

# 产业政策关联度对公司业绩影响研究<sup>①</sup>

覃 飞<sup>1</sup> 沈 艳<sup>2</sup>

(1. 北京工商大学数学与统计学院; 2. 北京大学国家发展研究院)

**研究目标:** 测算政府工作报告和企业年报的产业政策关联程度, 并评估政策关联对企业业绩的影响以及内在机制。**研究方法:** 基于 2009~2018 年的国务院政府工作报告和企业年报文本数据, 利用文本大数据分析技术测算企业年报和政府工作报告文本的政策相似度, 并使用面板回归模型评估政策效果以及机制分析。研究发现: 第一, 2012 年以后企业经营活动与政策关联度显著上升; 第二, 政策关联度较高的企业往往为民营、中小型、东部地区、轻资产、财务杠杆率低、资金状况良好, 并更敢于冒风险的企业; 第三, 高政策关联度的企业业绩水平高, 这一点对民营企业、中小企业和东部地区企业更为明显; 第四, 就影响机制而言, 本文发现政策关联程度高的企业融资约束更低, 也获得了更多的政府补贴和税收优惠; 第五, 由于在扣除政府补贴和税收优惠后, 政策关联度和企业业绩水平间仍存在正相关关系, 这一点对民营企业尤为明显, 表明企业业绩好并非仅仅通过获取政策相关优惠而采取策略性创新导致。**研究创新:** 本文创造性地结合文本大数据分析技术, 采用相似度算法测算年报和政府工作报告的政策关联程度, 为量化政策提供了新的视角。**研究价值:** 本文的方法为利用文本大数据分析技术研究经济学问题提供了借鉴, 研究结论对政府优化政策实施手段, 提高政策执行效率提供了建议。

**关键词** 政策关联 企业绩效 文本大数据分析 自然语言处理

**中图分类号** F420 **文献标识码** A

DOI:10.13653/j.cnki.jqte.2021.09.007

## 引 言

改革开放以来, 我国经济持续高速增长, 政府政策作为优化结构、推动技术升级等的重要手段被广泛运用在国民经济的各个领域, 成为企业发展不可忽视的外部力量。较多的观点认为政府政策可以提升企业的创新能力和创新效率 (Peters 等, 2012), 帮助企业提高生产率 (林毅夫等, 2018), 提高企业的全要素生产率和资源配置效率 (宋凌云和王贤彬, 2013), 降低企业信息收集成本 (谭劲松等, 2017), 有助于优化产业结构 (韩永辉等, 2017), 是政府利用有限的资源有效地帮助微观企业克服外部性和协调问题不可或缺的条件 (林毅夫, 2017)。但也有研究认为政府政策并不能够推动行业创新, 保护弱势企业 (Krugman, 1983); 只会激励企业策略性创新, 并未提高质量上的创新 (黎文靖和郑曼妮, 2016); 对单个企业的产出绩效作用不大 (陆国庆等, 2014); 降低了资源配置效率和企业的投资效率 (钱雪松等, 2018)。可见, 政府政策对微观企业发展的影响在学术界的观点并不统一。

<sup>①</sup> 本文获得北京工商大学青年教师科研启动基金项目 (QNJJ2021-03) 的资助。沈艳为本文通讯作者。

正如蔡庆丰和田霖(2019)指出的,当前学术研究的一个重点与难点是从经济政策出发落脚到微观企业行为,而将经济政策内化到企业微观决策行为中进行研究更有价值。那么,政府经济政策对微观企业会产生何种影响?微观企业会因政府政策作出何种调整?政府的政策执行往往伴随着财政补贴、税收优惠等,以及间接的融资支持等,有助于改善企业投资水平和短期业绩水平(王克敏等,2015),企业存在获取资源而进行政策套利的动机,那么考虑政策套利后政府政策对微观企业影响的有效性是否依然存在?这些问题都亟待进一步研究。近年来,为了推动中国经济往高质量经济增长方向转变、应对中美贸易摩擦和经济不确定性增加,积极的政策仍然是我国调控经济活动的重要工具,因此亟需对新时代政策有效性作进一步评估。

要通过实证研究来匡算政策影响程度并发掘影响的具体机制存在的一个主要挑战,就是政策实施有效性不易量化。现有文献对政策采取了多种度量方法,例如Aghion等(2015)使用政府税收优惠、贷款、关税等指标测度政策;宋凌云和王贤彬(2013)则手工收集了各个省份的“九五”“十五”“十一五”规划的重点政策,通过统计数量等方法构造重点产业指标;韩永辉等(2017)通过统计各省份各个年份政策的地方性法规和地方政府规章及其各自的文件累积数测度政策的力度;杨德明和刘泳文(2018)通过人工打分量化“互联网+”政策实施程度。更一般地,不少研究采用虚拟变量来综合检验政策对企业的影响(余明桂等,2016);或者按照特定政策实施前后分样本来估计(韩乾和洪永淼,2014)。

值得注意的是,政策的有效性不仅取决于政府对特定产业的支持力度,还取决于企业是否会因为政府相关政策而调整。文献中讨论了政策作用于微观企业行为的三种途径。一是改善企业生产的外部环境,降低企业创新投资的风险。采取特定产业政策的政府部门通常会加大对基础设施建设的投入,建立产业区、创业园、自贸区等,为企业的发展提供良好的外部条件;各级政府还会通过特定的专利保护政策、公共研发基础平台、人才支持计划等,来降低企业创新投资的风险(韩永辉等,2017)。二是改善企业的信息不对称问题和降低交易成本。一般情况下,政府政策都是政府职能部门投入大量的人力物力收集整理资料并反复研判的产物,代表着产业的未来发展方向,也是对市场发出的可信信号(韩永辉等,2017)。如果企业可以充分利用政府政策,就可能降低企业投资活动的事前和事后不确定性,进而推动企业形成长远的竞争优势(谭劲松等,2017)。三是提供融资支持、税收优惠、政府补贴等(张杰等,2016),能够间接和直接改善企业的现金流状况。同时,受政策支持的行业一般具有较好的发展前景,为了能够获取政府政策的支持,企业存在根据政策调整自身业务的动机,对于优化企业业务具有很强的指引效应。

上述三个途径起作用都有一个前提,即微观企业的经营活动对政府政策作出相应调整。因此,评估政府政策有效性的一个重要角度是量化企业经营行为对政府政策的敏感程度。实证研究虽有评估企业经营行为对政府政策作出调整的证据,如余明桂等(2016)发现受政策支持的企业在专利申请数量、科技人力资源占比和企业研发投入等方面均高于其他企业,但囿于数据可得性,目前尚缺乏评估企业经营行为对政府政策的敏感程度的指标。

