 (2): 737-753.
- [23] 肖凤秋, 郑志伟, 陈英和. 亲社会行为产生机制的理论演进 [J]. 心理科学, 2014 (5): 1263-1270.
- [24] 凌文铨, 李锐, 聂婧, 李爱梅. 中国组织情境下上司—下属社会交换的互惠机制研究——基于对价理论的视角 [J]. 管理世界, 2019, 35 (5): 134-148, 199-200.
- [25] Danthine J P, Kurmann A. The Macroeconomic Consequences of Reciprocity in Labor Relations [J]. The Scandinavian Journal of Economics, 2007, 109 (4): 857-881.
- [26] 陈叶烽, 周业安, 宋紫峰. 人们关注的是分配动机还是分配结果? ——最后通牒实验视角下两种公平观的考察 [J]. 经济研究, 2011 (6): 31-44.
- [27] 邹文箴, 田青, 刘佳. “投桃报李”——互惠理论的组织行为学研究述评 [J]. 心理科学进展, 2012 (11): 1879-1888.
- [28] Mahmood S, Zaman A. Monetary and Non-monetary Gift Exchange [J]. The Pakistan Development Review, 2010: 719-740.
- [29] Riedl A, Tyran J R. Tax Liability Side Equivalence in Gift-exchange Labor Markets [J]. Journal of Public Economics, 2005, 89 (11/12): 2369-2382.
- [30] Dennis A V. Dittrich & Anthony Ziegelmeyer. Laboratory Bilateral Gift Exchange: The Impact of Loss Aversion [Z]. Working Paper, 2006.
- [31] Cooper D J, Lightle J P. The Gift of Advice: Communication in a Bilateral Gift Exchange Game [J]. Experimental Economics, 2013, 16 (4): 443-477.
- [32] 骆欣庆, 刘剑锋, 谢俊杰. 惩罚、奖励与合作行为——基于礼物交换博弈实验的对比研究 [J]. 经济与管理, 2019 (3): 26-35.
- [33] Lin W, Meng J. Does Gift Competition Hurt? An Experimental Study of Multilateral Gift Exchange [J]. An Experimental Study of Multilateral Gift Exchange (June 12, 2019), 2019.
- [34] Gneezy U, Saccardo S, Van Veldhuizen R. Bribery: Behavioral Drivers of Distorted Decisions [J]. Journal of the European Economic Association, 2019, 17 (3): 917-946.
- [35] Kocher M G, Sutter M. Individual Versus Group Behavior and the Role of the Decision Making Procedure in Gift-exchange Experiments [J]. Empirica, 2007, 34 (1): 63-88.

(责任编辑: 邵霖 张安平)

# 失信被执行人名单制度影响企业银行借款机理探讨

## ——基于第三方声誉惩罚视角的检验

Research on the Influence Mechanism of the List of Persons Subject to Enforcement for Trust-breaking System on the Loan of Enterprise Banks: Test from the Perspective of Third-party Reputation Punishment

周 洲 周丹琪 冉 戎

ZHOU Zhou ZHOU Dan-qi RAN Rong

**[摘要]** 笔者基于第三方声誉惩罚理论,利用2010—2016年中国沪深A股上市公司数据,运用双重差分法检验了2013年出台的失信被执行人名单制度对企业银行借款的影响及其机理。检验结果证实:失信被执行人名单制度是由公权力担任执行者的第三方声誉惩罚,提供了声誉惩罚的“公共品”,提高了社会诚信水平并降低了市场交易成本,有助于企业获得银行借款。这种第三方声誉惩罚措施还能够通过两条路径对企业银行借款产生正向影响:一是第三方声誉惩罚增加了企业失信的声誉成本,减少了企业失信行为,提高了企业声誉;二是第三方声誉惩罚能增强企业信用意识,使企业更加重视并提高内部控制质量,有助于企业获得更多的银行借款。失信被执行人名单制度的第三方声誉惩罚效果在不同企业中存在差异,企业风险水平越低和地区金融发展水平越低,第三方声誉惩罚对企业银行借款的影响也越显著。失信被执行人名单制度在提高企业银行借款的同时减少了企业商业信用融资,改变了企业融资结构;即使设立第三方声誉惩罚退出机制也不会弱化其对企业银行借款的作用。本研究首次从第三方声誉惩罚视角实证检验了失信被执行人名单制度对企业银行借款的影响,丰富了有关第三方声誉惩罚理论以及企业银行借款影响因素等方面的文献,为政府正确认识并充分发挥第三方声誉惩罚对企业融资的作用提供了理论依据。

**[关键词]** 失信被执行人名单制度 银行借款 影响机理 第三方声誉惩罚 双重差分

**[中图分类号]** F832.4 D64 **[文献标识码]** A **[文章编号]** 1000-1549 (2022) 11-0102-13

**Abstract:** Based on the theory of third-party reputation punishment, the author uses the data of China's Shanghai-Shenzhen A-share listed companies from 2010 to 2016, and uses the difference method to test the impact and mechanism of the list of persons subject to enforcement for trust-breaking introduced in 2013 on corporate bank loans. The test results confirm that the system of persons subject to enforcement for trust-breaking is a third party reputation punishment with public power as the executor, which provides the “public good” of reputation punishment, improves the level of social integrity and reduces market transaction costs, and helps enterprises to obtain bank loans. Such third-party reputation punishment measures can also have a positive impact on corporate bank borrowing through two paths: first, third-party reputation punishment increases the reputation cost of corporate trust-breaking, reduces corporate trust-breaking behavior, and improves corporate reputation; second, the third party reputation punishment can also enhance the credit consciousness of enterprises, make enterprises pay more attention to and improve the quality of internal control, help enterprises to obtain more bank loans. The effect of third-party reputation punishment of the system of persons subject to enforcement for trust-breaking is different in different enterprises. The lower the enterprise risk level and the lower the regional financial development level, the more significant the impact of third-party reputation punishment on corporate bank borrowing. The system of the list of persons subject to enforcement for trust-breaking not only improves the loan of the enterprise bank, but also reduces the commercial credit financing and changes the financing structure of the enterprise. Even the establishment of the third party reputation penalty exit mechanism will not weaken its role in corporate bank borrowing. This study for the first time from the perspective of third-party reputation punishment empirically the faithless person subjected to execution list system influence on enterprise bank borrowing, enriches the literature on the third-party reputation punishment theory and the influencing factors of corporate bank loan, for the government to correctly understand and give full play to the third party reputation punishment provides theory basis for the role of corporate finance.

**Key words:** List system of persons subject to enforcement for trust-breaking Bank loan Influence mechanism Third party reputation penalty Double difference

**[收稿日期]** 2022-07-12

**[作者简介]** 周洲,男,1975年5月生,重庆大学公共管理学院副教授,重庆大学公共经济与公共政策研究中心研究员,主要研究方向为制度经济学、法经济学;周丹琪,女,1999年6月生,重庆大学公共管理学院硕士研究生,研究方向为制度经济学、产业经济学;冉戎,男,1978年12月生,重庆大学公共管理学院教授,主要研究方向为公共政策与企业社会责任。本文通讯作者为周洲,联系方式为 zzy\_75@cqu.edu.cn。

**[基金项目]** 国家自然科学基金面上项目“行政问责及治理政策工具影响企业非市场战略配置决策的机理及经济后果研究”(项目编号:72074035);中央高校基本科研业务费科研专项资助项目(项目编号:2019CJDK01XK09)。

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

银行借款是中国企业最主要的融资渠道,市场诚信环境是企业获取银行借款的重要影响因素(钱先航和曹廷求,2015<sup>[1]</sup>),而最权威的司法判决能否得到切实执行是反映市场诚信环境的“底线”,中国司法实践中的“执行难”是中国市场诚信环境有待改善的客观反映(黄忠顺,2017<sup>[2]</sup>)。最高人民法院于2013年发布了《关于公布失信被执行人名单信息的若干规定》(以下简称“失信被执行人名单制度”),试图利用政府实施的第三方声誉惩罚约束企业不诚信行为,为信贷市场创造更加诚信的环境。对于具备履行能力,但是通过各种手段拒绝履行法律判决的被执行人,人民法院有权对他们给予曝光并向各部门通报,直至失信被执行人履行生效法律文书规定的义务。失信被执行人名单制度将声誉与个人信用相结合,使拒不执行法律义务者承担相当的心理压力,迫使其主动履行法律义务。同时,对失信被执行人的第三方声誉惩罚将形成社会示范,使更多的人主动履行各种社会义务。既有研究证明了政府等权威第三方机构实施的声誉惩罚将给失信企业带来巨大的经济损失,同时对失信企业融资产生负向影响(Karpoff等,2008<sup>[3]</sup>;Verschoor等,2012<sup>[4]</sup>;Armour等,2017<sup>[5]</sup>)。那么,政府实施的第三方声誉惩罚是否还能对其他没有失信行为的企业产生影响呢?遗憾的是目前尚无这方面的实证研究。

基于此,笔者首先利用2010—2016年中国沪深A股上市公司的数据,基于第三方声誉惩罚理论对失信被执行人名单制度能否以及如何对企业银行借款产生影响进行实证检验;然后在此基础上,本研究还将探究企业自身风险以及所处地区金融发展差异对失信被执行人名单制度与企业银行借款关系的影响,并进一步检验失信被执行人名单制度对企业融资结构的影响以及设立退出机制是否会削弱失信被执行人名单制度对企业银行借款的影响。

## 二、文献综述及研究假设

(一)失信被执行人名单制度对企业银行借款的影响

Ang等(2015)<sup>[6]</sup>、雷宇(2016)<sup>[7]</sup>等发现声誉是一种非正式的治理机制,它有助于降低市场的交易

成本,促进市场主体间相互合作,而声誉惩罚是一种“负激励”,也有助于提高市场主体的合作水平(Gächter和Fehr,2002<sup>[8]</sup>;Sigmund,2007<sup>[9]</sup>;Raihani等,2010<sup>[10]</sup>),允许更大群体内实现稳定合作的演变。目前声誉惩罚的既有研究主要在二元互动的框架下进行,考察交易双方如何运用声誉惩罚来抑制机会主义行为,但在市场规模较大时,二元互动下的声誉惩罚仍存在诸多限制。第一,市场参与者可能并不了解自己所处市场上的每一次二元互动。作为直接观察的替代品,市场参与者可能会通过别人共享信息来了解合作伙伴的声誉(Gallo和Yan,2015<sup>[11]</sup>;Giardini和Vilone,2016<sup>[12]</sup>)。虽然共享信息被视为一种低成本的信息获取方式,但是它并不是完全免费的,因为它涉及评估信息真实性所花费的精力以及保留信息等成本(Hess和Hagen,2006<sup>[13]</sup>),这些成本阻止了市场上声誉惩罚信息的传播。第二,声誉惩罚可能会被机会主义者滥用。有时机会主义者会将声誉惩罚作为一种报复行为,利用散布谣言来对抗合作者(Oliver,1984<sup>[14]</sup>),被报复的威胁可能会削弱合作者进行声誉惩罚的意愿。第三,虽然既有研究通过理论分析普遍认为声誉惩罚是有效的市场诚信治理机制,但声誉惩罚也具有外部性,惩罚者自身也要承担高昂的成本,而第三方则从合作水平的提高中受益(Nowak等,2008<sup>[15]</sup>;Rankin等,2009<sup>[16]</sup>;Wu等,2009<sup>[17]</sup>)。因此,声誉惩罚通常被认为是真正的利他主义(Barclay,2006<sup>[18]</sup>),由市场参与者对机会主义者进行声誉惩罚代价高昂。

综上所述,既然声誉惩罚具有一定公共品属性,则可以由公权力担任声誉惩罚的“第三方”执行者来实施利他惩罚,提供声誉惩罚的“公共品”并承担声誉惩罚成本,而且,由公权力实施的声誉惩罚能够降低声誉惩罚错误率,提高声誉惩罚权威性和失信受罚的确定性。中国的失信被执行人名单制度就是典型的由公权力实施的第三方声誉惩罚机制,该制度不仅对失信被执行人起到了声誉惩罚作用,加大了企业的失信成本,还为人们未来的行为模式提供了相对确定的预期,有助于提高司法公信力和社会诚信水平(刘涛和朱燕,2013<sup>[19]</sup>),从而降低市场交易成本并有利于企业获得更多的银行借款(余泳泽等,2020<sup>[20]</sup>)。基于以上分析,本文提出假设1:

**H1:** 失信被执行人名单制度有助于企业获取银

行借款。

(二) 失信被执行人名单制度通过企业声誉对企业银行借款的影响

声誉在信贷市场中有重要意义, 具有良好声誉的企业可以向市场发送稳定经营且违约风险较低的信号, 从而获得银行更大的信任(叶康涛等, 2010<sup>[21]</sup>), 使企业更容易获得银行资金支持; 反之, 声誉较差的企业则要为失信行为付出沉重代价(石晓峰和仲秋雁, 2017<sup>[22]</sup>)。当企业受到负面声誉的冲击时, 这些负面信息将会向资本市场传递不利信号, 此时企业的债务违约风险上升, 为了避免债务违约事件发生, 银行的贷款意愿将会随之降低(陈雪等, 2021<sup>[23]</sup>)。失信被执行人名单制度出台后, 企业若出现失信行为将会受到更加严厉的第三方声誉惩罚, 其失信行为将通过网络、报纸等媒体曝光, 并在市场上迅速形成负面声誉, 降低银行及其他利益相关者对企业的信用评级, 大大增加企业失信的声誉成本。一方面, 失信被执行人名单制度的第三方声誉惩罚使失信企业更容易成为市场上的“反面典型”, 可以对其他企业发挥“替代强化”作用, 促使其他企业更重视保护和提升自己的声誉; 另一方面, 强化对失信行为的声誉惩罚使失信行为的声誉成本大于外部收益, 市场上的企业作为理性经济人可能改变决策偏好, 更倾向于诚信行为、严守契约, 从客观结果来看企业声誉得到了普遍提高, 有利于获取银行的信任。基于上述分析, 本文提出假设 2:

**H2:** 失信被执行人名单制度可以通过提高企业声誉帮助企业获得更多的银行借款。

(三) 失信被执行人名单制度通过企业内部控制质量对企业银行借款的影响

内部控制是企业治理中一项重要的机制, 对企业提升经营绩效和降低融资约束等方面有重要影响(赵玲和黄昊, 2022<sup>[24]</sup>)。由于企业内部控制质量在很大程度上受制于企业管理层的影响, 因此, 企业管理层所处社会文化环境将会对企业内部控制质量产生一定作用。失信被执行人名单制度通过第三方声誉惩罚改善了市场诚信环境, 增强了人们的信用意识, 推动“守信”成为一种普世价值观, 并内化为企业经营的道德标准(Pruckner 和 Sausgruber, 2013<sup>[25]</sup>), 约束了企业管理层的自利行为。因此, 失信被执行人名单制度作为第三方声誉惩罚还可以发挥“激励机

制”, 增强企业管理层诚实守信的心理激励, 提升企业管理层信用意识和道德素养, 倒逼企业强化内部治理, 实施更严格的内部控制(Doyle 等, 2007<sup>[26]</sup>), 进一步降低“机会主义”及“逆向选择”风险(杨雄胜, 2005<sup>[27]</sup>)。另外, 严格的内部控制还可以有效抑制企业管理层的盈余管理、报表粉饰等“失信行为”(陈汉文等, 2019<sup>[28]</sup>), 提高会计信息质量, 从而缓解信息不对称, 增加企业获得的银行借款。基于上述分析, 本文提出假设 3:

**H3:** 失信被执行人名单制度可以通过提高企业内部控制质量帮助企业获得更多的银行借款。

(四) 不同风险水平的企业中失信被执行人名单制度对企业银行借款的影响

除了外部制度, 资金实力、市场地位等企业自身风险水平也是影响银行借款的重要因素, 那么, 企业自身风险水平的高低是否会影响失信被执行人名单制度的第三方声誉惩罚效果呢? 李卓松(2018)<sup>[29]</sup>的研究发现, 较高的企业风险水平将会提升其融资成本。银行在放贷前会利用企业财务信息对企业的风险水平进行评估, 风险较高的企业更有可能因为业绩下行或现金流波动过大而无法履行还款义务(于晓红和卢相君, 2012<sup>[30]</sup>), 而风险较低的企业往往具备更好的还款能力且违约率较低(程砚秋, 2015<sup>[31]</sup>), 银行也可以减少对企业的事后监督。因此, 当政府通过第三方声誉惩罚降低了市场交易成本时, 风险水平较低的企业更容易成为银行的放贷对象, 此时失信被执行人名单制度更有助于风险水平较低的企业获得银行借款。基于上述分析, 本文提出假设 4:

**H4:** 企业自身风险水平较低时, 企业银行借款受到失信被执行人名单制度的影响更显著。

(五) 不同地区金融发展水平下失信被执行人名单制度对银行借款的影响

地区金融发展水平也是影响企业获取外部融资的重要因素(Claessens 和 Laeven, 2003<sup>[32]</sup>), 陈耿等(2015)<sup>[33]</sup>的研究证明, 当企业所处地区的金融发展水平较高时, 以银行为主的正规金融机构数量较多, 拓宽了企业的融资渠道, 缓解了企业的融资压力, 此时, 即使失信被执行人名单制度利用第三方声誉惩罚措施提升了市场诚信水平, 其对企业银行借款的作用也将相对减弱。此外, Diamond(1984)<sup>[34]</sup>、Demirgüç-Kunt 和 Maksimovic(2002)<sup>[35]</sup>等的研究发

现,金融发展水平较高地区的借贷配套体系完善,贷款的事前信息收集以及事后监督执行的成本相对较低,缓解了银企之间的信息不对称,在一定程度上弱化了失信被执行人名单制度的作用;反之,在金融发展水平较低的地区,银企之间存在较严重的信息不对称,失信被执行人名单制度可以通过第三方声誉惩罚更明显地降低交易成本,缓解制约银行借款的信息不对称。因此,对于企业银行借款而言,第三方声誉惩罚的制度环境和地区金融发展水平可以相互替代。基于上述分析,本文提出假设5:

**H5:** 企业所处地区的金融发展水平较低时,企业银行借款受到失信被执行人名单制度的影响更显著。

### 三、研究设计

#### (一) 样本与数据来源

最高人民法院于2010年发布了《关于限制被执行人高消费的若干规定》,开始对失信行为实施“经济型惩戒”,2013年建立的失信被执行人名单制度对失信行为追加第三方声誉惩罚,是对强制性消费约束的“补充惩戒”。为了剥离出失信被执行人名单制度对于企业银行借款的净影响,本文以2013年失信被执行人名单制度出台为实验事件,选取实验事件发生前后相同的时间区域,即2010—2016年的A股上市公司作为研究样本。由于失信被执行人名单中的企业一般是有限公司或有限责任公司,而上市公司受到严格的监管,进入失信被执行人名单的可能性较小。自2013年以来,失信被执行人的数量达到了691万,其中上市公司数量不足100家,且绝大多数为ST企业。因此,上市公司受失信被执行人名单制度的直接影响较小,这使我们可以利用上市公司数据研究失信被执行人名单制度对非失信企业银行借款的影响。本文所需数据来自CSMAR数据库,对样本进行了如下处理:(1)剔除ST或\*ST企业;(2)剔除金融业上市公司样本;(3)剔除重要变量缺失的样本;(4)对连续变量进行上下1%的Winsor缩尾处理。经过上述筛选,本文最终得到14260个观测值作为研究样本。

#### (二) 模型设定与变量定义

由于2013年在全国范围内实施的“失信被执行人名单制度”是一项对所有企业均有影响的制度,无法剥离出完全不受该制度影响的对照组。根据Vig(2013)<sup>[36]</sup>、Campello和Larrain(2016)<sup>[37]</sup>等的做法,

本文根据该制度对不同特质企业影响的差异性来构建实验组和对照组。具体地,本文基于企业固定资产占比高低来判断企业银行借款是否更容易受到失信被执行人名单制度的影响。由于抵押品可以大大缓解银企之间的逆向选择与道德风险问题,起到风险控制的作用(尹志超和甘犁,2011<sup>[38]</sup>),因此,企业可以通过向银行提供抵押或担保品提高企业的贷款信用评级,有助于其获得银行借款。在众多抵押形式中,固定资产的交易市场最为成熟,且价值波动小,变现容易,因此银行对于机器、土地等固定资产的抵押尤为重视(Gregory和Tenev,2001<sup>[39]</sup>)。如果企业的固定资产占比较低,则往往缺乏抵押品,可能无法顺利获得银行借款(Ayyagari等,2010<sup>[40]</sup>)。因此,固定资产占比较低的企业所能获得的银行借款对失信被执行人名单制度更加敏感,适合作为本研究的实验组。

在模型设定上,本文将更容易受到失信被执行人名单制度影响的企业作为“实验组”,将不容易受到失信被执行人名单制度影响的企业设为“对照组”。由于在该制度开始实施的2013年企业有可能已经对制度做出了一定反应,为了减少选择性偏误,本文选取2010—2012年企业的固定资产平均占比作为分组依据。为消除时间和个体之间的差异,本文运用双重差分法进行实证检验,模型如下:

$$Loan_{it} = \beta_0 + \beta_1 Treat_i \times Post_t + \beta_2 Treat_i + \beta_3 Post_t + \beta_4 Controls_{it} + \delta_i + \gamma_t + \varepsilon_{it} \quad (1)$$

其中,Loan表示企业的银行借款,Treat为分组虚拟变量,为了凸显失信被执行人名单制度对企业银行借款影响的差异性,本文将2010—2012年固定资产占比平均值最低的1/3的企业设为实验组,取值为1,反之将固定资产占比平均值最高的1/3的企业作为对照组,取值为0。Post为年份虚拟变量,本文将2010—2013年取值为0,2014—2016年取值为1。Controls为一系列企业层面控制变量。模型中 $\beta_1$ 系数反映了失信被执行人名单制度实施前后实验组和对照组企业所获得的银行借款的变化情况。

被解释变量企业银行借款利用总借款与总资产的比值来衡量。参照徐思等(2019)<sup>[41]</sup>的做法,企业层面控制变量包括企业规模(Size)、企业负债水平(Lev)、企业会计业绩(Roe)、成长能力(Growth)、股权集中度(Top1)、两职合一(Dual)以及产权性质(Soe)。具体变量定义参照表1。

表1 变量定义

被解释变量	<i>Loan</i>	企业获得银行借款	总借款/总资产
主要解释变量	<i>Treat</i>	分组虚拟变量	固定资产占比最低的1/3的企业为1, 固定资产占比最高的1/3的企业为0
	<i>Post</i>	时间虚拟变量	2014年及之后取1
控制变量	<i>Size</i>	企业规模	企业总资产取对数
	<i>Lev</i>	企业负债水平	总负债/总资产
	<i>Roe</i>	企业会计业绩	净利润/股东权益
	<i>Growth</i>	企业成长能力	(本期营业收入-上期营业收入)/上期营业收入
	<i>Top1</i>	股权集中度	第一大股东持股比例
	<i>Dual</i>	两职合一	董事长和总理由同一人担任时取1, 反之为0
	<i>Soe</i>	产权性质	国有企业为1, 非国有企业为0

#### 四、实证结果及分析

##### (一) 描述性统计

表2为主要变量的描述性统计。样本中所有企业

总借款占比的平均值约为0.158, 标准差为0.143, 最大值是平均值的3.72倍, 说明不同企业的银行借款水平具有较大差异。企业获得银行借款水平的平均值和中位数相差不大, 说明样本数据基本呈正态分布。

表2 描述性统计

变量	样本数	平均值	中位数	标准差	最小值	最大值
<i>Loan</i>	14 260	0.158	0.132	0.143	0.000	0.587
<i>Treat</i>	9 354	0.481	0.000	0.500	0.000	1.000
<i>Law</i>	14 260	0.442	1.000	0.497	0.000	1.000
<i>Size</i>	14 260	22.037	21.871	1.301	19.251	25.979
<i>Roe</i>	14 260	0.064	0.068	0.119	-0.697	0.388
<i>Lev</i>	14 260	0.451	0.448	0.219	0.047	0.960
<i>Growth</i>	14 260	0.214	0.115	0.559	-0.573	3.894
<i>Top1</i>	14 260	0.358	0.338	0.151	0.084	0.751
<i>Dual</i>	14 260	0.238	0.000	0.426	0.000	1.000
<i>Soe</i>	14 260	0.432	0.000	0.495	0.000	1.000

##### (二) 实证结果分析

本文从失信被执行人名单制度出台这一准自然实验出发, 运用双重差分的方法考察失信被执行人名单制度对企业银行借款的影响, 结果如表3所示。本文在列(1)中只检验解释变量的影响, 列(2)控制了年度固定效应和行业固定效应, 列(3)在加入控制变量的同时控制了年度固定效应和行业固定效应。从表3列(1)到列(3)可以发现, 核心解释变量  $Treat \times Post$  的系数和显著性均没有发生太大变化。列(3)的回归结果证明: 失信被执行人名单制度通过第三方声誉惩罚显著提高了

企业的银行借款, 验证了 H1。

表3 失信被执行人名单制度与企业银行借款的回归结果

	(1)	(2)	(3)
$Treat \times Post$	0.037 2*** (7.37)	0.037 7*** (7.52)	0.018 6*** (4.64)
<i>Treat</i>	-0.099 5*** (-13.86)	-0.083 2*** (-9.81)	-0.052 7*** (-8.54)
<i>Post</i>	-0.030 6*** (-8.32)	-0.031 2*** (-5.52)	-0.044 2*** (-9.35)
<i>Size</i>			0.009 7*** (3.81)



续前表

	(1)	(2)	(3)
<i>Roe</i>			-0.115 9*** (-7.06)
<i>Lev</i>			0.418 5*** (24.95)
<i>Growth</i>			0.000 4 (0.23)
<i>Top1</i>			-0.021 1 (-1.43)
<i>Dual</i>			0.005 5 (1.30)
<i>Soe</i>			-0.021 7*** (-4.11)
<i>Ind</i>	NO	YES	YES
<i>Year</i>	NO	YES	YES
<i>_Cons</i>	0.224 1*** (42.22)	0.203 4*** (7.55)	-0.169 5*** (-3.13)
<i>R</i> <sup>2</sup>	0.084 7	0.245 6	0.576 5
<i>N</i>	9 354	9 354	9 354

注：括号内是经过企业层面群聚调整的 *t* 值；\*、\*\*、\*\*\* 分别表示 10%、5%、1% 的显著性水平。下同。

### (三) 稳健性检验

#### 1. 平行趋势检验。

双重差分模型有效的潜在前提是实验组和对照组在制度实施之前具有相同的变化趋势。本文通过构建模型 (2) 来检验对照组和实验组是否符合平行趋势假设 (Bertrand 和 Mullainathan, 2003<sup>[42]</sup>; Fang 等, 2017<sup>[43]</sup>)。模型设定如下：

$$\begin{aligned}
 Loan_{it} = & \alpha + \beta_1 Treat_i \times Before^{-3} + \beta_2 Treat_i \times Before^{-2} \\
 & + \beta_3 Treat_i \times Current + \beta_4 Treat_i \times After^1 + \beta_5 Treat_i \\
 & \times After^2 + \beta_6 Treat_i \times After^3 + \tau_1 Before^{-3} \\
 & + \tau_2 Before^{-2} + \tau_3 Current + \tau_4 After^1 + \tau_5 After^2 \\
 & + \tau_6 After^3 + \gamma_1 Treat_i + \gamma_2 Controls_{it} \\
 & + \delta_i + \gamma_i + \varepsilon_{it} \quad (2)
 \end{aligned}$$

其中, *Before*<sup>-3</sup> 在 2010 年时取值为 1, 其余为 0; *Before*<sup>-2</sup> 在 2011 年时取值为 1, 其余为 0; 2013 年为失信被执行人名单制度开始实施年份, 此时 *Current* 取值为 1, 其余为 0; *After*<sup>1</sup> 在 2014 年时取值为 1, 其余为 0; *After*<sup>2</sup> 在 2015 年时取值为 1, 其余为 0; *After*<sup>3</sup> 在 2016 年时取值为 1, 其余为 0。

从表 4 的回归结果中可以看出, 制度实施前的系数  $\beta_1$  和  $\beta_2$  均不显著, 说明制度实施前对照组和实验

组没有明显差异。制度实施后的系数  $\beta_3$ - $\beta_5$  均显著为正, 说明失信被执行人名单制度实施后对照组和实验组企业的银行借款产生显著差异, 该制度对固定资产较低的企业产生了影响, 总体上平行趋势成立。

表 4 平行趋势检验

	<i>Loan</i>
<i>Treat</i> × <i>Before</i> <sup>-3</sup>	-0.007 1 (-1.32)
<i>Treat</i> × <i>Before</i> <sup>-2</sup>	-0.004 1 (-1.08)
<i>Treat</i> × <i>Current</i>	0.008 8** (2.49)
<i>Treat</i> × <i>After</i> <sup>1</sup>	0.016 5*** (3.94)
<i>Treat</i> × <i>After</i> <sup>2</sup>	0.014 1*** (2.98)
<i>Treat</i> × <i>After</i> <sup>3</sup>	0.009 8 (1.62)
<i>Before</i> <sup>-3</sup>	0.025 1*** (6.74)
<i>Before</i> <sup>-2</sup>	0.012 2*** (4.42)
<i>Current</i>	-0.008 2*** (-3.14)
<i>After</i> <sup>1</sup>	-0.024 5*** (-7.76)
<i>After</i> <sup>2</sup>	-0.029 4*** (-7.99)
<i>After</i> <sup>3</sup>	-0.039 3*** (-8.23)
<i>Controls</i>	YES
<i>Ind</i>	YES
<i>Year</i>	YES
<i>R</i> <sup>2</sup>	0.363 1
<i>N</i>	9 354

#### 2. 剂量效应检验。

为了保证实证结果的稳健性, 本文改变实验组和对照组的构造方法 (钱雪松和方胜, 2017<sup>[44]</sup>)。本文保持对照组不变, 将固定资产处于中间 1/3 的企业作为实验组, 重新检验失信被执行人名单制度对企业银行借款的影响, 结果如表 5 列 (1) 所示。本文还将

对照组改为固定资产最高的 1/2 的企业，实验组改为固定资产最低的 1/2 的企业进行回归，其结果如表 5 列 (2) 所示。从结果可以看出，在两种新分组的情况下，该制度仍然在 1% 的水平上显著提高了企业的银行借款。同时，两种结果都表现出剂量效应，即对照组和实验组之间的差异逐渐缩小时，制度所产生的效果在两组间的差异也会随之减少，说明失信被执行人名单制度对企业银行借款的提升作用存在“中等区间”。表 5 列 (1) 与列 (2) 中  $Treat \times Post$  的系数均小于表 4 列 (3) 基准回归中的系数，说明制度作用因分组不同被削弱，与理论预期相符。

3. 倾向得分匹配法。

因为制度本身可能会存在非随机选择因素，从而产生一定的内生性问题，为此，本文进一步使用倾向得分匹配法和双重差分相结合的方法 (PSM-DID) 进行检验。倾向得分匹配法有助于改善样本选择偏差问题，保证处理组和对照组的可比性。其基本思想

是：通过匹配构建一个与固定资产占比较低企业的主要特征“尽可能一致”的固定资产占比高企业作为对照组，使匹配后两个样本组的配对企业之间仅在固定资产占比方面有所不同。具体地，本文先将控制变量用于倾向得分匹配法，对样本进行 1 : 4 有放回近邻匹配方法匹配，再用双重差分法进行估计，其结果如表 5 列 (3) 所示，可见结果并未发生实质性变化。

4. 添加地区层面控制变量。

考虑到地区层面差异可能会影响企业获得的银行借款，本文添加一系列地区层面控制变量，包括：(1) 经济发展水平。由于银行借款受到国家宏观经济发展的影响较大，本文使用平减后的城市人均 GDP 取对数 ( $GDP\_per$ ) 进行衡量。(2) 法制化水平。金融契约的界定和实施很大程度上依赖于法律的保护，本文使用各省份每万人律师数 ( $Law$ ) 进行衡量。表 5 列 (4) 的回归结果依然和基准回归相符。

表 5 稳健性检验

	剂量效应		PSM-DID	添加地区层面控制变量
	(1)	(2)	(3)	(4)
$Treat \times Post$	0.015 1 *** (4.04)	0.011 3 *** (3.62)	0.020 5 *** (4.91)	0.019 7 *** (4.75)
$Treat$	-0.028 6 *** (-6.08)	-0.033 8 *** (-7.92)	-0.055 1 (-8.87)	-0.054 4 *** (-8.45)
$Post$	-0.039 0 *** (-8.65)	-0.032 4 *** (-8.53)	-0.045 7 *** (-9.21)	-0.041 9 *** (-8.13)
$Law$				-0.001 3 ** (-1.98)
$GDP\_per$				0.002 4 (0.67)
$Controls$	YES	YES	YES	YES
$Ind$	YES	YES	YES	YES
$Year$	YES	YES	YES	YES
$R^2$	0.606 9	0.570 2	0.575 7	0.590 3
$N$	9 040	14 260	8 740	6 375

(四) 作用机制检验

为了检验失信被执行人名单制度对企业银行借款的作用渠道，本文根据 Baron 和 Kenny (1986)<sup>[45]</sup> 提出的依次检验法构建中介效用模型：

$$Mediacy_{it} = \alpha_0 + \alpha_1 Treat_{it} + \alpha_2 Post_{it} + \alpha_3 Treat_{it} \times Post_{it} + \alpha_4 Controls_{it} + \delta_i + \gamma_t + \varepsilon_{it} \quad (3)$$

$$Loan_{it} = \beta_0 + \beta_1 Treat_{it} + \beta_2 Post_{it} + \beta_3 Treat_{it} \times Post_{it} + \beta_4 Mediacy_{it} + \beta_5 Controls_{it} + \delta_i + \gamma_t + \varepsilon_{it} \quad (4)$$

其中, *Mediacy* 为一系列中介变量, 其他变量含义同模型 (1), 所有回归结果的标准误差均在企业层面进行聚类调整。

1. 替代强化机制检验。

本文参考甄红线和王三法 (2021)<sup>[46]</sup>的做法, 选择了 14 个与企业声誉相关的指标<sup>①</sup>, 采用因子分析法计算出企业声誉得分 (*REP*)。企业声誉得分越高, 说明企业声誉越好。模型 (3) 的回归结果如表 6 列 (1) 所示, 模型 (4) 的回归结果如表 6 列 (2) 所示。 $\alpha_3$ 、 $\beta_3$  和  $\beta_4$  的系数都通过了显著性检验, 说明失信被执行人名单制度可以利用第三方声誉惩罚对企业起到威慑作用并约束企业的失信行为, 提高企业声誉, 从而帮助企业获得银行借款, 验证了 H2。

2. 激励机制检验。

企业内部控制质量是一个宽泛的概念, 为保证实证结果的可靠性, 本文利用三种变量来衡量: (1) 利用内部控制是否存在缺陷 (*Inc*)。当企业自我评价报告中表明自身存在重大缺陷时, 即认为企业内部控制存在缺陷, 虚拟变量 *Inc* 取值为 0, 反之取值为 1。(2) 迪博-中国上市公司内部控制质量指数 (*Dib*)。迪博-中国上市公司内部控制质量指数是外界对企业内部控制质量的评价, 反映了企业风险管控水平和内部控制能力, 该指数的值越大, 表明内部控制目标越能得到贯彻落实, 内部控制质量越高。(3) 借鉴 Roychowdhury (2006)<sup>[47]</sup>的模型来测量企业真实盈余

管理水平 (*RM*), 该指标的值越大, 代表真实盈余管理水平越高, 企业内部控制质量越差。

表 7 的回归结果证明, 无论采用哪种方式衡量企业内部控制质量,  $\alpha_3$ 、 $\beta_3$  和  $\beta_4$  的系数都显著, 说明存在部分中介效应, 失信被执行人名单制度所代表的第三方声誉惩罚机制将会提升企业信用意识, 倒逼企业提高内部控制质量, 降低企业盈余管理, 从而帮助企业获得银行借款, 验证了 H3。

表 6 替代强化机制检验

被解释变量	(1)	(2)
	<i>REP</i>	<i>Loan</i>
<i>Treat</i> × <i>Post</i>	0.025 5* (1.73)	0.013 9*** (3.06)
<i>Treat</i>	-0.011 3 (-0.57)	-0.045 1*** (-7.02)
<i>Post</i>	-0.124 0*** (-7.35)	-0.048 5*** (-8.63)
<i>REP</i>		0.014 1** (1.98)
<i>Controls</i>	YES	YES
<i>Ind</i>	YES	YES
<i>Year</i>	YES	YES
<i>R</i> <sup>2</sup>	0.716 0	0.588 5
<i>N</i>	7 123	7 123

表 7 激励机制检验

被解释变量	内部控制是否存在缺陷		迪博指数		真实盈余管理水平	
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Inc</i>	<i>Loan</i>	<i>Dib</i>	<i>Loan</i>	<i>RM</i>	<i>Loan</i>
<i>Treat</i> × <i>Post</i>	0.072 7*** (3.76)	0.016 9*** (4.12)	0.160 3*** (2.84)	0.017 8*** (4.41)	-0.026 3*** (-2.73)	0.021 4*** (5.15)
<i>Treat</i>	-0.006 5 (-0.38)	-0.048 0*** (-7.93)	-0.094 22* (-1.72)	-0.053 2*** (-8.46)	0.030 2*** (2.50)	-0.054 9*** (-8.44)
<i>Post</i>	-0.388 6*** (-18.20)	-0.034 5*** (-6.61)	-0.575 3*** (-7.99)	-0.041 2*** (-8.56)	0.013 7 (1.21)	-0.045 7*** (-9.33)
<i>Inc</i>		0.008 3** (2.47)				

① 这些指标具体包括: 消费者和社会角度的企业资产、主营业务收入、净利润、总资产收益率和主营业务收入市场占有率; 债权人角度的资产负债率、流动比率、长期负债比; 股东角度的每股收益、每股股利、股东大会出席率和是否为国际四大会计师事务所审计; 企业角度的董事会规模; 声誉投入角度的无形资产净额的对数。

续前表

被解释变量	内部控制是否存在缺陷		迪博指数		真实盈余管理水平	
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Inc</i>	<i>Loan</i>	<i>Dib</i>	<i>Loan</i>	<i>RM</i>	<i>Loan</i>
<i>Dib</i>				0.006 3*** (4.58)		
<i>RM</i>						-0.012 5* (-1.69)
<i>Controls</i>	YES	YES	YES	YES	YES	YES
<i>Ind</i>	YES	YES	YES	YES	YES	YES
<i>Year</i>	YES	YES	YES	YES	YES	YES
<i>R<sup>2</sup></i>	0.146 8	0.596 0	0.139 6	0.575 9	0.149 0	0.585 7
<i>N</i>	8 363	8 363	9 256	9 256	8 287	8 287

(五) 异质性检验

为了考察失信被执行人名单制度对不同风险水平以及不同地区企业的作用差异，本文在双重差分模型的基础上继续构建三重差分模型进行进一步的识别检验。具体模型构建如下：

$$\begin{aligned}
 Loan_{it} = & \beta_0 + \beta_1 Treat_i \times Post_t \times dz + \beta_2 Treat_i \times Post_t \\
 & + \beta_3 Treat_i + \beta_4 Post_t + \beta_5 Treat_i \times dz \\
 & + \beta_6 Controls_{it} + \delta_i + \gamma_t + \varepsilon_{it} \quad (5)
 \end{aligned}$$

其中，*dz* 依次表示企业风险与地区金融发展水平。本文利用企业财务杠杆 (*DFL*) 与行业竞争度 (*HHI*) 衡量企业风险，利用普惠金融指数 (*Fin*) 衡量地区金融发展水平，其他变量同模型 (1)。本文主要关注三重差分交乘项  $\beta_1$ ，若其通过显著性检验，则说明失信被执行人名单制度在不同特征的企业中存在明显差异。

1. 企业风险水平的影响分析。

传统企业风险主要包括财务风险和经营风险，本文使用企业财务杠杆率的大小来衡量其财务风险的高低，财务杠杆率越高则相应的企业财务风险越大；同时，利用 *HHI* 指数<sup>①</sup> 衡量企业所处行业的竞争度，*HHI* 指数越高，说明市场集中度越高，行业内的市场竞争度越小，此时企业经营风险较低。本文以财务杠杆的年度行业均值为界将样本分为财务风险较低的企业

(赋值为 1) 和财务风险较高的企业 (赋值为 0)，生成虚拟变量 *DFL*，再根据企业所处行业的 *HHI* 指数生成变量 *HHI*，分别和双重差分项交乘构建三重差分模型。结果如表 8 列 (1) 和列 (2) 所示，三重差分交乘项系数显著为正，说明当失信被执行人名单制度通过第三方声誉惩罚降低市场交易成本时，风险水平较低的企业更容易获得银行借款，失信被执行人名单制度的第三方声誉惩罚机制主要对企业银行借款起到“锦上添花”的作用，验证了 H4。

表 8 异质性检验

	<i>Loan</i>		
	企业风险		地区金融发展水平
	财务风险	经营风险	普惠金融指数
	(1)	(2)	(3)
<i>Treat</i> × <i>Post</i> × <i>DFL</i>	0.025 6*** (4.18)		
<i>Treat</i> × <i>Post</i> × <i>HHI</i>		0.065 2* (2.12)	
<i>Treat</i> × <i>Post</i> × <i>Fin</i>			-0.046 6** (-2.08)
<i>Treat</i> × <i>Post</i>	-0.000 8 (-0.14)	0.013 3*** (3.04)	0.257 6** (2.15)
<i>Controls</i>	YES	YES	YES

① 本文基于所有者权益构建市场竞争度指标。由于文中样本数据只保留了固定资产占比最高 1/3 和最低的 1/3 企业，衡量行业的市场竞争度时若只根据回归样本进行计算则会产生一定误差，故本文使用全样本 (即所有上市公司 A 股企业) 对市场竞争度 (*HHI*) 进行计算。

续前表

	Loan		
	企业风险		地区金融发展水平
	财务风险	经营风险	普惠金融指数
	(1)	(2)	(3)
<i>Ind</i>	YES	YES	YES
<i>Year</i>	YES	YES	YES
<i>R</i> <sup>2</sup>	0.582 6	0.577 3	0.588 3
<i>N</i>	8 429	9 340	6 831

注：省略了 *Treat*、*Post*、*Treat*×*dz* 和常数项的回归结果。

## 2. 地区金融发展水平的影响分析。

本文使用北京大学数字普惠金融指数（郭峰等，2020<sup>[48]</sup>）来衡量地区金融发展水平，取其自然对数定义为 *Fin*，构造三重差分项进行检验。表 8 列（3）的结果显示三重差分系数显著为负，说明地区金融发展水平和第三方声誉惩罚的制度环境呈现替代关系，在金融发展水平较低地区更需要第三方声誉惩罚机制弥补金融体系的“缺位”，失信被执行人名单制度对银行借款的作用更显著，验证了 H5。

## 五、进一步检验

（一）通过其他融资渠道对失信被执行人名单制度的银行借款提升作用进行检验

除了银行借款，商业信用也是企业的一种重要融资渠道<sup>①</sup>。商业信用的双方是密切往来的商业伙伴，拥有共同的商业利益，属于信用融资范畴。相比银行借款而言，商业信用融资具有手续便利、门槛低的特点（Biais 和 Gollier，1997<sup>[49]</sup>；Petersen 和 Rajan，1997<sup>[50]</sup>），但也有期限较短、融资金额较少等缺点。那么，失信被执行人名单制度对企业商业信用融资有何影响呢？

本文利用 2010—2016 年的中国沪深 A 股上市公司作为研究样本，被解释变量“企业商业信用”利用企业商业信用与总资产的比值来衡量，模型设定同模型（1）。从表 9 列（1）的回归结果中可以看出，失信被执行人名单政策的建立显著降低了企业获得的商业信用融资。可能的原因是：石晓军和李杰（2009）<sup>[51]</sup>、赵胜民和张博超（2019）<sup>[52]</sup>等发现，在既定的企业融资需求下，银行借款和商业信用融资是

相互替代的关系，对于那些难以获得银行融资的企业，他们更倾向于获取商业信用融资。失信被执行人名单制度出台后，由政府实施的第三方声誉惩罚提高了司法公信力和社会诚信水平，企业内部控制质量提高，从资本市场中获得的银行借款增加，此时对于商业信用融资的需求显著减少。因此，失信被执行人名单制度增加了企业获得的银行借款，减少了商业信用融资，改变了企业融资结构。

（二）失信被执行人名单制度银行借款提升作用的动态检验

中国对失信被执行人的声誉惩罚是一个动态变化的过程，失信被执行人名单制度也经历了数次修改，不同阶段对失信被执行人的惩戒措施有所不同，因此，对失信被执行人名单制度的银行借款提升作用还需进行动态检验。

2017 年 3 月，最高人民法院通过了《最高人民法院关于修改〈最高人民法院关于公布失信被执行人名单信息的若干规定〉的决定》（以下简称《修改》），在失信被执行人的第三方声誉惩罚中导入了“退出”机制。《修改》明确了纳入失信被执行人名单的期限为两年，若情节严重或具有多项失信行为可延长一至三年。那么，这种带有退出机制的第三方声誉惩罚是否会降低失信被执行人名单制度的惩罚力度，削弱第三方声誉惩罚对企业银行借款的提升作用呢？

本文使用 2015—2019 年的中国沪深 A 股上市公司数据检验第三方声誉惩罚退出机制是否会对企业银行借款产生影响，设置新的时间虚拟变量 *law*2017，在 2017 年之前取 0，2017 年及以后取 1。表 9 列（2）的结果显示，*Treat*×*Law*2017 的系数并不显著，说明第三方声誉惩罚退出机制对于企业银行借款没有显著影响。由于网络、电视等现代传媒具有传播速度快、范围广等特点，数年的“曝光”足以产生“公告效应”，让失信被执行人名誉扫地，形同“社会死亡”。即使《修改》让失信被执行人在不履行法律义务的前提下 2~5 年后仍然能自动从失信被执行人“黑名单”中撤销，但曾经失信的不良记录无法抹去，在信息时代中失信行为更加无所遁形，失信会使企业声誉受到几乎不可挽回的重大损失。因此，时间限定的第三方声誉惩罚并不会弱化声誉惩罚机制的作用。