近年来,文本大数据分析在经济学和金融学中得到了广泛的运用(沈艳等,2019),例如Hoberg和Phillips(2016)相似度算法计算不同上市公司年报披露的产品描述内容的相似度,构造新的行业分类方法(有别于传统固定不变的行业分类);Hassan等(2019)以上市公司年报披露后的电话会议文本记录为数据,通过词频等算法构造企业层面的政治风险;伊志宏等(2019)借助文本分类方法度量上市公司的分析师研究报告中公司特质信息,研究分

析师报告内容对股价同步性的影响。这些进展为从企业经营活动对政策敏感度角度度量政府政策提供了新视角。通过将反映政府政策和企业经营活动的文本这类非结构化数据转化为结构化数据,并进一步度量这些结构化数据之间的关系,就有可能从微观层面刻画不同时间、不同行业的企业对于政府统一政策的反应程度,进而可以进一步考察企业经营活动对政策的反应程度与其绩效之间的关系。

本文旨在从量化企业经营活动对政府政策反应程度出发,评估这一关联度对企业绩效的影响及其内在作用机制,进而考察政府政策的有效性。要量化企业经营行为对政府政策的敏感程度,首先需要选择能从较长期反映政府政策的文件。国务院政府工作报告是中央政府每年召开全国人民代表大会时发布的公文,主要包括过去一年工作回顾、当年工作任务(经济、社会、劳动等)、政府自身建设和其他内容,集中记录了当年政府鼓励并推动的政策信息。例如,2014年3月发布的政府工作报告中首次提及“互联网金融”,之后四年的政府工作报告中均有提及“互联网金融”,体现了政府对“互联网金融”的重视。虽然在政府办公的数字化下,越来越多的政策性文件以电子形式发布,也应是政策分析的语料来源,但各个地方政府和部门发布的政策性文件较分散、重复度高、充斥着较多无关信息;而与之相反,政府工作报告则集中地体现了中央政府的政策态度,因为一旦发布将受到国内外的广泛关注,因此语言一致、连贯、简洁、信息含量大,是政策类文本数据的良好来源。鉴于此,在综合考虑政策文本来源权威性、文本质量、获取难易程度等因素后,本文以国务院政府工作报告为政策类文本数据来源,以提取政府对过去一年以及未来一段时期内各个产业的政策态度信息。

要度量企业对于政府政策的敏感程度,也需要选择能恰当反映企业经营行为对政策反映程度的文本。为了考察较长时期并且让不同地区不同行业企业的行为具有可比性,我们选择上市公司年报作为企业经营行为与政策关联度的主要语料。上市公司年报是上市公司信息披露的重要组成部分,根据《公开发行证券的公司信息披露内容与格式准则第2号——年度报告的内容与格式(2017年修订)》,证监会要求上市公司年报中应当提供报告期内公司从事的主要业务、主要产品及其用途、经营模式、主要的业绩驱动因素等内容,并且应当重点突出其报告期内发生的重大变化等。相比于企业为了宣传自身产品和服务而在诸多媒介(诸如企业网站)上提供的业务范围描述,为满足监管要求,上市公司年报总体上对企业主要业务、产品和服务的介绍更为详细和准确,语言更规范,行文更严谨,格式也更专业化,因此是较为理想的能在全中国范围内反映较长时期企业经营行为对政策反应程度的语料。

本文的分析策略如下。首先,我们采用国务院政府工作报告作为度量政府政策力度的文本数据来源,又采用上市公司年报作为度量其经营活动与政府政策关联度的数据来源,结合文本大数据分析技术,构建和描述企业经营活动与政策之间的关联程度,即上市公司政策关联度。如果上市公司政策关联度越高,可以认为企业更关注政府相关政策,企业的经营活动就更多反映了中央政府政策鼓励的方向。其次,本文采用规范的面板数据方法来估计政策关联程度对企业绩效的影响,并从企业所有制、规模、行业等角度区分影响程度的大小。最后,本文从政府补贴、税收返还等财政政策手段和信贷约束程度等货币政策角度评估这类影响存在的机制,并进一步考察扣除这些因素后高政策关联度是否仍然可以影响企业绩效。

本文的主要贡献有三点。第一,利用政府工作报告和企业年报等文本信息来量化企业经营活动的政策关联程度,为度量政策有效性提供了新视角。第二,多角度评估政策关联程度与企业绩效之间的关系。本文发现政策关联程度与上市公司净资产收益率(ROE)和资产收益率(ROA)均呈现正相关。进一步分析显示,企业业绩对政策关联程度更为敏感的企

业有如下特征：民营企业、中小规模企业和高政策关联行业的企业。第三，本文进一步从融资约束和政府补贴、税收返还等角度来评估政策关联程度影响企业业绩的机制。研究发现，政策关联度高的企业缓解了融资约束，获得了更多的政府补贴和税收返还；但即便扣除补贴和税收返还，高政策关联度企业仍然有更好的绩效。

### 一、上市公司政策关联度指标的构建

构建上市公司政策关联程度指标主要涉及两类文本：一是2009~2018年国务院政府工作报告<sup>①</sup>；二是同时期上市公司年报。我们选择2009年作为起始年份是因为当年为应对国际金融危机而出台了十大产业的调整和振兴规划，这可能使上市公司面临的政策在2009年前后出现结构性变化。为避免政策结构性变化和上市公司政策关联程度间产生混淆效应，我们将研究重点放在2009~2018年。

本部分分三步完成对上市公司政策关联度指标的构建和描述。我们在第一节介绍文本准备、词典构建；第二节介绍关联度指数的具体计算方法；第三节对政策关联度指数给出基本的统计描述。

#### 1. 文本数据准备、词典构建和分词

本文使用的国务院政府工作报告来源于中国政府网和新华网；上市公司年报则从巨潮资讯网获得，该网站是中国证券监督管理委员会指定的上市公司信息披露网站，其缺失部分用Wind数据库来补足。国务院政府工作报告和上市公司年报都有对上年状况和业绩的回顾以及对未来的展望，因此我们认为同年的政府工作报告与上市公司年报在时间上可以匹配，当然为了评估早期政府工作报告是否会跨年影响企业经营活动，我们在后续分析中也构建了跨期政策关联度指标。