① 商业信用是指在商品交易中由于延期付款或预收货款所形成的企业间的借贷关系。具体形式包括应付账款、应付票据、预收账款等。

表9 失信被执行人名单制度的  
企业银行借款提升作用的进一步检验

被解释变量	商业信用	2017年《修改》 (2015—2019年)
	(1)	(2)
<i>Treat</i> × <i>Law</i>	-0.015 2*** (-4.51)	
<i>Treat</i> × <i>Law</i> 2017		-0.002 4 (-0.54)
<i>Treat</i>	0.052 4*** (8.82)	-0.041 1*** (-6.67)
<i>Law</i>	0.011 3** (2.17)	
<i>Law</i> 2017		0.011 9*** (3.28)
<i>Controls</i>	YES	YES
<i>Ind</i>	YES	YES
<i>Year</i>	YES	YES
<i>R</i> <sup>2</sup>	0.506 3	0.530 5
<i>N</i>	9 354	5 929

## 六、研究结论与展望

### (一) 研究结论

虽然已有学者证明了当企业受到政府声誉惩罚时,其损失将远远高于罚款金额(Karpoff, 2008<sup>[3]</sup>; Verschoor, 2012<sup>[4]</sup>),但迄今缺乏关于政府的声誉惩罚是否会对其他未受罚企业产生影响的实证。本文基于第三方声誉惩罚理论,将2013年失信被执行人名单制度的出台作为自然实验,以2010—2016年中国大陆A股上市公司作为研究样本,实证检验了失信被执行人名单制度对企业银行借款的影响,得出如下主要研究结论。

第一,失信被执行人名单制度有助于企业获得银行借款。失信被执行人名单制度利用政府担当第三方声誉惩罚的执行人,提高了声誉惩罚的权威性,有助于提高社会诚信水平,降低市场交易成本,从而提升了企业银行借款。更换分组方法、改用倾向得分匹配法、添加地区层面控制变量等一系列稳健性检验均支持了这一结论。

第二,失信被执行人名单制度通过倒逼企业提高声誉、强化内部控制质量提高企业银行借款。失信被执行人名单制度通过第三方声誉惩罚增加了企业失信

的声誉成本,企业作为理性经济人可能会改变决策偏好,从而在客观上提高企业声誉,增强银行放贷意愿;失信被执行人名单制度通过政府实施第三方声誉惩罚,改善了市场诚信环境,增强了企业管理层的信用意识,使其实施更严格的内部控制,从而提高会计信息质量,帮助企业获得银行借款。

第三,企业风险水平越低,失信被执行人名单制度对企业银行借款的作用越显著。企业拥有较低的风险水平意味着其具备较好的还款能力且违约率较低,此时失信被执行人名单制度通过第三方声誉惩罚改善市场环境后,风险水平低的企业更容易成为银行的放款对象。另外,地区金融发展水平与失信被执行人名单制度之间存在替代关系,在金融不发达的地区,失信被执行人名单制度能够通过政府主导的第三方声誉惩罚降低市场交易成本,弥补金融发展的不足,企业银行借款受到失信被执行人名单制度的影响也更显著。

第四,失信被执行人名单制度也影响了企业其他融资渠道。本文通过实证检验证明了失信被执行人名单制度显著降低了企业的商业信用融资规模,改变了企业融资结构。在对失信被执行人名单制度修改产生的影响进行检验后,本文发现声誉具有“建难毁易”的特征,声誉惩罚的退出机制并不会弱化第三方声誉惩罚的企业银行借款提升作用。

### (二) 管理启示

基于失信被执行人名单制度对企业银行借款影响机理的研究结论,得到如下主要管理启示。

第一,确保失信被执行人名单制度的有效实施。应严格按照程序将失信人员纳入“黑名单”,同时充分利用信息技术进行数据互联,与个人和企业的信用档案实现失信信息共享,确保失信被执行人名单制度对企业生产经营活动的规范作用。

第二,完善政府主导的第三方信用惩罚措施。政府主导的第三方信用惩罚是缓解企业融资约束、激活市场活力的重要途径。2019年7月国务院办公厅发布了《关于加快推进社会信用体系建设构建以信用为基础的新型监管机制的指导意见》,指出要加强完善失信联合惩戒机制,深入开展失信联合惩戒。本文结论为构建良好的失信联合惩戒提供了理论依据,以消费约束为代表的经济型惩戒和以声誉惩罚为代表的社会型惩戒相结合是未来第三方信用惩戒的正确发展

方向,有助于形成“一处失信、处处受限”的信用惩戒格局。

第三,降低企业自身风险水平。解决企业融资难问题不仅依靠外部制度环境的改善,还依赖于企业自身财务信息质量改善以及市场竞争力提高。因此,企业在日常经营中应注意降低其经营风险和财务风险,使制度环境改善发挥更大的企业银行借款提升作用。

### (三) 局限与展望

本文实证检验了失信被执行人名单制度对企业银行借款的影响,丰富了第三方声誉惩罚的相关研究,

未来在该领域可以从以下两个方面进行更加深入的研究:其一,本文以中国企业为样本研究了失信被执行人名单制度对企业银行借款的影响,但不同国家对于失信被执行人的声誉惩罚方式有所不同,因此,本文结论对于其他国家是否适用仍需进一步实证检验,未来研究可以将其他国家企业作为研究样本对本文结论进行进一步验证。其二,本文检验了失信被执行人名单制度对企业银行借款的影响,但对受信用环境影响的企业其他方面尚未加以研究,未来可以针对企业生产经营合规性、会计信息披露质量等方面进行拓展研究。

## 参考文献

- [1] 钱先航,曹廷求.法律、信用与银行贷款决策——来自山东省的调查证据[J].金融研究,2015(5):101-116.
- [2] 黄忠顺.民事执行机构改革实践之反思[J].现代法学,2017(2):3-18.
- [3] Karpoff J M, Lee D S, Martin G S. The Cost to Firms of Cooking the Books [J]. Journal of Financial and Quantitative Analysis, 2008, 43 (3): 581-611.
- [4] Verschoor W, Kemp R, Van den Broek S, et al. Reputational Penalties to Firms in Antitrust Investigations [J]. Journal of Competition Law & Economics, 2012, 8 (2): 231-258.
- [5] Armour J, Mayer C, Polo A. Regulatory Sanctions and Reputational Damage in Financial Markets [J]. Journal of Financial and Quantitative Analysis, 2017, 52 (4): 1429-1448.
- [6] Ang J S, Cheng Y, Wu C. Trust, Investment, and Business Contracting [J]. Journal of Financial and Quantitative Analysis, 2015, 50 (3): 569-595.
- [7] 雷宇.声誉机制的信任基础:危机与重建[J].管理评论,2016(8):225-237.
- [8] Gächter S, Fehr E. Altruistic Punishment in Humans [J]. Nature (London), 2002, 415 (6868): 137-140.
- [9] Sigmund K. Punish or Perish? Retaliation and Collaboration among Humans [J]. Trends in Ecology & Evolution (Amsterdam), 2007, 22 (11): 593-600.
- [10] Raihani N J, Grutter A S, Bshary R. Punishers Benefit from Third-Party Punishment in Fish: Fig. 1 [J]. Science (American Association for the Advancement of Science), 2010, 327 (5962): 171.
- [11] Gallo E, Yan C. The Effects of Reputational and Social Knowledge on Cooperation [J]. Proceedings of the National Academy of Sciences-PNAS, 2015, 112 (12): 3647-3652.
- [12] Giardini F, Vilone D. Evolution of Gossip-based Indirect Reciprocity on a Bipartite Network [J]. Scientific Reports, 2016, 6 (1): 37931.
- [13] Hess N H, Hagen E H. Psychological Adaptations for Assessing Gossip Veracity [J]. Human Nature (Hawthorne, N. Y.), 2006, 17 (3): 337-354.
- [14] Oliver P. Rewards and Punishments as Selective Incentives: An Apex Game [J]. The Journal of Conflict Resolution, 1984, 28 (1): 123-148.
- [15] Nowak M A, Fudenberg D, Rand D G, et al. Winners Don't Punish [J]. Nature (London), 2008, 452 (7185): 348-351.
- [16] Rankin D J, Dos Santos M, Wedekind C. The Evolutionary Significance of Costly Punishment Is Still to Be Demonstrated [J]. Proceedings of the National Academy of Sciences-PNAS, 2009, 106 (50): E135.
- [17] Wu J, Zhang B, Zhou Z, et al. Costly Punishment Does Not Always Increase Cooperation [J]. Proceedings of the National Academy of Sciences-PNAS, 2009, 106 (41): 17448-17451.
- [18] Barclay P. Reputational Benefits for Altruistic Punishment [J]. Evolution and Human Behavior, 2006, 27 (5): 325-344.
- [19] 刘涛,朱燕.《关于公布失信被执行人名单信息的若干规定》的理解与适用[J].人民司法,2013(19):25-30.
- [20] 余泳泽,郭梦华,胡山.社会失信环境与民营企业成长——来自城市失信人的经验证据[J].中国工业经济,2020(9):137-155.
- [21] 叶康涛,张然,徐浩萍.声誉、制度环境与债务融资——基于中国民营上市公司的证据[J].金融研究,2010(8):171-183.
- [22] 石晓峰,仲秋雁.媒体报道、所有权性质与上市公司债务融资期限结构[J].大连理工大学学报(社会科学版),2017(3):11-17.
- [23] 陈雪,孙慧莹,王雨鹏,等.媒体声誉与企业债务融资——基于媒体文本情绪大数据的证据[J].中央财经大学学报,2021(1):

54-69.

- [24] 赵玲, 黄昊. 数字化时代董事信息技术背景对内部控制有效性影响研究 [J]. 云南财经大学学报, 2022 (5): 80-101.
- [25] Pruckner G J, Sausgruber R. Honesty on the Streets: A Field Study on Newspaper Purchasing [J]. *Journal of the European Economic Association*, 2013, 11 (3SI): 661-679.
- [26] Doyle J, Ge W, Mcvay S. Determinants of Weaknesses in Internal Control over Financial Reporting [J]. *Journal of Accounting & Economics*, 2007, 44 (1): 193-223.
- [27] 杨雄胜. 内部控制理论研究新视野 [J]. 会计研究, 2005 (7): 49-54.
- [28] 陈汉文, 廖方楠, 韩洪灵. 独立董事联结与内部控制对盈余管理的治理效应 [J]. 经济管理, 2019 (5): 171-191.
- [29] 李卓松. 企业风险承担、高管金融背景与债券融资成本 [J]. 金融评论, 2018 (2): 73-84.
- [30] 于晓红, 卢相君. 基于行业环境条件下的上市公司创新战略与资本结构 [J]. 经济管理, 2012 (2): 50-56.
- [31] 程砚秋. 基于违约判别度的小企业信用风险评价研究 [J]. 科研管理, 2015 (S1): 510-517.
- [32] Claessens S, Laeven L. Financial Development, Property Rights, and Growth [J]. *The Journal of Finance (New York)*, 2003, 58 (6): 2401-2436.
- [33] 陈耿, 刘星, 辛清泉. 信贷歧视、金融发展与民营企业银行借款期限结构 [J]. 会计研究, 2015 (4): 40-46.
- [34] Diamond D W. Financial Intermediation and Delegated Monitoring [J]. *The Review of Economic Studies*, 1984, 51 (3): 393-414.
- [35] Demirgüç-Kunt A, Maksimovic V. Funding Growth in Bank-based and Market-based Financial Systems: Evidence from Firm-level Data [J]. *Journal of Financial Economics*, 2002, 65 (3): 337-363.
- [36] Vig V. Access to Collateral and Corporate Debt Structure: Evidence from a Natural Experiment; Access to Collateral and Corporate Debt Structure [J]. *The Journal of Finance (New York)*, 2013, 68 (3): 881-928.
- [37] Campello M, Larrain M. Enlarging the Contracting Space: Collateral Menus, Access to Credit, and Economic Activity [J]. *The Review of Financial Studies*, 2016, 29 (2): 349-383.
- [38] 尹志超, 甘犁. 信息不对称、企业异质性与信贷风险 [J]. 经济研究, 2011 (9): 121-132.
- [39] Gregory N, Tenev S. The Financing of Private Enterprise in China [J]. *Finance & Development*, 2001, 38 (1): 14.
- [40] Ayyagari M, Demirgüç-Kunt A, Maksimovic V. Formal Versus Informal Finance: Evidence from China [J]. *The Review of Financial Studies*, 2010, 23 (8): 3048-3097.
- [41] 徐思, 何晓怡, 钟凯. “一带一路”倡议与中国企业融资约束 [J]. 中国工业经济, 2019 (7): 155-173.
- [42] Bertrand M, Mullainathan S. Enjoying the Quiet Life? Corporate Governance and Managerial Preferences [J]. *The Journal of Political Economy*, 2003, 111 (5): 1043-1075.
- [43] Fang L H, Lerner J, Wu C. Intellectual Property Rights Protection, Ownership, and Innovation: Evidence from China [J]. *The Review of Financial Studies*, 2017, 30 (7): 2446-2477.
- [44] 钱雪松, 方胜. 担保物权制度改革影响了民营企业负债融资吗? ——来自中国《物权法》自然实验的经验证据 [J]. 经济研究, 2017 (5): 146-160.
- [45] Baron R M, Kenny D A. The Moderator-mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations [J]. *Journal of Personality and Social Psychology*, 1986, 51 (6): 1173-1182.
- [46] 甄红线, 王三法. 企业精准扶贫行为影响企业风险吗? [J]. 金融研究, 2021 (1): 131-149.
- [47] Roychowdhury S. Earnings Management through Real Activities Manipulation [J]. *Journal of Accounting & Economics*, 2006, 42 (3): 335-370.
- [48] 郭峰, 王靖一, 王芳, 等. 测度中国数字普惠金融发展: 指数编制与空间特征 [J]. 经济学 (季刊), 2020 (4): 1401-1418.
- [49] Biais B, Gollier C. Trade Credit and Credit Rationing [J]. *The Review of Financial Studies*, 1997, 10 (4): 903-937.
- [50] Petersen M A, Rajan R G. Trade Credit: Theories and Evidence [J]. *The Review of Financial Studies*, 1997, 10 (3): 661-691.
- [51] 石晓军, 李杰. 商业信用与银行借款的替代关系及其反周期性: 1998—2006年 [J]. 财经研究, 2009 (3): 4-15.
- [52] 赵胜民, 张博超. 商业信用与银行信贷能相互替代吗——基于2000—2018年中国上市公司数据的实证分析 [J]. 上海金融, 2019 (1): 16-23.

(责任编辑: 邵 霖 张安平)



# 土地要素空间配置对区域经济发展的影响： 基于建设用地指标跨省域调剂政策的分析

The Effect of Land Allocation on Regional Economic Development:  
An Analysis Based on the Inter-provincial Adjustment of  
Construction Land Indicators

张 莉 黄 伟

ZHANG Li HUANG Wei

**[摘要]** 现阶段,中国城市建设用地的空间供需矛盾严重制约了区域经济发展,土地资源的配置亟待优化。党的十九大以来,中央及各级政府出台了大量政策法规来缓解土地要素市场的扭曲,探索建设用地指标跨省域调剂就是其中的重要举措之一。基于2018年在东部八省和西部“三区三州”地区试点的城乡建设用地增减挂钩节余指标跨省域调剂政策,本文试图回答:建设用地指标的跨省域调剂是否对试点地区的经济发展产生了积极影响?实证结果显示,该政策促进了指标调出省份和指标调入省份的人均GDP增长,且对指标调出省份经济的提振作用更明显。这说明优化土地要素的空间配置可以同时促进发达地区和欠发达地区的经济发展,达到“双赢”的政策效果。此外,土地指标调出省份获得的土地调剂资金有效发挥了减贫效应,指标调出省份的农村贫困发生率显著降低。本文的结论为进一步推进建设全国统一的土地要素市场提供了经验支撑。

**[关键词]** 跨省域调剂 空间配置 减贫

**[中图分类号]** F127 F299.23 **[文献标识码]** A **[文章编号]** 1000-1549(2022)11-0115-14

**Abstract:** At this stage, the contradiction between the supply and demand of urban construction land has severely restricted the regional economic development in China, thus the allocation of land needs to be promoted urgently. Since the 19th National Congress of the Communist Party of China, both the central government and local governments have taken measures to correct the distortion of the land market, and conducting the inter-provincial adjustment of construction land indicators is one of these efforts. Based on the inter-provincial adjustment of surplus indicators of urban and rural construction land between eight eastern provinces and “San Qu San Zhou” in 2018, we aim to answer the following questions: does the inter-provincial adjustment of construction land indicators boost the economy of the pilot areas? Our empirical results show that the above policy has promoted the GDP per capita of both the eastern provinces and “San Qu San Zhou”, with a larger effect for the latter. These results imply that we can create a win-win situation by optimizing the spatial allocation of land elements. In addition, the indicator-adjustment funds obtained by the provinces from which the land indicators are transferred have played an important role in reducing the rural poverty incidence in these regions. The conclusions of our paper provide empirical evidence for the necessity of constructing a unified national land market.

**Key words:** Inter-provincial adjustment Spatial allocation Poverty reduction

**[收稿日期]** 2022-07-30

**[作者简介]** 张莉,女,1982年4月生,中山大学国际金融学院教授,博士生导师,延安大学乡村发展研究院兼职研究员,研究方向为城市与区域经济学;黄伟,男,1988年8月生,中山大学国际金融学院博士研究生,研究方向为城市与区域经济学。本文通讯作者为黄伟,联系方式为 huangwee-2006@163.com。

**[基金项目]** 国家自然科学基金面上项目(项目编号:71973158);广东省哲学社会科学创新工程2022年度特别委托项目;招商局慈善基金会和中国人口福利基金会资助项目。

感谢“畅通国内大循环与促进国内国际双循环的基础设施建设”国际学术研讨会各位与会者的建设性意见。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

资源配置是理解中国经济增长的关键视角。经过近四十年的改革探索，我国资本和劳动力的市场化程度已得到极大提高，要素配置效率明显改善。然而，相比于资本和劳动，土地要素的配置过程中仍存在诸多扭曲因素，严重制约了经济发展。政府垄断土地一级市场的现实背景以及土地要素的空间不可流动性决定了土地配置效率的提升需要更多的顶层设计。如果将资本和劳动配置效率的提升理解为市场在资源配置中起决定性作用，从而降低制度性成本的结果，那么土地要素市场的优化则更多依赖于在我国既定土地制度框架下用好政府“看得见的手”。具体而言，尽管现行的土地配置方式不能获得帕累托最优的结果，但通过优化政府规制，可以一定程度上缓解土地要素配置中的扭曲。探索建设用地指标的跨省域调剂即是其中的重要举措之一，本文以该政策为切入点分析土地要素空间配置对区域经济发展的影响，进而为建设全国统一要素市场提供政策参考。

一直以来，中国的土地出让市场由政府主导。中央政府在土地利用总体规划、年度土地利用计划以及基本农田保护这一约束体系下实施建设用地供给的计划指标管理，中央政府下达的建设用地指标再经各级地方政府层层分解，从而形成了中国建设用地的空间配置格局。由于存在信息不对称、激励扭曲、计划的时滞等，指标供给并不能保证土地要素得到有效配置。已有大量学者指出中国经济中存在土地资源错配 (Brandt 等, 2013<sup>[1]</sup>; 李力行等, 2016<sup>[2]</sup>; 陆铭, 2015<sup>[3]</sup>; 张莉等, 2022<sup>[4]</sup>)。特别是自 2003 年后，建设用地指标的分配表现出明显的地区偏向。具体来说，东部沿海省份具有生产率优势，但没有得到相应的建设用地指标数量；相反，中西部欠发达省份被配置了过多的建设用地。土地空间配置上的扭曲降低了全国层面的全要素生产率，进而阻碍了经济增长 (陆铭和向宽虎, 2014<sup>[5]</sup>)。在上述背景下，若能以现实中土地空间供需为依据实行跨区域甚至全国范围内的建设用地指标再配置，将极大缓解土地空间错配带来的效率损失。2022 年 4 月《中共中央国务院关于加快建设全国统一大市场的意见》强调了构建全国统一大市场对提升经济增长动能的决定性作用，其中建成统

一、规范的土地要素市场至关重要。这进一步表明，促进建设用地指标在全国范围内流转是顺应我国经济发展客观规律的必然要求。

事实上，在党的十九大以前，中央及各级政府已经在稳步推进相关政策，出台了一系列促进建设用地指标跨区域流转的举措，旨在优化土地空间配置，为经济高质量发展注入新动力。城乡建设用地增减挂钩结余指标流转就是其中的典范，2016 年以来集中连片的特困地区、国家级贫困县以及省级贫困县被允许将其增减挂钩节余指标在省域范围内流转使用，一定程度上提高了省域范围内的土地利用效率。然而，仅仅依靠省域内的建设用地指标流转并不能有效地解决我国土地和住房的空间供需错配问题，其原因是跨省域的人口流动仍然规模巨大。根据 2020 年第七次全国人口普查数据推算，我国流动人口总数约为 3.76 亿，其中跨省流动人口数约为 1.25 亿，占全部流动人口数的比例为 33%<sup>①</sup>。只有突破省域边界，推动全国范围内建设用地指标流转，才能真正实现“以人定地、地随人走”并最终提高土地要素的整体利用效率。因此，将增减挂钩结余指标由省域内流转扩围至跨省域流转，是实现我国高水平 and 高质量城市化的客观需要。2018 年 3 月，国务院办公厅印发《城乡建设用地增减挂钩节余指标跨省域调剂管理办法》，对深度贫困地区与对口帮扶省份间建设用地指标的跨省调剂作出了明确部署。通俗地讲，跨省域土地指标调剂，即将贫困地区多余的建设用地指标，分流给建设用地不足的东部省份，具体的操作过程可用图 1 来简述。其主要特征是中央政府统筹操作，内涵是通过土地指标在空间上的重新配置让指标调入和调出省份同时获得与其经济发展实际需要相匹配的建设用地数量，从而提高土地资源的整体利用效率。

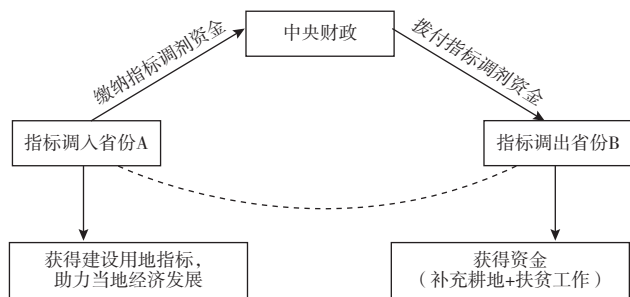


图 1 建设用地指标跨省域调剂示意图

① 流动人口数据由第七次全国人口普查资料中表 7-1 及相关假设推算得到，因缺乏微观层面的人口流动数据，该数据与准确结果相比可能存在一定偏差。

和大多数经济政策的推进过程一样,建设用地指标跨省域流转首先在部分地区进行了试点,意图通过评价其前期成效为政策的进一步推广提供经验支撑。本文主要考察 2018—2020 年时间段内在东部八省(广东、浙江、江苏、天津、山东、福建、上海、北京)和西部三区三州地区(西藏自治区、四川滇黔四省藏区、南疆四地州、四川凉山州、云南怒江州、甘肃临夏州)之间进行的建设用地指标跨省域调剂试点。理论上讲,将土地指标从三区三州地区调剂至东部发达省份,至少能产生两方面的积极作用:第一,建设用地指标的调入大大缓解了东部高生产率地区的土地资源约束,有利于其进一步发挥区位优势 and 集聚效应,扩大经济规模。第二,土地指标的调剂并不是无偿的,而是中央政府统筹管理下的指标交易。具体而言,三区三州地区在调剂出本省建设用地指标的同时,也获得了东部省份支付的调剂资金。考虑到该调剂资金的主要用途是帮扶深度贫困地区脱贫和农村基础设施建设,对于三区三州地区而言,跨省域建设用地指标调剂可能在降低其农村贫困发生率的同时促进当地经济增长。

本文将建设用地指标跨省域调剂试点作为外生的政策冲击,使用双重差分(DID)模型考察其对区域经济发展的影响。首先,我们利用地级市面板数据估计了该政策对东部八省和三区三州省份的经济增长效应。其次,基于省级层面的贫困发生率数据,我们也考察了该政策对三区三州省份的减贫效应。实证研究发现,跨省域的建设用地指标调剂使得三区三州省份的人均 GDP 提高了约 8.2%,而对于东部八省的人均 GDP 则有微弱的提振作用。由此可见,土地要素的空间再配置具有显著的经济增长效应,并且对于土地指标调入和调出省份具有“双赢”的政策效果。关于减贫效应的实证结果则显示,跨省域建设用地指标调剂使得三区三州地区的农村贫困发生率显著降低,这意味着该政策起到了较好的扶贫作用。

本文的主要贡献表现在两个方面:(1)土地资源的配置是区域经济学领域的重要话题,对于中国建设用地指标跨省域调剂政策的研究,已有文献大多局限于观点阐述和描述性工作,而相关的实证研究则较为缺乏。本文的政策评估提供了跨省域土地指标调剂能助力区域经济发展和减贫的经验证

据,对这些研究作出了补充。(2)本文的研究结论表明,政府可以通过合理有效的机制设计弥补土地要素市场化程度的不足,实现土地空间配置的优化。这为进一步深化跨省域建设用地指标流转提供了理论依据,具有较强的政策参考价值。(3)本文的研究与共同富裕有关,跨省域土地指标调剂可视为一种特定的转移支付,提供了一种缩小收入差距的新途径。

后文安排如下:第二部分介绍我国建设用地指标管理体制的演变历程以及跨省域建设用地指标调剂政策出台的背景;第三部分回顾和评述相关文献并提出本文的研究假说;第四部分是本文的研究设计以及核心变量的定义;第五部分呈现回归结果并对其进行简要分析;第六部分是结论与政策启示。

## 二、政策背景

本部分将系统梳理我国建设用地指标管理体系的发展演变过程,并具体介绍建设用地指标跨省域流转政策的相关背景。

### (一)我国城乡建设用地指标管理沿革

二十多年来,随着工业化和城市化发展,控制建设用地的规模和确保耕地红线成为我国建设用地指标管理的主基调,整体上经历了从项目管理到区域管理的转变。

项目管理,即由中央部门对特定项目所需占用额定土地的面积进行管理和审批,由于实践上的低效率,以及各地区、各项目的独特差异,从项目层面控制建设用地指标在后期变得难以为继。2008年,国务院明确了对各个地区新增建设用地指标实行“统一分配、层层分解、指令性管理”的方法。一方面,中央每年对各地新增建设项目用地指标进行总量管控;另一方面,严格农转用程序报中央审批。于是,以各级政府为主体,以总量控制为核心的建设用地指标管理体系初步形成。首先,根据《全国土地利用总体规划纲要(2006—2020年)》及调整方案,该管理体系确定了各省份在规划期内的建设用地总规模、耕地保有量及基本农田保护面积,各省再向下分解至市及区县;其次,各地方政府先根据土地利用总体规划,对一定时期内的建设用地总规模进行控制,之后每一年通过土地利用年度计划进行管理;最后,在每一年的实际工作中,县以上地方政府对本地建设

用地总规模进行动态管理。尽管目前有部分地区进行了用地审批权下放的试点<sup>①</sup>，但总的来说，建设用地指标的区域管理特征没有发生实质性变化。

在区域管理体制下，目前破解建设用地指标困境的主要手段包括增量与存量两大类。就增量（即常规的新增建设用地计划指标）而言，由于土地本身属于存量资源，新增建设用地一般意味着其他用地向建设用地的转换。而根据《中华人民共和国土地管理法（2014年修正）》，我国对耕地实行特殊保护；因此，若新增建设用地涉及农用地中的耕地，则需考虑占补平衡制度。在此背景下，盘活城乡存量建设用地成为解决城镇和工业园区建设用地不足的另一重要手段。通过逐步推进城乡建设用地增减挂钩、工矿废弃地复垦利用和城镇低效用地再开发可以有效破解建设用地指标限制。城乡建设用地增减挂钩，即在拆旧地块将农村建设用地进行拆除及复垦，增加的建设用地指标定向用于建新地块使用。例如，A地使用了一定面积建设性用地，具有对口关系的B地就垦复增加等面积耕地；这样一来，增减挂钩使得建设性用地的总增量是零，形成一个平衡。

2005年，为了缓解城市用地供需矛盾，原国土资源部按照《国务院关于深化改革土地管理的决定》有关要求，启动了增减挂钩的试点工作，并严格限定在县域范围内实施。2015年自然资源部（原国土资源部）又出台了《关于用好用活增减挂钩政策积极支持扶贫开发及异地扶贫搬迁工作的通知》，允许集中连片的特困地区、国家扶贫开发工作重点的县将增减挂钩结余指标在省域内流转使用。尽管这些尝试一定程度上缓解了用地供需矛盾，但由于我国建设用地需求和耕地后备资源数量的空间不均衡性更多地体现在东、中、西部这一较大地理尺度间，因此跨省域的建设用地增减挂钩节余指标调剂显得更为紧迫。

## （二）城乡建设用地增减挂钩结余指标跨省域流转

在经济集聚效应较强的东部地区，耕地保护与保障建设用地之间的矛盾导致重大项目落地难，经济发展受到制约。相反，土地资源富集的中西部地区，可补充的耕地后备资源充足，但经济上的欠发达使得土地价值难以实现。基于此，中央政府从2017年开始

逐步放开耕地占补平衡条件，并推行城乡建设用地增减挂钩节余指标跨省域调剂的试点，探索建立全国性建设用地指标跨区域交易的机制。2018年，增减挂钩政策进一步拓展，国务院办公厅印发《城乡建设用地增减挂钩节余指标跨省域调剂管理办法》（国办〔2018〕16号）以探索“三区三州”及深度贫困县增减挂钩结余指标在东西部扶贫协作和对口支援框架内开展交易。调出节余指标的地区限定在“三区三州”，即西藏自治区、四川滇甘四省藏区、南疆四地州和四川凉山州、云南怒江州、甘肃临夏州。根据经济承担能力，调入节余指标的主要帮扶省份确定为广东、浙江、江苏、天津、山东、福建、上海和北京东部八省市。至此，增减挂钩结余指标的跨省流转和交易得到了制度保障。