要将该语料转换成可用于计量分析的关联度指标，需要先将文本库内所有文本（非结构化数据）转化数据矩阵（结构化数据），再通过计量或者统计方法，将数据矩阵转换成目标信息序列（沈艳等，2019）。要完成非结构化数据向结构化数据的转换，我们首先对文本作分词处理。由于中文是由连续的单个汉字组成，相比于采用空格分隔单词的英文，中文分词则需要根据语境完成。较为常见的汉语分词技术包括基于字符串匹配、基于理解、基于统计三大类，当前经济金融类文献中的分词技术往往结合上述三种方法形成分词工具（沈艳等，2019），如Jieba分词、清华THULAC、Yaha分词、盘古分词等。本文采用Python软件提供的业内认为中文分词较好的“Jieba”分词组件（<https://github.com/fxsjy/jieba>）。

为提高文本分词的准确度，需要根据具体研究问题首先明确专有名词词典。文献中常见的特定名词词典有Harvard IV-4心理社会词典、Loughran和Mcdonald（2011）构建的金融情感词典等。构建中文自然语言处理词典的研究早期采取翻译经典特定名词词典、并在其上做适当调整以符合中文语言规范的策略。例如，曾庆生等（2018）以Loughran和Mcdonald（2011）提供的金融情感词典为基础，翻译成相应的中文词汇并研究年报语调与内部人交易行为直接的关系。由于中文语境跟英文语境存在较大的差异，又由于不同研究问题可能涉及不同专有名词，直接翻译英文词典未必可以满足研究需求，就需要学者在现有词典基础上构建符合研究实际的词典。例如，Chen等（2018）在直接翻译Loughran和Mcdonald（2011）

<sup>①</sup> 由于当年发布的企业年报是上一年经营行为的总结，同样对于当年3月份发布的政府工作报告实际上是反映上年政府政策行为，为了跟企业年报年份对应，如无特殊说明，我们将当年发布的政府工作报告特指上年政府工作报告。

词典、结合国内已有金融词典的基础上,又根据人工标注股吧帖子情绪而挑选出能反映中国股市论坛情绪的新词汇构建金融情感词典,他们发现在情绪预测准确率方面,新词典较直接使用翻译词典准确率提高了30%。还有一些问题国外文献尚不存在专有词典,也需要研究者根据实际需要构建。

本文研究采取在现有词典基础上根据研究需要拓展词典的策略,其中专有名词词典包括:刻画区域和机构的专有名词,如城市名称、省份名称、地区名称、国家名称、政府机关名称等;灵格斯(Lingoes)的《汉英会计词典》和《英汉汉英会计金融词典》中的词汇;以国家统计局发布的《2017年国民经济行业分类》为依据,结合Wind数据库中上市公司经营范围和主营产品类型文本数据,采用启发式分词算法构建的专有名词词典。在具体分词中,我们按照通行规范,去除了无实际意义的停顿词和标点符号,比如“的”“地”“了”等;删除单个字;并删除了数字信息而只保留文字文本内容。由此我们建立起基于政府工作报告文本数据和企业年报文本数据的语料库,共涉及9548个词和词组。

完成分词后,我们需要还构建适合度量政策关联程度的政策类词汇词典。借鉴Hassan等(2017)构建企业政治风险语料库的方法,我们首先剔除政府工作报告中非政策类词汇,这包括删除:已有分词中不具有实际意义的动词、形容词、名词等非政策类词汇,如到位、进一步、今年等;可能在政府工作报告和年报中存在歧义的词汇,如固定资产投资、库存等为年报财务指标。最终构建的政策类词典一共包含3145个词<sup>①</sup>组成。

## 2. 政策关联程度指标构建

完成词典构建和分词之后,我们需要进一步将所有 $t=1, \dots, T$ 个文本转化为数字矩阵,这涉及如何将每篇文档转换为一个向量。我们分别采用三种方法来生成这一向量,即词典法等权、词典法TF-IDF(Term Frequency-Inverse Document Frequency)和Word2Vec。词典法主要依靠独热法(One-Hot Presentation)将语料视作从所有词语库中挑选若干独立词汇形成的集合,根据语料中出现的全部词语构建一个词表,并将每个词按顺序编号 $1, 2, 3, \dots, N$ 。然后,将词语 $j$ 用一个 $N$ 维向量 $w_j$ 来表示,该向量的第 $j$ 个位置的元素为1其余均为0。这一方法忽略语法和语序等要素,在每一个词都转换为向量后,通过加总所有词的向量,将文本 $t$ 转化为 $1 \times N$ 的向量 $W_t$ ,其中 $w_{j_t}$ ( $j=1, \dots, N$ )是第 $j$ 个词语是否在文本中出现,或者在文本 $t$ 中出现的频率。独热法最早应用在自然语言处理和信息检索领域,在金融领域,Manela和Moreira(2017)使用该方法,根据华尔街日报新闻的标题和摘要中全部词语出现的频率来提取新闻数据的特征。

在决定矩阵内每个具体数值时,就需要选择合适的加权方法(Jegadeesh和Wu, 2013)。常用的加权方法有等权重和词频-逆文档(TF-IDF)加权。顾名思义,等权重法假定文本中每个词语的重要程度相同。例如,Hoberg和Phillips(2016)认为是否用一个词描述业务范围比该词出现的词频更重要,因此在向量中只记录一个词是否属于事先给定的词典。TF-IDF加权方法则同时考虑词语在文本中出现的次数(频率)和多少文档包含该词语这两个维度,对在文本中频繁出现但并没有实际含义的词语赋予较少的权重,而给予有重要含义但出现次数较少的词语较大权重。

本文采用两种权重来计算相似度。一是借鉴Hoberg和Phillips(2016),只记录一个词是否属于政策词典而不关注词频;二是采用TF-IDF来加权。我们用一个简单例子来说明。假定从政府工作报告中构建的政策词典包括四个词:[农业,工业,高科技,第三产业]。根

<sup>①</sup> 对该政策词典感兴趣的读者可与作者联系。

据这四个词,独热法词典的维度是  $1 \times 4$ 。其中,农业对应的向量是  $[1, 0, 0, 0]$ ,工业对应的向量是  $[0, 1, 0, 0]$ ,以此类推。假定整个语料包括一份政府工作报告和一份企业年报,其中政府工作报告包括的词是 [推动, 高科技, 产品, 发展, 第三产业],企业年报分词后得到 [我司, 工业, 高科技, 销售, 喜人, 加大, 投入]。在等权重法下,政府工作报告有“高科技”“第三产业”和政策词典重合,因此对应的向量为  $[0, 0, 1, 1]$ ,企业年报有“高科技”“工业”一词和政策词典重合,其对应的向量为  $[0, 1, 1, 0]$ 。

独热法的优势是操作简单,但有两个弱点。一是词汇量大时转换后的矩阵往往是高维稀疏数据矩阵。高维是因其维度由词典包含的词个数决定。例如本文中政策词典涉及 3145 个词,那么在独热法下每份政府工作报告和每份企业年报都会被转换为  $1 \times 3145$  的向量。稀疏是由于大部分词语出现频率低,导致向量中绝大部分元素的值为零。二是独热法可能因忽略上下文结构而产生歧义。词语嵌入技术 (Word Embedding) 的出现有助于解决这两个问题。该技术是把一个维数为所有词的数量的高维空间“嵌入”到一个维数低得多的连续向量空间中,即  $e_j = E \cdot W_j$ , 其中  $e_j$  表示第  $j$  个词通过嵌入矩阵  $E$  映射到实数域上的词向量,  $W_j$  为第  $j$  个词的独热向量表示。独热法可以看作是最简单的词嵌入方法,即  $e_j = E \cdot W_j = W_j$ 。由于该向量的每个元素值可以是连续值而不只是 0 或者 1,  $e_j$  的维度  $N_e$  可以远低于  $N$ 。常用词嵌入算法包括 Mikolov 等 (2013) 提出的 Word2Vec 技术和 Pennington 等 (2014) 开发的 GloVe (Global Vectors for Word Representation) 技术,其中 Word2Vec 的应用更为广泛。Word2Vec 的主要思想是先用向量代表各个词,然后通过神经网络模型,在大量的文本语料数据上来学习词向量的参数。训练后的模型不仅可以每个词语映射到一个低维的空间上 (通常为  $100 \sim 1000$  维),每个维数上的取值为连续值;并且根据不同词语的向量距离可以度量词语间的相似程度 (谢宗彦等, 2018),也解决了独热法下不同词语相互独立的问题。本文实际训练中利用 Google 开源的深度学习工具 Word2Vec 生成  $1 \times 200$  维的词向量。

在将每篇政府工作报告和每份企业年报都转换为词向量后,就可以计算文本相似度了。文本相似度的计算在文本分类、文本摘要、机器翻译、自动问答等方面均有广泛的应用 (田星等, 2018),本文采用较为常用的余弦相似度计算政府工作报告和企业年报的相似度。若将分词后的政府工作报告和企业年报分别以向量  $A$  和  $B$  表示,则余弦相似度度量了政府工作报告向量和企业年报向量之间的夹角:

$$SIM_{\cos}(A, B) = \frac{A \cdot B}{\|A\| \cdot \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (1)$$

其中,  $n$  是向量的维度,  $A_i$  是在第  $i$  位置向量的值。由式 (1) 可知,当两个文本的向量方向完全相反时,  $SIM_{\cos}(A, B)$  取值为  $-1$ ; 两个文本的向量相互垂直则取值为  $0$ ; 两个向量完全重合则取值为  $1$ 。也就是说,余弦相似度越接近  $0$  则两个文本越不相似;越接近  $1$  则文本相似度越高 (Madylova 和 Oguducu, 2009)。由于文本长度的变化不会影响相似度的计算,余弦相似度在实证中有较为广泛的运用。

### 3. 政策关联程度指标描述

本文分别利用词典法等权、词典法 TF-IDF 以及 Word2Vec 形成的词向量来计算政府工

作报告和企业年报之间的余弦相似度，得到三套企业政策关联程度指标。由于三种方法对应的词向量不同，政策关联度的绝对值在规模上有较大差异，词典法等权和 TF-IDF 得到的政策关联度均值分别为 0.3 和 0.049，远低于采用 Word2Vec 的均值 0.68。由此可见，忽略了上下文的词典法两序列在刻画文本相似度的能力上总体低于 Word2Vec。值得注意的是，由于 TF-IDF 降低高频出现词的权重的逻辑使这些词的信息含量低，但对政府工作报告而言，高频出现的词恰恰体现了政府关切，是民生发展重点。因此，采用 TF-IDF 为词语赋权不适用于本文研究。

图 1 给出 0—1 标准化后的采用词典法等权和 Word2Vec 构造词向量时的政策关联度年度均值走势。图 1 显示，词典法等权和 Word2Vec 的趋势相近。另外，比对词典法等权序列和 Word2Vec 计算得出的政策关联度可知，基于 Word2Vec 的政策关联度在各年间的变化比较平缓。因此，我们在后续分析中，以 Word2Vec 计算得出的政策关联度作为主要度量指标，而采用词典法等权作为稳健性检验指标。从图 1 可以看到，无论是词典法等权还是 Word2Vec，都可以看到 2012 年以后，企业年报的政策关联度的年度均值呈现逐年上升的趋势，也就是说企业经营活动对于政府政策的方向更为敏感。

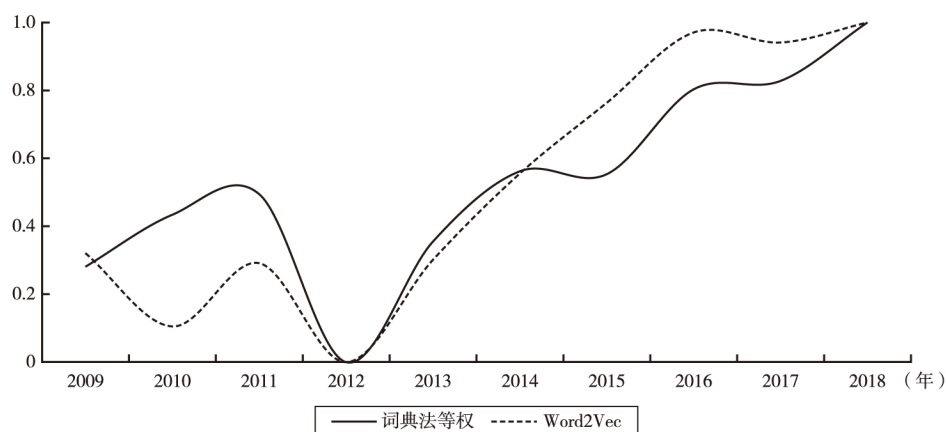


图 1 上市公司政策关联度

注：词典法等权指采用词典法对不同词语采用简单加权的方式构成词向量进而用余弦相似度计算得出的关联度。Word2Vec 值是采用该方法将词向量降维到 200 维后计算的相似度。

## 二、政策关联程度与企业绩效实证检验

### 1. 数据描述与计量模型

本文研究 2009~2018 年 A 股上市公司年报与政府工作报告的政策关联程度及其对企业绩效的影响。其中，上市公司财务数据来源于 Wind 数据库。借鉴已有研究，我们剔除金融类上市公司样本、已退市的公司和存在数据缺失的样本后共有观测值 24371 个<sup>①</sup>。