按照指导文件，城乡建设用地增减挂钩节余指标必须由国家统筹跨省域调剂使用。其中，自然资源部会同财政部、国家发改委和农业部等相关部门确定调剂规模、激励措施和监管要求。财政部会同自然资源部等相关部门制定资金使用管理办法，统一资金收取和支出。根据近年来中西部地区节余指标省域内流转的平均成交价格（每亩24万元），确定跨省域指标调出价格。具体标准为：复垦为一般耕地和其他农用地的，每亩为30万元；对部分有条件复垦为高标准农田的，适当提高价格标准，每亩为40万元，可以满足中西部地区拆旧复垦安置的需求且有结余。节余指标调入价格根据地区差异相应确定，北京、上海每亩70万元，天津、江苏、浙江、广东每亩50万元，福建、山东等其他省份每亩30万元；附加规划建设用地规模的，每亩再增加50万元。总而言之，本轮建设用地指标跨省域调剂试点的核心特征是中央政府统一下达调剂任务，统一实施调剂价格，统一资金收取和支出，进而实现东西部地区间增减挂钩数量和调剂资金的整体平衡。

由于东西部地区之间存在资金要素和土地要素上的互补优势，经过国家统筹调剂后，一方面，人均城市建设用地水平偏低、规划建设用地规模不足的东部省份，可以使用跨省域调剂节余指标少量增加规划建

<sup>①</sup> 2020年3月12号，国务院颁布了《关于授权和委托用地审批权的决定》，将永久基本农田以外的农用地转为建设用地审批事项授权省一级地方政府批准，同时试点将永久的基本农田转化为建设用地和国务院批准土地征收审批事项委托部分省级政府批准。首批进入试点的地区有北京、天津、上海、江苏、浙江、安徽、广东和重庆。

设用地规模,并在新一轮土地利用总体规划编制时予以调整。这一定程度上缓解了东部地区面临的土地资源约束,有助于其更好发挥区位优势和利用集聚外部性。另一方面,输出用地指标的西部欠发达省份所得调剂资金收益,除一部分给承担补充耕地任务的省份优先用于补充耕地外,其余全部用于脱贫攻坚和乡村振兴。因此,跨省域建设用地指标流转成为解决脱贫攻坚资金难题的重要渠道,为实现区域均衡发展和共同富裕目标创造了又一重要的制度条件。

### 三、文献回顾与理论假说

接下来,我们将对与本文联系较为紧密的相关文献进行回顾和评述,并在此基础上提出本文的研究假说。

建设用地指标的跨省域调剂以土地存量在空间上的非均匀分布特征为起点,通过土地要素的空间再配置实现全国层面的土地利用效率最大化。因此,本文的研究丰富了土地配置如何影响经济增长的相关文献。对比资本和劳动这两种要素,关于土地配置的研究相对较少,且研究对象主要聚焦于农业用地(Restuccia和Rogerson,2013<sup>[6]</sup>;Adamopoulos等,2014<sup>[7]</sup>)。有关建设用地配置的文献则主要关注城市间或城市内的土地错配(陆铭,2011<sup>[8]</sup>;陈晓光,2017<sup>[9]</sup>)。Duranton等(2015)<sup>[10]</sup>基于印度的研究认为,土地资源的错配给印度经济带来了效率上的损失,并且城市内土地错配和城市间土地错配具有不同的成因,因此纠正这些错配需要不同的政策建议。陆铭等(2015)<sup>[11]</sup>探究了建设用地指标分配对于不同城市居住成本的影响,认为差异化的土地供应政策导致受到土地供应限制的东部城市房价和工资上涨,进而对中国经济效率和竞争力产生了负面影响。在中国的城市内部,同样存在大量土地错配(黄忠华和杜雪君,2014<sup>[12]</sup>),体现在不同用途间、不同行业间、企业间等多个层面。李力行等(2016)<sup>[2]</sup>认为,中国各地大规模的开发区建设,导致大量稀缺的土地资源被低效利用,在降低企业进入门槛的同时,也产生了严重的生产效率损失。

就我国实际情况而言,土地资源禀赋和经济效率的不均衡更多地体现在不同省份之间,这就决定了土地供需错配也会更多地发生在省份之间。陆铭(2011)<sup>[14]</sup>指出,建设用地指标在2003年后更多配置

给中西部地区,导致了土地利用效率降低。邵挺等(2011)<sup>[13]</sup>则发现东部地区建设用地的边际生产率要远远高于中西部地区,各地级市间的土地利用效率差异主要来自省间差距,这意味着跨省区的土地错配严重。在方法上,Fang等(2021)<sup>[14]</sup>利用空间一般均衡模型量化了地区间建设用地指标错配对中国经济增长的影响,结论是该错配使得全国层面的生产率和总产出分别降低了7.3%和2.4%,这与赵扶扬和陈斌开(2021)<sup>[15]</sup>的研究结论类似。在中国的土地制度下,政府作为建设土地指标的供给者和管理者,主导土地资源配置(蒋省三等,2007<sup>[16]</sup>)。因此,如何利用好“政府之手”,通过合理有效的机制设计促进土地要素在空间上的再配置,是提升土地利用效率的关键途径。

土地作为一种稀缺资源,其达到帕累托有效配置的条件是土地要素的边际产出在各省份间相等。而目前的情形是高生产率的东部地区建设用地供给严重不足,土地边际产出较高;相反,生产率较低的中西部地区被配置了过多的建设用地,导致土地边际产出较低。这明显偏离了理论上的土地最优配置。跨省域建设用地指标调剂打破了土地要素的空间不可流动性,通过等量增加(减少)东部地区(西部地区)的建设用地指标数量来实现不同省份土地边际产出的趋同。无论对于指标调入还是调出地区,更好地供需匹配和更高的土地利用效率将促进其经济发展。基于以上分析,我们提出本文的第一个假说:

**假说1:**跨省域建设用地指标调剂可以同时促进指标调入和调出省份的经济增长。

除了带来可能的经济增长效应外,建设用地指标跨省域调剂的政策定位更多的是解决西部深度贫困地区的脱贫攻坚难题、促进区域均衡发展,是共同富裕目标在土地政策领域的具体表现。欠发达地区在流转出建设用地指标的同时也获得了用于脱贫的资金,事实上,我们可以将其理解为一种特殊形式的转移支付。关于转移支付是否能有效减贫,大部分学者持肯定观点。雷根强等(2015)<sup>[17]</sup>利用地理断点回归法发现转移支付有效促进了区域均衡发展。卢盛峰等(2018)<sup>[18]</sup>探讨了转移支付在缩小居民收入不平等中的效果,其结论是中国现行转移支付体系对市场收入不平等具有正向的调节作用。解垠和李敏(2020)<sup>[19]</sup>的研究发现政府公共转移支付有利于降低相对贫困,

且转移支付再分配效应大于社保缴费和个税的再分配效应。徐爱燕等(2021)<sup>[20]</sup>基于微观家庭数据实证研究了转移支付对西藏农牧民的减贫效应,发现转移支付能促进农牧民进行非农就业进而间接减贫。刘明慧和章润兰(2021)<sup>[21]</sup>认为财政转移支付对相对贫困的影响存在规模效应,地方财政民生性支出占比的提高有助于减缓相对贫困。李丹等(2019)<sup>[22]</sup>则指出转移支付可以有效提升相对贫困地区的“造血”能力,对极度贫困地区则更多地体现“输血”。上述研究的考察对象都是通常意义下的政府转移支付,而关于跨省域建设用地指标调剂资金,目前鲜有文献对其减贫效应进行评价。

土地不仅是农业生产的基础,还是城市建设发展的重要资源。正是由于土地的特殊重要性和空间不可流动性,现实中土地资源存量高的欠发达地区缺少发展资金,同时发达地区旺盛的土地需求得不到满足,即“地多的少钱、钱多的少地”。虽然东西部之间存在巨大的土地价差,但缺乏助力西部贫困农村发展的有效机制。允许建设用地指标跨省域调剂,正是从制度上为乡村振兴和农村脱贫攻坚筹集资金创造了条件。对建设用地相对充裕的贫困地区而言,这一举措可以让土地产生更大效益,换取急需的扶贫资金。理论上讲,把土地流转半径从省内扩大到跨省域,在更大空间范围对土地要素进行配置,这不仅直接拓宽了扶贫资金的来源渠道,而且会增加支持贫困农村地区的资金总量,对增加农村人均收入具有重要意义。此外,由于中央政府明确规定了调剂资金的大部分用于深度贫困地区的脱贫工作,因而相比于一般意义的财政转移支付,其资金的使用效率可能更高,扶贫效果也可能更为直接。综上所述,我们提出本文的第二个假说:

**假说 2:** 跨省域建设用地指标调剂有助于缓解指标调出地区的农村贫困。

#### 四、研究设计与变量定义

2018年3月,国务院办公厅印发《城乡建设用地增减挂钩节余指标跨省域调剂管理办法》(国办[2018]16号),正式对跨省域建设用地指标调剂展开试点。为了推进相关工作顺利实施,同年7月财政部制订、印发《城乡建设用地增减挂钩节余指标跨省域调剂资金收支管理办法》。本次试点期限为3年

(有效期截止到2020年年底),其中,调出节余指标的地区限定在“三区三州”,即西藏自治区、四川滇甘四省藏区、南疆四地州和四川凉山州、云南怒江州、甘肃临夏州,调入节余指标的主要帮扶省份为广东、浙江、江苏、天津、山东、福建、上海和北京东部八省市。这一政策背景为我们评估跨省域建设用地指标调剂的经济影响提供了一个合适的准自然实验。具体而言,我们可以将该政策视为一次外生冲击,进而使用双重差分法(DID)来分别估计政策的经济增长效应和扶贫效应。

##### (一) 计量模型

为了识别跨省域建设用地指标调剂对指标调入和调出地区经济增长的影响,我们选取未受试点政策影响的其他省份的地级市作为对照组,而分别将东部八省和三区三州所在省份的地级市作为实验组。具体的计量模型如下:

$$per\_gdp_{i,t} = \beta_0 + \beta_1 treat_i \times post_t + \beta_2 X_{i,t} + u_i + v_t + \varepsilon_{i,t} \quad (1)$$

其中,下标*i*和*t*分别代表地级市和年份。被解释变量 $per\_gdp_{i,t}$ 是地级市层面的人均GDP; $treat_i$ 是代表实验组别的虚拟变量,若某一地级市属于东部八省(或三区三州所在省份),取值为1,否则取0。 $post_t$ 是代表政策实施年份的虚拟变量,2018年及之后取值为1(由于2020年数据难以获得,本文的研究区间皆为2010—2019年),否则为0。交互项 $treat_i \times post_t$ 即是该双重差分模型的核心解释变量,其系数 $\beta_1$ 反映了跨省域建设用地指标调剂对指标调入(调出)地区的经济增长效应。 $X_{i,t}$ 是一系列地级市层面的控制变量, $u_i$ 和 $v_t$ 分别代表地级市固定效应和年份固定效应, $\varepsilon_{i,t}$ 是随机扰动项。各变量的具体定义和描述性统计将在后文介绍。

为了识别跨省域建设用地指标调剂对指标调出地区的扶贫效果,我们将未受试点政策影响的其他省份作为对照组,而将三区三州所在省份作为实验组。具体的计量模型如下:

$$poverty\_hap_{i,t} = \gamma_0 + \gamma_1 treat_i \times post_t + \gamma_2 X_{i,t} + u_i + v_t + \varepsilon_{i,t} \quad (2)$$

由于只能获得被解释变量的省级层面数据,因此方程(2)中的下标*i*代表省份。被解释变量 $poverty\_hap_{i,t}$ 是省级层面的贫困发生率。 $treat_i$ 是代表实验组别的虚拟变量,若某一省份位于三区三州所在省,取值为

1, 否则取 0。除了变量层级不同外, 其他变量的定义方式与方程 (1) 中类似。系数  $\gamma_1$  捕捉的是跨省建设用地指标调剂对指标调出地区的扶贫效应。各变量的具体定义和描述性统计见后文。

(二) 变量定义

1. 被解释变量。

计量模型 (1) 中, 被解释变量  $per\_gdp$  为地级市人均 GDP, 具体定义为地级市当年的国民生产总值与年末总人口的比值再取对数。数据来自《城市统计年鉴》, 时间区间为 2010—2019 年。计量模型 (2) 中, 被解释变量为某一省份当年的农村贫困发生率  $poverty\_hap$ 。具体定义分别是  $poverty\_hap = (\text{贫困线以下的农村人口数}) / (\text{农村总人口})$ , 该数据的来源为国家统计局农村住户调查和住户收支与生活状况调查, 时间区间为 2010—2019 年。

2. 核心解释变量。

计量模型 (1) 中, 核心解释变量  $treat_i \times post_i$  中  $treat_i$  取值为 1 的条件是某一地级市属于东部八省 (或三区三州所在省份)。计量模型 (2) 中,  $treat_i$  取值为 1 的条件是某一省份位于三区三州所在省。

$post_i$  的定义在这两组分析中相同且与前文保持一致。

3. 控制变量。

为了缓解可能存在的遗漏变量问题, 在考察试点政策的经济增长效应时, 我们在回归模型中加入了地级市层面的控制变量, 数据来自 2010—2019 年《城市统计年鉴》。具体来说, 参考郑世林等 (2014)<sup>[23]</sup>、李世刚和尹恒 (2017)<sup>[24]</sup>, 我们分别加入了以下变量: 城市规模  $city\_size$  (年末总人口取对数)、第二产业占比  $industry2$ 、第三产业占比  $industry3$ 、高等院校数量  $high\_edu$ 、取对数后的财政总收入  $gov\_income$  和年度财政支出  $gov\_exp$ 。在考察试点政策的扶贫效应时, 我们加入了省级层面的上述控制变量, 数据来自 2010—2019 年《中国统计年鉴》和各省统计局。此外, 在模型 (1) 和 (2) 的估计中, 我们分别控制了 (城市固定效应+年份固定效应) 以及 (省份固定效应+年份固定效应)。

为了排除样本极端值的影响, 本文对除虚拟变量外的所有变量进行了左右 1% 缩尾 (Winsor) 处理, 变量的描述性统计见表 1。

表 1 计量模型 (1) 对应的描述性统计

变量	实验组: 东部八省地级市		对照组: 不受政策影响的其他地级市		
	(1)	(2)	(3)	(4)	(5)
	<i>N</i>	平均值	标准差	最小值	最大值
<i>per_gdp</i>	2 374	10.69	0.56	9.19	12.29
<i>treat×post</i>	2 511	0.06	0.24	0	1
<i>city_size</i>	2 456	5.91	0.64	3.26	8.14
<i>industry2 (%)</i>	2 198	47.43	10.55	10.60	82.24
<i>industry3 (%)</i>	2 198	40.59	9.67	14.36	83.52
<i>high_edu</i>	2 402	9.23	15.36	1	93
<i>gov_income</i>	2 449	14.00	1.06	10.87	18.09
<i>gov_exp</i>	2 449	14.84	0.77	12.97	18.24
	实验组: 三区三州地级市		对照组: 不受政策影响的其他地级市		
<i>per_gdp</i>	2 057	10.42	0.54	8.58	11.46
<i>treat×post</i>	2 184	0.04	0.19	0	1
<i>city_size</i>	2 118	5.76	0.81	2.94	8.14
<i>industry2 (%)</i>	1 891	47.51	11.79	10.60	89.75
<i>industry3 (%)</i>	1 891	38.80	10.09	9.76	79.23
<i>high_edu</i>	2 042	7.79	13.31	1	84
<i>gov_income</i>	2 109	13.61	0.93	10.87	16.94
<i>gov_exp</i>	2 109	14.66	0.68	11.71	17.70

表 2 计量模型 (2) 对应的描述性统计

变量	(1)	(2)	(3)	(4)	(5)
	<i>N</i>	平均值	标准差	最小值	最大值
<i>poverty_hap</i>	221	0.11	0.09	0.01	0.49
<i>treat×post</i>	230	0.03	0.18	0	1
<i>city_size</i>	230	8.03	0.87	5.70	9.17
<i>industry2 (%)</i>	230	44.42	8.02	20.73	59.08
<i>industry3 (%)</i>	230	43.76	6.86	28.64	58.97
<i>gov_income</i>	230	6.76	0.87	3.23	7.97
<i>gov_exp</i>	230	8.10	0.61	6.31	9.25