在变量选择上，本文用文献常用的净资产收益率（ROE）和资产收益率（ROA）两个指标来衡量企业业绩水平。政策关联度指标为关键解释变量，控制变量包括：财务杠杆率

<sup>①</sup> 因缺失删除的观测值为 21 个。为了消除极端值的影响，对本文主要连续变量采用 1% 的 Winsorize 缩尾处理。

(*LEV*), 以资产负债率来度量; 资金状况 (*LIQ*), 以流动资产与总资产之比来反映; 企业规模 (*SIZE*), 以总资产的自然对数衡量; 企业员工规模 (*STAFF*), 以企业员工总人数的自然对数为指标; 管理者风险厌恶程度 (*RA*), 以所有者权益与总资产之比为替代指标衡量<sup>①</sup>。除了上述变量外, 我们还从企业所有制、规模和行业政策关联程度三个角度来分析政策关联度的影响。

首先, 我们在表 1 刻画政策关联度高的上市公司的基本特征, 这主要是通过将主要变量控制省份虚拟变量后, 对高政策关联度虚拟变量回归而得, 其中高政策关联度是指政策关联度高于全样本中值的那些观测值。其次, 我们按照所有制将所有企业分为国有和民营两类<sup>②</sup>; 又按照行业政策关联度高低将样本分为高政策关联行业和低政策关联行业。表 1 显示, 政策关联度高的上市公司和政策关联度低的公司在企业绩效 (*ROE* 和 *ROA*) 和所有控制变量上都具有显著差异。政策关联度高的公司, 他们的绩效往往更好 (*ROE* 和 *ROA* 更高), 这些公司更可能是民营企业、中西部地区企业、中小企业、财务杠杆率低的企业、资金状况良好的企业、轻资产企业以及敢于冒风险的企业。

表 1 政策关联度高公司的基本特征

变量	关联度高	关联度低	差	标准误
	(1)	(2)	(3)	(4)
<i>ROE</i>	0.067	0.058	0.238***	0.0205
<i>ROA</i>	0.044	0.035	0.924***	0.0468
国有企业 = 1	0.298	0.431	-0.086***	0.0061
东部地区 = 1	0.646	0.693	-0.304***	0.0466
企业规模	12.714	12.820	-0.034***	0.0022
杠杆率	0.393	0.460	-0.329***	0.0130
资金状况	0.573	0.580	0.032**	0.0135
员工规模	7.622	7.504	0.007***	0.0022
风险厌恶程度	0.607	0.540	0.329***	0.0129

注: \*\*、\*\*\* 分别表示系数估计值在 5%、1% 水平上显著。“差”是指分组后 (政策关联度高的企业和政策关联度低的企业) 系数在两组之间是否存在显著差异。

本文采用面板数据固定效应模型来估计政策关联度和其他控制变量对企业绩效 (*ROE* 或 *ROA*) 的影响:

$$ROE(ROA)_{it} = \beta_0 + \beta_1 SIM_{it} + \beta_2 LEV_{it} + \beta_3 LIQ_{it} + \beta_4 SIZE_{it} + \beta_5 STAFF_{it} + \beta_6 RA + \lambda_t + \mu_i + \varepsilon_{it} \quad (2)$$

其中, 下标  $i$  和  $t$  分别表示公司和年度,  $\lambda_t$  和  $\mu_i$  分别代表不随个体改变的时间固定效应和不随时间改变的上市公司个体固定效应。由于同一个公司不同时期的扰动项可能存在自相关, 本文所有回归均采用聚类稳健标准误。

① 上述指标均借鉴相关文献中的构建方式, 如戴璐和宋迪 (2018)、杨继生和阳建辉 (2015)、谭之博和赵岳 (2012)、吴超鹏等 (2019)、Maudos 和 De Guevara (2004)。

② Wind 数据库中所有制形式分为国有企业、民营企业、外资企业、集体企业和公众企业。外资企业、集体企业和公众企业共有 2368 个观测值, 为避免混淆效应, 所以在所有制分类子样本中未被包括在内。



## 2. 政策关联度与企业业绩

由于国务院政府工作报告的内容不仅包含政府对上年工作的总结,还包含对未来的展望。政府工作报告提及的政策具有前瞻性和连贯性,对未来政府工作内容以及推行的政策具有指导意义,影响着企业下一年乃至更长时间的经营行为。为此,本文计算了当年政府工作报告与当年企业年报之间的政策关联度、一年前政府工作报告与当年企业年报之间的政策关联度,以及两年前政府工作报告与当年企业年报之间的关联度,并估计这三列指标对企业绩效的影响。表2的第(1)和(2)列、第(3)和(4)列及第(5)和(6)列分别报告了当年政府工作报告、一年前政府工作报告及两年前政府工作报告与当前企业年报的关联度对当年企业绩效的影响。对于每一个序列的政策关联度,我们都分别估计了不控制其他解释变量的总效应和控制解释变量的净效应模型。从表2可以看到,加入控制变量后当年企业政策关联度系数在5%的水平上显著为正,也就是当期政策关联度越高,则企业绩效也越好。一年前和两年前的政府工作报告对当年的绩效也均产生显著正向影响。这显示出政府工作报告所提及的政策具有延续性。

从控制变量看,财务杠杆率对企业绩效具有显著负向效应,即财务杠杆越高则企业绩效越差,这与杨德明和刘泳文(2018)等人的研究结论一致;资金状况对企业绩效具有显著正向效应,即资金状况越好企业绩效越高;公司规模存在规模效应;管理者风险厌恶程度刻画了管理者的风险态度和企业管理效率之间的关系,管理者风险厌恶程度越高的企业管理效率越低,绩效水平也越低,与杨继生和阳建辉(2015)的发现一致。因此本文使用的控制变量对绩效的影响或者符合预期,或者和现有文献发现一致,在后文中不再赘述。

表2 政策关联度与企业绩效(ROE):固定效应回归

变量	(1)	(2)	(3)	(4)	(5)	(6)
	当期报告		报告滞后一期		报告滞后两期	
政策关联度	0.127*** (0.0341)	0.0782** (0.0333)	0.129*** (0.0344)	0.0783** (0.0336)	0.133*** (0.0348)	0.0798** (0.0339)
财务杠杆		-0.843*** (0.161)		-0.845*** (0.161)		-0.844*** (0.161)
资金状况		0.0656*** (0.0127)		0.0654*** (0.0127)		0.0653*** (0.0127)
企业规模		0.0129*** (0.00447)		0.0129*** (0.00447)		0.0129*** (0.00447)
员工规模		0.00196 (0.00349)		0.00195 (0.00349)		0.00192 (0.00349)
风险厌恶		-0.677*** (0.159)		-0.680*** (0.159)		-0.679*** (0.160)
截距项	-0.00213 (0.0236)	0.569*** (0.166)	-0.00313 (0.0238)	0.572*** (0.166)	-0.00323 (0.0234)	0.571*** (0.166)
R <sup>2</sup>	0.295	0.312	0.295	0.312	0.295	0.312
观测值	24310	24310	24310	24310	24310	24310