## 五、实证结果

本文第三部分的分析认为,建设用地指标跨省域流转一方面能促进指标调出省份和调入省份的经济增长,另一方面也能缓解指标调出地区的农村贫困。本部分将为上述观点提供经验证据。

### (一) 建设用地指标跨省域流转的经济增长效应 1. 基准结果。

我们利用回归方程(1)~(2)列和(3)估计了建设用地指标跨省域流转对东部八省和三区三州的经济增长效应,结果分别如表3中列(1)、列(4)所示。

表 3 建设用地指标跨省域流转的经济增长效应

	(1)	(2)	(3)	(4)
	<i>per_gdp</i>	<i>per_gdp</i>	<i>per_gdp</i>	<i>per_gdp</i>
<i>treat×post</i>	0.050** (0.024)	0.010 (0.021)	0.095*** (0.026)	0.082*** (0.022)
<i>city_size</i>		-0.203** (0.086)		-0.103 (0.124)
<i>industry2</i>		0.037*** (0.003)		0.033*** (0.003)
<i>industry3</i>		0.026*** (0.004)		0.021*** (0.004)
<i>high_edu</i>		0.007*** (0.002)		0.008*** (0.003)
<i>gov_income</i>		0.165*** (0.029)		0.204*** (0.028)
<i>gov_exp</i>		0.144** (0.063)		0.143*** (0.049)
<i>Constant</i>	10.282*** (0.009)	4.371*** (0.836)	10.095*** (0.010)	3.605*** (0.864)
<i>CityFE</i>	Yes	Yes	Yes	Yes
<i>YearFE</i>	Yes	Yes	Yes	Yes
观测值	2 374	2 155	2 057	1 831
<i>within-R<sup>2</sup></i>	0.683	0.828	0.686	0.861

注:括号内为聚类到地级市层面的标准误,\*\*\*、\*\*和\*分别表示在1%、5%和10%的水平上显著。下同。

列(1)中政策变量的估计系数显著为正,但列(2)在加入城市层面的控制变量后交互项的系数大幅下降且统计上不显著。这意味着建设用地指标跨省域流转对东部八省经济增长的提振作用是有限的。结合现实来看,虽然理论上东部发达地区可以通过建设用地指标跨省域调剂缓解土地资源约束,但由于现阶段调剂规模和范围都偏小,因此该政策对东部地区经济增长的提升幅度不明显。这也间接说明,如果要通过土地要素的再配置给发达地区注入增长动力,需要在更大规模和范围内实现建设用地指标流转。列(3)的回归结果显示,在不加入控制变量时政策变量的系数显著为正且在1%的统计水平上显著,即建设用地指标跨省域流转促进了指标调出地区(三区三州)的经济增长。在控制了城市层面的特征变量后,核心解释变量的系数值变化幅度不大且仍在1%的水平上显著,如列(4)所示。以列(4)的结果为基准,平均而言该政策使三区三州地级市的人均GDP增长了约8.2%,这在经济意义上也是不小的幅度。综上所述,跨省域建设用地指标调剂能够通过优化土地资源的空间配置同时促进指标调出省份和指标调入省份的经济增长,即本文的假说1得到验证。此外,从估计系数的大小来看,该政策对指标调出省份经济增长的影响更显著;考虑到跨省域建设用地指标调剂的扶贫政策定位,这一结果与现实是相符的。

### 2. 稳健性分析。

#### (1) 平行趋势检验。

为了增强上述回归结果的可靠性,参照大多数使用双重差分法的实证研究,我们对回归方程(1)进行了平行趋势的检验。具体做法是先定义各个年份的虚拟变量,然后将其与实验组别虚拟变量 $treat_i$ 相乘得到一组政策变量交互项,用这些交互项替换模型



(1) 中的核心解释变量后重新估计即可得到政策的动态效应(将政策发生的前一年设定为基期)。估计结果如图2所示,2018年之前,政策变量的估计系数在统计上均不显著,平行趋势假设得到满足。2018

年之后,政策试点地区的人均GDP得到明显提升;与此同时,对比图2(左)和图2(右),不难发现跨省域建设用地指标调剂试点对指标调出地区的经济增长具有更强的促进作用。

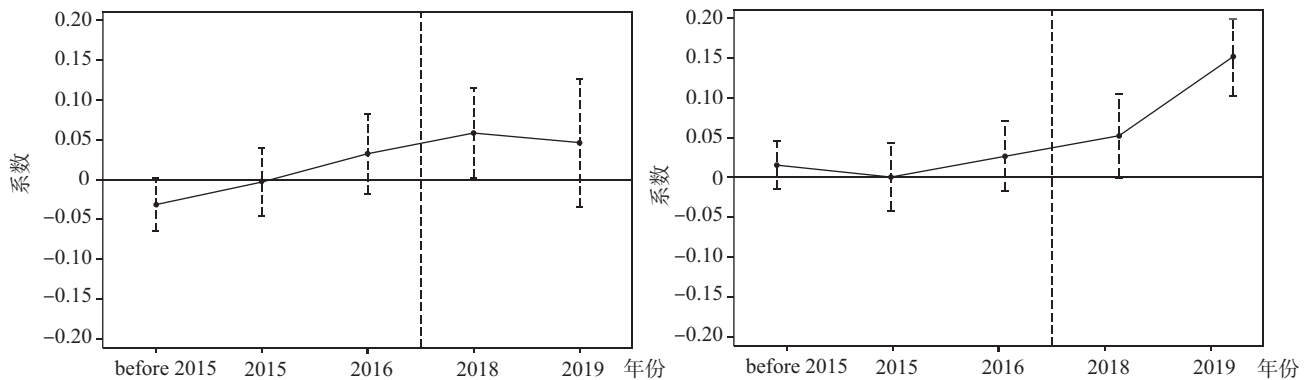


图2 建设用地指标跨省域流转的经济增长效应:平行趋势检验

注:左右两图分别对应东部八省和三区三州的经济增长动态效应。

(2) 核心变量的其他度量方法

基准回归中,我们将地级市的人均GDP作为衡量经济发展水平的指标。为了检验实证结果的稳健性,我们使用地级市单位建设用地面积的产值  $gdp\_land$  作为被解释变量重新估计模型(1)。人均GDP除了衡量经济发展水平外,一定程度上还反映了土地要素的利用效率,因此在本文的背景下是一个合理的被解释变量替换指标。此外,考虑到政策试点效果可能存在的时滞,我们将被解释变量提前一期( $t+1$ 期)后重新估计模型(1)。表4的回归结果显示,采用以上两种方式更换被解释变量后试点政策的经济增长效应仍然存在。

表4 建设用地指标跨省域流转的  
经济增长效应:替换被解释变量

	东部八省		三区三州	
	(1)	(2)	(3)	(4)
	$gdp\_land$	$per\_gdp(t+1)$	$gdp\_land$	$per\_gdp(t+1)$
$treat \times post$	0.045** (0.020)	0.014* (0.008)	0.095*** (0.022)	0.056** (0.026)
Controls	Yes	Yes	Yes	Yes
CityFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
观测值	2 155	1 873	1 831	1 549
$within-R^2$	0.883	0.830	0.869	0.820

(3) 控制地区时变特征

在前文的分析中,尽管我们已经控制了一系列城市层面的特征变量,但仍可能面临遗漏变量问题。具体来说,地区层面存在某些不可观测的特征,且这些特征是动态变化的,若不加以考虑可能会对我们的实证结果造成干扰。为了缓解上述担忧,我们参考 Vig (2013)<sup>[25]</sup>的做法,在计量模型(1)中加入省乘以年份的交互项来控制省级层面的时变特征,尽可能地排除遗漏变量问题,估计结果如表5所示。控制省份时变特征后,政策变量的系数仍显著异于零,这进一步验证了前文结论的稳健性。

表5 建设用地指标跨省域流转的  
经济增长效应:控制地区时变特征

	东部八省		三区三州	
	(1)	(2)	(3)	(4)
	$per\_gdp$	$per\_gdp$	$per\_gdp$	$per\_gdp$
$treat \times post$	0.048** (0.023)	0.012 (0.017)	0.092*** (0.021)	0.068*** (0.026)
Controls	No	Yes	No	Yes
Province×Year	Yes	Yes	Yes	Yes
CityFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
观测值	2 374	2 155	2 057	1 831
$within-R^2$	0.883	0.830	0.869	0.820

## (二) 建设用地指标跨省域流转的扶贫效应

### 1. 一些案例。

建设用地增减挂钩节余指标跨省域调剂政策是为支持脱贫攻坚而建立的新的制度机制。2018年中央1号文件明确“建立城乡建设用地增减挂钩节余指标跨省域调剂机制，将所得收益全部用于巩固脱贫攻坚成果和支持实施乡村振兴战略……”，中央政府旨在通过以城带乡、以工补农的方式支持深度贫困地区的脱贫工作，促进区域均衡发展。2019年国务院批准调剂资金由中央对地方一般性转移支付调整为专项转移支付，进一步保障了调剂资金全部用于脱贫攻坚。

政策试点以来，已有不少贫困地区通过调剂出富余的建设用地指标而获得了助力本地农村脱贫的资金要素。例如，2017年年底四川马边彝族自治县与浙江绍兴市越城区签订全国首个城乡建设用地增减挂钩节余指标流转协议，此次签约流转节余指标共计7000亩，总金额达50.4亿元，资金量相当于马边县当年财政总收入的12倍。2018年1月13日，浙江嘉善与四川木里藏族自治县签订框架协议，嘉善拟在3年内以每亩72万元价格向木里购买增减挂钩节余指标3000亩，且协议签订后嘉善将在1个月内向木里支付5000万元前期工作经费。截至2020年10月，建设用地增减挂钩节余指标跨省域调剂政策已为深度贫困地区筹集资金超过1800亿元<sup>①</sup>（图3对比了该调剂资金总量与各指标调出省份2017—2019年的财政收入总和），在国家财政转移支付之外为脱贫攻坚开辟了新的资金来源。

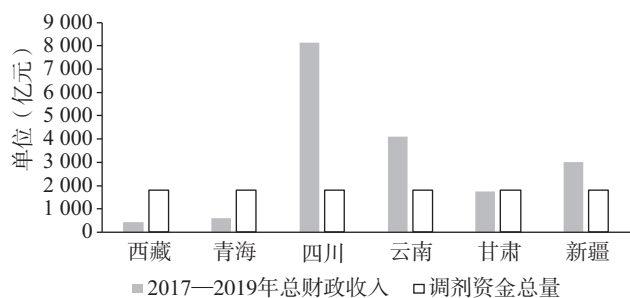


图3 截至2020年10月的建设用地增减挂钩节余指标跨省域调剂资金总量

注：各省份2017—2019年的财政收入总和来自省级财政年鉴。

### 2. 计量分析。

以上案例说明建设用地指标的跨省域调剂有效增加了指标调出地区的扶贫资金，从而解决了农村脱贫

攻坚面临的首要问题。接下来我们将实证检验该政策的扶贫效果，在呈现正式的回归结果之前，我们绘制了三区三州省份（实验组）和未受政策影响的其他中部省份（对照组）的平均农村贫困程度变化趋势，如图4所示。图4的纵轴变量是不同组别的平均农村贫困发生率，其定义如本文第四部分所述。首先，我们可以发现2010—2019年期间无论中部省份还是三区三州所在省份，平均农村贫困程度都逐年下降，这说明随着我国经济的发展和扶贫政策的深化，农村贫困问题得到了一定的缓解。其次，位于西部的三区三州省份，其农村贫困程度明显高于中部省份，这也是政府将一系列帮扶政策（包括本文讨论的建设用地指标的跨省域调剂）向西部地区倾斜的原因。

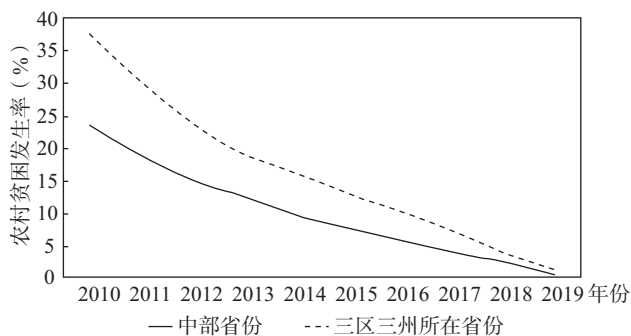


图4 不同地区农村贫困程度的变化趋势

建设用地指标的跨省域调剂是否以及多大程度缓解了指标调出省份的农村贫困问题？为了提供定量证据，我们利用计量模型（2）估计了试点政策对三区三州所在省份的扶贫效应，表6报告了回归结果。列（1）、列（2）中被解释变量为省份层面的农村贫困发生率，无论是否加入控制变量，政策交互项的系数都显著为负且在5%的统计水平上显著，这表明建设用地指标的跨省域调剂确实起到了扶贫作用。以列（2）的估计结果来计算，平均而言该政策使得指标调出省份的农村贫困发生率降低了5个百分点。基于以上估计结果，我们认为跨省域的建设用地指标调剂通过为贫困地区提供脱贫攻坚急需的“真金白银”，大幅降低了其农村贫困程度，也即本文的假说2得到验证。

### 3. 稳健性测试。

#### (1) 平行趋势检验。

采用与前文经济增长效应研究中类似的方法，我们对模型（2）对应的平行趋势假设进行了检验。图

<sup>①</sup> 该数据由2020国家扶贫日系列论坛扶贫用地政策分论坛公布。

5的结果显示,在政策实施之前的年份处理组省份并不存在事前趋势,且扶贫效应在政策实施后的第二年更加显著。

表6 建设用地指标跨省域流转的扶贫效应

	(1)	(2)
	<i>poverty_hap</i>	<i>poverty_hap</i>
<i>treat</i> × <i>post</i>	-0.070*** (0.023)	-0.050** (0.019)
<i>city_size</i>		-0.772* (0.383)
<i>industry2</i>		0.563* (0.324)
<i>industry3</i>		0.707* (0.363)
<i>gov_income</i>		-0.110* (0.064)
<i>gov_exp</i>		-0.000 (0.100)
<i>Constant</i>	0.212*** (0.012)	6.607** (2.926)
<i>ProvinceFE</i>	Yes	Yes
<i>YearFE</i>	Yes	Yes
观测值	208	208
<i>within-R</i> <sup>2</sup>	0.800	0.862

(2) 排除其他扶贫政策的干扰。

作为我国扶贫工作的重点对象和脱贫攻坚的“硬骨头”,三区三州所在省份往往受到来自中央层面的不同政策(包括直接转移支付和教育扶贫项目等)帮扶。如果在本文所考察的政策试点时间窗口附近同时存在其他的扶贫政策,那么我们得到的处理效应将是不准确的。由于无法精确地控制某一项具体扶贫政策的影响,为了一定程度上排除这些政策对本文结论的干扰,我们进行了以下的安慰剂检验。首先,我们保持处理组和控制组的组别定义方式不变,但假定政策冲击分别发生在*t*-2期、*t*-3期和*t*-4期,如果此时政策变量的估计系数不再显著,则说明并不是某项事前的扶贫政策导致了前文的估计结果。其次,我们从样本中随机选择处理组和控制组来估计本文的双重差分模型(2),图6描绘了

经过500次随机抽样的估计系数值分布。结合表7和图6的结果可以看出,前文估计得到的扶贫效应并不是由某项其他扶贫政策引起,从而增强了本文结论的稳健性。

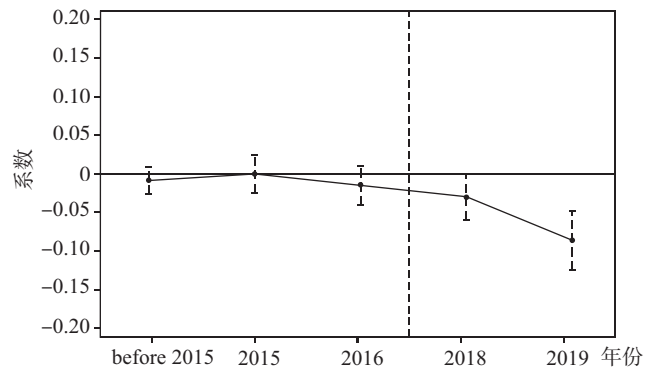


图5 建设用地指标跨省域流转的扶贫效应:平行趋势检验

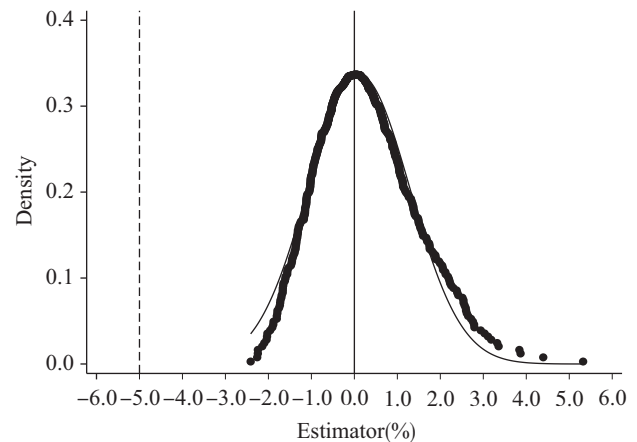


图6 安慰剂检验

表7 建设用地指标跨省域流转的扶贫效应:排除其他政策的干扰

	<i>t</i> -2	<i>t</i> -3	<i>t</i> -4
	(1)	(2)	(3)
	<i>poverty_hap</i>	<i>poverty_hap</i>	<i>poverty_hap</i>
<i>treat</i> × <i>post</i>	0.018 (0.033)	0.012 (0.027)	0.022 (0.021)
<i>Controls</i>	Yes	Yes	Yes
<i>ProvinceFE</i>	Yes	Yes	Yes
<i>YearFE</i>	Yes	Yes	Yes
观测值	208	208	208
<i>within-R</i> <sup>2</sup>	0.813	0.830	0.839

(三) 进一步分析: 跨省域建设用地指标调剂能否抑制东部发达地区房价上涨?

改革开放以来, 我国的工业化和城市化进程表现出明显的区域不平衡特征。相比于中西部地区, 东部沿海城市依靠先天区位优势和部分政策红利吸引了大量生产要素向该地流动, 其中以劳动力从中西部向东部集聚的过程最为突出。伴随该过程的一个现实是东部地区(特别是一线城市)的房价增速也远高于其他地区(韩立彬和陆铭, 2018<sup>[26]</sup>)。理论上, 在劳动力流入引致住房需求上升后, 若同步增加东部城市的土地供给数量, 一定程度上可以起到降或稳房价的作用。但现实中东部城市的土地供给并不与其高涨的需求相匹配, 已有学者指出中国东部城市房价高速上涨的根本原因是其土地供给更加缺乏弹性(刘修岩等, 2019<sup>[27]</sup>; 刘学良, 2014<sup>[28]</sup>)。

在本次试点政策框架下, 东部经济发达城市可以从西部欠发达地区调入建设用地指标从而放松本地土地资源约束, 从供给端减缓房价上涨压力。因此一个

很自然的问题是: 建设用地指标的跨省域调剂能否抑制东部城市房价上涨? 为了回答这一问题, 我们将地级市的年度平均房价  $h\_price$  作为被解释变量, 利用本文第四部分的计量方程(1)估计了该政策对东部地区房价的影响。其中实验组是东部八省份的地级市, 对照组是除三区三州地区以外其他省份的地级市, 回归结果如表8所示。列(1)、列(2)中的样本包括所有实验组和对照组的地级市, 列(3)、列(4)剔除了“北上广深”这四个一线城市, 列(5)、列(6)则同时剔除了“北上广深”和东部八省的省会城市。表8的实证结果显示, 尽管政策变量的系数估计值为负, 但在统计意义上均不显著, 这表明该政策并不能有效地降低房价。一个可能的原因是目前跨省域建设用地的调剂数量相比于各地方政府每年的土地出让总量仍然偏少, 因而对房价的抑制作用有限。基于这一解释, 要使得该政策除了扶贫外还能有效降低东部城市房价, 接下来的政策方向是进一步扩大跨省域建设用地指标调剂的范围和数量。

表8 建设用地指标跨省域流转的房价抑制效应?

	(1)	(2)	(3)	(4)	(5)	(6)
	$h\_price$	$h\_price$	$h\_price$	$h\_price$	$h\_price$	$h\_price$
$treat \times post$	-0.085 (0.097)	-0.050 (0.091)	-0.099 (0.102)	-0.059 (0.095)	-0.113 (0.109)	-0.074 (0.103)
$city\_size$		-0.100 (0.329)		-0.210 (0.338)		-0.230 (0.336)
$industry2$		0.010 (0.009)		0.011 (0.009)		0.012 (0.009)
$industry3$		0.020 (0.012)		0.021* (0.012)		0.021* (0.012)
$high\_edu$		0.003 (0.005)		0.002 (0.005)		0.007 (0.006)
$gov\_income$		0.017 (0.042)		0.007 (0.042)		0.009 (0.043)
$gov\_exp$		0.139*** (0.033)		0.136*** (0.033)		0.128*** (0.032)
$Constant$	8.133*** (0.010)	5.386*** (1.309)	8.110*** (0.010)	6.132*** (1.249)	8.090*** (0.010)	6.231*** (1.253)
$CityFE$	Yes	Yes	Yes	Yes	Yes	Yes

续前表

	(1)	(2)	(3)	(4)	(5)	(6)
	<i>h_price</i>	<i>h_price</i>	<i>h_price</i>	<i>h_price</i>	<i>h_price</i>	<i>h_price</i>
<i>YearFE</i>	Yes	Yes	Yes	Yes	Yes	Yes
观测值	2 374	2 155	2 334	2 115	2 284	2 065
<i>within-R</i> <sup>2</sup>	0.260	0.250	0.256	0.246	0.253	0.243

资料来源：地级市房价数据来自国信房地产网。

到目前为止，我们的实证分析表明，对贫困地区而言，建设用地指标的跨省域调剂既可以加快其脱贫进程又能推动其经济增长，即同时具备“输血”和“造血”功能。对东部发达省份而言，由于放松了土地资源约束，该政策对其经济发展也具备一定的提振作用。此外，我们也注意到该政策对东部地区的房价并不能起到明显的抑制作用。

### 六、结论与政策启示

提升资源配置效率是实现我国经济高质量发展的重要突破口。当前，我国土地要素利用效率仍然不高，地区间的错配为主要表现之一。土地作为一种空间上不可流动的要素，其配置效率的提升很大程度上有赖于政府的政策设计。建设用地指标跨区域流转是一项制度创新，突破了土地要素空间不可流动性，提高其配置效率。政策的核心目标一方面是帮扶调出建设用地指标的西部省份，助力其实现脱贫。另一方面是为东部发达地区解决计划指标紧、项目落地难的问题，进而为其提供更大的发展空间。本文以2018年在三区三州和东部八省之间试点的城乡建设用地增减挂钩结余指标跨省域调剂为自然实验，利用双重差分法评估了该政策的经济增长效应和扶贫效应。实证结果显示，该政策同时促进了指标调出和指标调入省份的人均GDP增长，且对指标调出省份的经济发展具有更强的推动作用。与此同时，该政策有效减缓了指标调出地区的农村贫困程度，扶贫效果明显。本文的结论能为我们提供以下政策启示。

第一，在我国要素配置市场化改革的进程中，政策制定者要思考和理顺政府与市场的关系。具体地，政府与市场的边界并不是一成不变的，二者也并不对立，而是相辅相成。价格机制可以充分发挥调节作用的领域，应该由市场来配置资源。当价格机制无法引导要素流动时，则需要政府适度干预以弥补市场功能

的缺失。本文考察的土地指标跨省域调剂就是政府适度干预进而优化资源配置的典型案列。因此，提升我国生产要素的配置效率，需要同时用好市场主导和政府协调这“两只手”。

第二，尽管目前跨省域的建设用地指标调剂数量有限，但已经显示出丰厚的政策红利。就经济整体而言，该政策优化了区域、城乡的用地结构和布局，实现了土地要素和资金要素的双向流动和优势互补，促进了区域协调发展，可谓“一举多得”。这意味着，实现更大范围和更大规模的建设用地流转，可以充分释放要素配置效率改善的潜能。考虑到我国特殊的土地制度和土地利用现状，提升土地要素的配置效率将是下一阶段提振中国经济的有效途径之一。

第三，当前的建设用地指标跨省域调剂仍然存在着不可忽视的问题，例如，指标交易数量受到严格限制从而不能反映市场供求，指标流入地区占用的耕地质量与流出地区复垦的耕地质量无法准确对应，增减挂钩节余指标的收益分配不合理，等等。政府如何在政策制定过程中充分考虑各方主体的利益诉求？这需要更深入的理论研究和更完善的机制设计。

要补齐当前土地指标跨区域流转制度的短板，总体的改革方向是进一步提高土地指标跨区域流转的市场化程度。具体到政策措施上，可以从以下三个方面发力。首先，中央政府要转变指标分配者角色，避免对土地指标流转价格和数量的过度干预。在政府制定价格范围和数量范围之后，由交易双方自主商议交易价格和交易数量，自行完成节余指标跨省交易。其次，政府要充当“监管者”的角色，在指标分配、流转、审批过程中做到监管到位，做好建新和复垦工作的跟踪监管。与此同时，扩大跨省土地流转的交易范围，逐步从“三州三区”和深度贫困县向其他地市推行跨省交易，乃至统筹做好农村集体经营性建设

用地入市与增减挂钩节余指标跨省流转工作。最后,搭建全国性的节余指标交易平台,下放市场交易主体。一方面,要发展和完善用地指标跨地区交易的市场机制,实现用地指标在区域间有效的二次分配。另

一方面,适当下放交易主体,增强交易自主性。市级人民政府在获得指标复垦成本、农用地质量、指标紧张程度等信息方面具有优势,方便控制复垦规模和质量,可以成为跨省流转的主体。

## 参考文献

- [1] Brandt L, Trevor T, Zhu X D. Factor Market Distortions Across Time, Space and Sectors in China [J]. *Review of Economic Dynamics*, 2013, 16: 39-58.
- [2] 李力行, 黄佩媛, 马光荣. 土地资源错配与中国工业企业生产率差异 [J]. *管理世界*, 2016 (8): 86-96.
- [3] 陆铭. 中国土地政策如何影响了经济竞争力 [N]. *东方早报*, 2015-01-13.
- [4] 张莉, 陆铭, 刘雅丽. 税收激励与城市商住用地结构——来自“营改增”的经验证据 [J]. *经济学 (季刊)*, 2022 (4): 1425-1446.
- [5] 陆铭, 向宽虎. 破解效率与平衡的冲突——论中国的区域发展战略 [J]. *经济社会体制比较*, 2014 (4): 1-16.
- [6] Restuccia D, Rogerson R. Misallocation and Productivity [J]. *Review of Economic Dynamics*, 2013, 16: 1-10.
- [7] Adamopoulos T, Brandt L, Leight J, Restuccia D. Misallocation, Selection and Productivity: A Quantitative Analysis with Panel Data from China [J]. NBER Working Paper, No. 23039, 2014.
- [8] 陆铭. 建设用地使用权跨区域再配置: 中国经济增长的新动力 [J]. *世界经济*, 2011 (1): 107-125.
- [9] 陈晓光. 土地错配、抵押效应与生产效率损失——基于企业数据的测算 [C]. 第二届空间政治经济学研讨会. 上海: 上海交通大学中国发展研究院, 2017.
- [10] Duranton G, Ejaz G, Arti G, William K. The Misallocation of Land and Other Factors of Production in India [J]. *World Bank Policy Research Working Paper*, No. 7221, 2015.
- [11] 陆铭, 张航, 梁文泉. 偏向中西部的土地供应如何推升了东部的工资 [J]. *中国社会科学*, 2015 (5): 59-83.
- [12] 黄忠华, 杜雪君. 土地资源错配研究综述 [J]. *中国土地科学*, 2014 (8): 80-87.
- [13] 邵挺, 崔凡, 范英, 许庆. 土地利用效率、省际差异与异地占补平衡 [J]. *经济学 (季刊)*, 2011 (3): 1087-1104.
- [14] Fang M, Han L B, Huang Z B, Lu M, Zhang L. Regional Convergence or Just An Illusion? Place-based Land Policy and Spatial Misallocation [J]. SSRN Working Paper, No. 3846313, 2021.
- [15] 赵扶扬, 陈斌开. 土地的区域间配置与新发展格局——基于量化空间均衡的研究 [J]. *中国工业经济*, 2021 (8): 94-113.
- [16] 蒋省三, 刘守英, 李青. 土地制度改革与国民经济成长 [J]. *管理世界*, 2007 (9): 1-9.
- [17] 雷根强, 黄晓虹, 席鹏辉. 转移支付对城乡收入差距的影响——基于我国中西部县域数据的模糊断点回归分析 [J]. *财贸经济*, 2015 (12): 35-48.
- [18] 卢盛峰, 陈思霞, 时良彦. 走向收入平衡增长: 中国转移支付系统“精准扶贫”了吗? [J]. *经济研究*, 2018 (11): 49-64.
- [19] 解垚, 李敏. 相对贫困、再分配与财政获益: 税收和转移支付的作用如何? [J]. *上海财经大学学报*, 2020 (6): 3-20.
- [20] 徐爱燕, 何松泽, 辛馨. 转移支付、非农就业与减贫——以西藏自治区为例 [J]. *社会科学家*, 2021 (3): 82-89.
- [21] 刘明慧, 章润兰. 财政转移支付、地方财政收支决策与相对贫困 [J]. *财政研究*, 2021 (4): 34-49.
- [22] 李丹, 裴育, 陈欢. 财政转移支付是“输血”还是“造血”——基于国定扶贫县的实证研究 [J]. *财贸经济*, 2019 (6): 22-39.
- [23] 郑世林, 周黎安, 何维达. 电信基础设施与中国经济增长 [J]. *经济研究*, 2014 (5): 77-90.
- [24] 李世刚, 尹恒. 政府-企业间人才配置与经济增长——基于中国地级市数据的经验研究 [J]. *经济研究*, 2017 (4): 78-91.
- [25] Vig V. Access to Collateral and Corporate Debt Structure: Evidence from a Natural Experiment [J]. *Journal of Finance*, 2013, 68 (3): 881-928.
- [26] 韩立彬, 陆铭. 供需错配: 解开中国房价分化之谜 [J]. *世界经济*, 2018 (10): 126-149.
- [27] 刘修岩, 杜聪, 李松林. 自然地理约束、土地利用规制与中国住房供给弹性 [J]. *经济研究*, 2019 (4): 99-115.
- [28] 刘学良. 中国城市的住房供给弹性、影响因素和房价表现 [J]. *财贸经济*, 2014 (4): 125-137.

(责任编辑: 李 晟 张安平)