注: \*、\*\*、\*\*\* 分别表示系数估计值在 10%、5%、1% 水平上显著, 变量括号内的值为聚类稳健标准误, 所有回归均控制了年度虚拟变量。

上述分析表明,政策关联度的总效果和净效果方向和显著性一致,并且一年后和两年后政策关联度对企业绩效影响的效果和当年类似。为节省篇幅并集中讨论,我们以 Word2Vec

当期政策关联度对  $ROE$  的影响作为主要讨论对象,并且直接报告控制了解释变量的净效果模型而不再报告总效果。后续在稳健性检验中我们再分别考察企业绩效变为  $ROA$  和关联度指标采用词典法等权方式构造时的稳健性。

### 3. 政策关联程度的异质性影响

上述分析从全样本角度评估了政策关联度对绩效的影响,但这一效果对不同规模、不同所有制、不同区域的企业可能都有所不同。在本小节我们采用分组回归的方式,检验不同维度下 Word2Vec 政策关联度对净资产收益率 ( $ROE$ ) 的影响。大量基于微观企业生产和贸易数据的研究发现,企业之间在生产规模、效率等方面存在显著差异,因此有必要评估不同规模企业的政策关联程度对企业绩效的影响。本文视资产规模大于行业平均规模(中位数)的企业为大规模企业,否则为中小规模企业。由于政府在对不同产权企业的资源分配中可能存在差异,本文将样本分为国有企业样本和民营企业样本,以考察不同产权企业的政策关联程度对绩效的异质性影响。另外,地区间经济发展差异也是影响政策制定和执行的重要因素,有必要考察不同地区企业政策关联度的影响力是否不同。在本文中,我们按照企业注册地址将样本分为东部地区和中西部地区<sup>①</sup>,以考察不同区域下政策关联程度对绩效的异质性影响。

表3重点考察政策关联度效果的异质性。其中,第(1)列和第(2)列基于企业所有制,第(3)列和第(4)列基于企业规模,而第(5)列和第(6)列则基于企业所在地区。这些估计结果表明,政策关联度对企业绩效的影响对民营企业、中小企业和东部地区企业都是正向,并且在1%或者5%的显著性水平上显著,但对国有企业、大企业和中西部地区企业绩效影响并不显著。

表3 政策关联程度对企业业绩影响

变量	(1)	(2)	(3)	(4)	(5)	(6)
	国有企业	民营企业	大企业	中小企业	东部地区企业	中西部地区企业
政策关联度	0.0175 (0.0456)	0.105** (0.0510)	0.00734 (0.0528)	0.107*** (0.0409)	0.0857** (0.0371)	0.0576 (0.0669)
财务杠杆	4.753 (231.2)	-0.871*** (0.165)	-1.156*** (0.389)	-0.844*** (0.163)	-1.386* (0.732)	-0.790*** (0.161)
资金状况	0.143*** (0.0211)	0.0313* (0.0177)	0.121*** (0.0308)	0.0560*** (0.0141)	0.0481*** (0.0141)	0.0926*** (0.0247)
企业规模	0.0212*** (0.00725)	0.00923 (0.00620)	0.0129 (0.00869)	0.0137*** (0.00520)	0.0134*** (0.00516)	0.0141* (0.00853)
员工规模	-0.00791 (0.00529)	0.0105** (0.00521)	0.00715 (0.00532)	0.000320 (0.00422)	0.00902** (0.00368)	-0.0111 (0.00741)
风险厌恶	4.968 (231.2)	-0.750*** (0.162)	-0.982** (0.397)	-0.681*** (0.161)	-1.246* (0.735)	-0.580*** (0.158)

<sup>①</sup> 参照财政部预算司地方财政分析资料(2004)中的分类(不含港澳台地区),东部地区包括北京、天津、辽宁、上海、江苏、浙江、福建、山东、广东9个省份;中部地区包括河北、山西、吉林、黑龙江、安徽、江西、河南、湖北、湖南、海南10个省份;西部地区包括内蒙古、广西、重庆、四川、贵州、云南、西藏、陕西、甘肃、青海、宁夏、新疆12个省份。

(续)

变量	(1)	(2)	(3)	(4)	(5)	(6)
	国有企业	民营企业	大企业	中小企业	东部地区企业	中西部地区企业
截距项	-5.091 (231.1)	0.630*** (0.172)	0.877** (0.400)	0.553*** (0.170)	1.087 (0.740)	0.558*** (0.180)
R <sup>2</sup>	0.315	0.327	0.368	0.297	0.326	0.293
观测值	8879	13002	4864	19446	16282	8028

注：同表 2。

从所有制角度看，随着民营企业对中国经济社会的贡献越来越大，国家政策对民营企业的支持力度也不断加强，民营企业获得越来越多的政策支持；加上民营企业更多地以市场导向为主，因此政策关联程度对民营企业的影响程度高于国有企业。从规模角度看，可能的解释是小规模企业现金流、营收等规模相对较小，对政府补贴、税收优惠带来的业绩增量更敏感，而大企业在跟进政策的过程中虽然也可以获得更多政府补贴、税收优惠等，但相对其营收规模而言这些优惠的影响并不显著。从地区角度看，东部地区享有更多的政策支持，意味着政府将更多的资源向该地区倾斜，政策关联程度对东部地区的企业业绩提升作用更为显著。

### 三、进一步机制分析

在前文实证分析中，高政策关联度企业的绩效水平更好，这一点对民营企业、中小企业和东部地区企业尤为明显。值得进一步追究的是高政策关联度作用于企业绩效的机制：是因为高政策关联度的企业本身是更优质的企业，还是因为这些企业可以获得更多的政府补贴、税收优惠，或者融资约束更少？例如，王克敏等（2015）发现地方政府通过财政补助等形式短期内增加新上市公司现金流，提升企业投资水平，改善短期业绩水平。陈冬华等（2010）也发现受到政策影响的产业，其 IPO 家数、融资额、再融资机会等均高于未受政策影响的产业。如果仅仅是由于高政府补贴、税收优惠和融资约束环节带来了绩效的改善，那么这种改进就不具备自生能力（王克敏等，2018）；如果去除优惠政策后高政策关联度企业仍然可以有更好的业绩，则表明这些企业本身更有可能是优质企业。在本部分，我们进一步从高政策关联度企业在政府补贴、税收返还和融资约束等方面的差异角度，来评估高政策关联度作用于企业绩效的机制。

#### 1. 政府补贴、税收返还与政策关联度

张杰等（2016）认为，政府补贴和税收返还是中国政府影响企业活动的一种普遍方式，2001~2010 年超过 50% 的上市公司得到了政府补贴或者税收优惠。如果政策关联高的企业能获得更多的以政府补贴和税收返还为表现的政策支持；这就会导致企业营业外收入增加，从而直接提升企业的绩效水平。为了检验政策关联高的企业是否会获取更多的政府补贴和税收返还，本文将政府补贴或者税收返还对政策关联度以及其他控制变量作回归：

$$GOV(TAX)_i = \beta_0 + \beta_1 SIM_i + \beta \sum Control_i + \lambda_i + \mu_i + \varepsilon_i \quad (3)$$

其中，GOV 和 TAX 分布是企业获取的政府补贴和税收返还（分别进行了对数化处理），其他符号和式（1）一致。我们在表 4、表 5、表 6 中分别报告以政府补贴和税收返还、

仅考虑政府补贴和仅考虑税收返还为因变量时,对于全样本,以及按是否是国有、是否是大企业、是否是东部地区子样本的回归结果。

从表4可以看出,就全样本而言,政策关联度高的企业获得了更多的政府补贴和税收返还,这一观察在1%显著性水平上显著。分样本后,可以看到无论企业所有制、规模和所在地区,只要政策关联度高,就会获得更多来自政府的优惠。

表4 政府补贴和税收返还与政策关联度

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	全样本	国有企业	民营企业	大企业	中小企业	东部地区企业	中西部地区企业
政策关联度	1.830*** (0.327)	1.476*** (0.541)	2.340*** (0.427)	1.512** (0.687)	1.920*** (0.372)	1.640*** (0.396)	2.149*** (0.573)
财务杠杆	-1.335 (1.157)	-510.8 (1.010)	-1.342 (1.192)	-18.85** (7.443)	-1.255 (1.163)	-13.12*** (3.550)	-1.249 (1.117)
资金状况	-0.00319 (0.114)	-0.418* (0.213)	-0.0450 (0.151)	-0.187 (0.301)	0.00926 (0.124)	0.233* (0.136)	-0.434** (0.200)
企业规模	0.682*** (0.0343)	0.710*** (0.0607)	0.727*** (0.0464)	0.563*** (0.0684)	0.712*** (0.0398)	0.644*** (0.0420)	0.754*** (0.0588)
员工规模	0.277*** (0.0306)	0.280*** (0.0470)	0.233*** (0.0435)	0.205*** (0.0589)	0.292*** (0.0353)	0.303*** (0.0377)	0.225*** (0.0525)
风险厌恶	-1.395 (1.151)	-511.1 (1.010)	-1.295 (1.179)	-19.55*** (7.462)	-1.191 (1.154)	-13.20*** (3.565)	-1.240 (1.096)
截距项	-4.630*** (1.237)	504.6 (1.010)	-4.895*** (1.321)	15.52** (7.435)	-5.250*** (1.267)	7.546** (3.579)	-5.446*** (1.344)
R <sup>2</sup>	0.713	0.709	0.717	0.704	0.675	0.718	0.707
观测值	24305	8877	13001	4863	19442	16278	8027

注:同表2。

表5显示,就政府补贴而言,总体上政策关联度高的企业获得了更高的政府补贴。分样本来看,无论是否为国有企业,也无论规模和地区,这一现象都存在。通过分别引入政策关联度和所有制、规模、地区的交叉项进一步检验政策关联度系数是否存在组间差异时发现,政策关联度高带来更多政府补贴这一点并不存在规模和地区上的差异(交叉项对应的系数分别为0.756和-0.0536,且均不显著)。

表5 政府补贴与政策关联度

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	全样本	国有企业	民营企业	大企业	中小企业	东部地区企业	中西部地区企业
政策关联度	1.868*** (0.330)	1.636*** (0.543)	2.275*** (0.434)	1.529** (0.687)	1.969*** (0.376)	1.727*** (0.400)	2.079*** (0.577)
财务杠杆	-1.250 (1.147)	-350.1 (1.018)	-1.196 (1.181)	-18.72** (7.293)	-1.124 (1.151)	-13.03*** (3.599)	-1.055 (1.100)
资金状况	0.0298 (0.114)	-0.383* (0.217)	0.00471 (0.150)	-0.309 (0.301)	0.0705 (0.124)	0.231* (0.136)	-0.344* (0.200)
企业规模	0.678*** (0.0344)	0.686*** (0.0612)	0.738*** (0.0466)	0.557*** (0.0682)	0.708*** (0.0400)	0.643*** (0.0422)	0.743*** (0.0586)

(续)							
变量	(1) 全样本	(2) 国有企业	(3) 民营企业	(4) 大企业	(5) 中小企业	(6) 东部地区企业	(7) 中西部地区企业
员工规模	0.278*** (0.0309)	0.287*** (0.0474)	0.229*** (0.0441)	0.223*** (0.0588)	0.288*** (0.0358)	0.306*** (0.0380)	0.225*** (0.0531)
风险厌恶	-1.338 (1.141)	-350.4 (1,018)	-1.199 (1.169)	-19.46*** (7.312)	-1.089 (1.142)	-13.12*** (3.613)	-1.127 (1.079)
截距项	-4.767*** (1.228)	344.0 (1,018)	-5.166*** (1.312)	15.34** (7.286)	-5.437*** (1.257)	7.348** (3.626)	-5.575*** (1.330)
R <sup>2</sup>	0.716	0.714	0.719	0.706	0.679	0.720	0.710
观测值	24306	8878	13001	4863	19443	16279	8027

注：同表 2。

表 6 反映了税收返还和政策关联度之间的关系。总体来看，仍然存在政策关联度高的企业税收返还多这一特征。分样本回归显示，这一发现对所有制、规模、地区差异稳健，同样通过分别引入政策关联度和所有制、规模、地区的交叉项进一步对不同规模和不同地区中政策关联变量的回归系数进行了组间差异检验，发现不同规模组间和不同地区组间政策关联系数存在显著差异（交叉项对应的系数为 0.66 和 1.494，且分别在 10% 和 1% 的水平上显著），结果表明政策关联度高的大规模企业和中西部地区企业获得更多的税收返还。

表 6 税收返还与政策关联度

变量	(1) 全样本	(2) 国有企业	(3) 民营企业	(4) 大企业	(5) 中小企业	(6) 东部地区企业	(7) 中西部地区企业
政策关联度	0.633*** (0.172)	0.565** (0.239)	0.615** (0.265)	1.012*** (0.319)	0.509** (0.205)	0.546*** (0.203)	0.814** (0.323)
财务杠杆	-0.278 (0.349)	71.47 (350.7)	-0.120 (0.162)	-2.893 (3.443)	-0.294 (0.336)	-16.36 (20.71)	-0.269 (0.210)
资金状况	0.0489 (0.0566)	-0.0896 (0.114)	-0.0166 (0.0767)	0.291** (0.145)	-0.00265 (0.0622)	0.181** (0.0708)	-0.168* (0.0940)
企业规模	0.0319* (0.0167)	0.0804*** (0.0285)	0.00878 (0.0234)	0.00532 (0.0322)	0.0398** (0.0194)	0.0162 (0.0213)	0.0635** (0.0281)
员工规模	-0.00360 (0.0143)	-0.0113 (0.0222)	-0.0111 (0.0216)	-0.000390 (0.0278)	-0.00590 (0.0166)	-0.0118 (0.0154)	-0.00178 (0.0297)
风险厌恶	-0.215 (0.344)	71.63 (350.7)	-0.0407 (0.144)	-2.981 (3.447)	-0.203 (0.328)	-16.32 (20.71)	-0.149 (0.185)
截距项	-0.324 (0.406)	-72.63 (350.6)	-0.0919 (0.345)	2.236 (3.393)	-0.265 (0.406)	15.99 (20.69)	-0.781** (0.398)
R <sup>2</sup>	0.522	0.428	0.568	0.517	0.523	0.571	0.356
观测值	24306	8877	13000	4863	19443	16278	8028

注：同表 2。

前文实证结果显示政策关联度高的上市公司有更好的绩效，如果仅仅是由于这些企业得到了更多的补贴或者税收返还，那么政府推行政策、发展新产业的初衷就可能被歪曲。为

此,我们扣除政府补贴和税收返还后重新计算净资产收益率,再估计政策关联度对绩效的影响,结果见表7。表7显示,全样本下,政策关联度系数值大小和显著程度跟表2第2列回归基本保持一致,政策关联度越高的企业净资产收益率仍然更高。这表明总体来看,政策关联度高的企业绩效好并不仅仅因为获得了政府补贴和税收返还。分样本来看,民营企业、中小企业和东部地区企业,在扣除政府补贴和税收返还后,仍然可以看到高政策关联度的企业有更高的绩效水平。

表7 扣除政府补贴和税收返还后资产收益率后检验

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	全样本	国有企业	民营企业	大企业	中小企业	东部地区企业	中西部地区企业
政策关联度	0.0742** (0.0340)	0.0174 (0.0472)	0.101* (0.0513)	-0.000227 (0.0531)	0.104** (0.0418)	0.0703* (0.0378)	0.0771 (0.0681)
财务杠杆	-0.878*** (0.170)	-71.23 (242.7)	-0.903*** (0.174)	-1.003** (0.486)	-0.880*** (0.173)	-1.477* (0.790)	-0.832*** (0.171)
资金状况	0.0639*** (0.0129)	0.140*** (0.0217)	0.0293 (0.0178)	0.121*** (0.0312)	0.0541*** (0.0143)	0.0466*** (0.0144)	0.0909*** (0.0250)
企业规模	0.0184*** (0.00457)	0.0293*** (0.00744)	0.0130** (0.00630)	0.0163* (0.00879)	0.0199*** (0.00533)	0.0185*** (0.00525)	0.0200** (0.00879)
员工规模	0.00233 (0.00355)	-0.00773 (0.00538)	0.0112** (0.00529)	0.00826 (0.00537)	0.000456 (0.00430)	0.00916** (0.00372)	-0.0105 (0.00757)
风险厌恶	-0.680*** (0.168)	-70.97 (242.7)	-0.756*** (0.171)	-0.794 (0.493)	-0.685*** (0.171)	-1.306* (0.793)	-0.587*** (0.167)
截距项	0.513*** (0.174)	70.76 (242.7)	0.594*** (0.180)	0.650 (0.494)	0.491*** (0.179)	1.100 (0.798)	0.486*** (0.188)
R <sup>2</sup>	0.333	0.349	0.337	0.384	0.319	0.340	0.320
观测值	24310	8879	13002	4864	19446	16282	8028

注:同表2。

## 2. 企业融资约束与政策关联度

融资约束制约是中国企业发展面临的主要限制条件之一(邓可斌和曾海舰,2014)。因此,政策常常也会包含信贷支持,即相关企业更容易获取银行或者其他金融机构的贷款,从而缓解企业融资约束,推动业绩提升。鉴于此,我们在本小节考察政策关联程度与融资约束之间的关系。

度量融资约束有不同的分类标准,早期分类标准主要依据与信息成本相关的企业特征。但由于只考虑企业融资的某一类指标因缺乏对融资整体的描述而存在缺陷(刘莉亚等,2015)。文献常用Lamont等(2001)构造的KZ指数、Whited和Wu(2006)根据季度财务报表构造的WW指数来度量融资约束,但是上述两种方法均包含了很多具有内生性的金融变量。为解决内生性问题,Hadlock和Pierce(2010)在Kaplan和Zingales(1995)方法的基础上构造了SA指数,目前已被广泛运用。本文采用两种方法来刻画融资约束,一是借鉴Hadlock和Pierce(2010)的SA指数公式<sup>①</sup>,SA即为融资约束程度,其值越大,企业面临

<sup>①</sup> SA指数构造公式是 $SA = -0.737 \times SIZE + 0.043 \times SIZE^2 - 0.04 \times AGE$ ,其中SIZE是企业总资产对数值,AGE为企业上市年限。

的融资约束越严重。二是除了直接度量融资约束外，我们也采用经营活动产生的现金流量净额占总资产比例反映企业的融资约束状况。由于企业的现金流明显影响着企业的融资约束，企业的现金流越大，企业的外部融资成本越低，面临的融资约束也越低（鞠晓生等，2013）。为检验政策关联程度能否缓解企业融资约束，我们估计：

$$SA(CASH)_i = \beta_0 + \beta_1 SIM_i + \beta \sum Control_i + \lambda_t + \mu_i + \varepsilon_{it} \quad (4)$$

其中，下标  $i$  和  $t$  分别表示公司和年度， $\lambda_t$  和  $\mu_i$  分别代表不随个体改变的时间固定效应和不随时间改变的上市公司个体固定效应， $Control$  是公司层面的特征变量，并在表 8 报告回归结果。

表 8 第 (1) 列显示，企业政策关联程度和融资约束呈反向关系，即企业政策关联程度越高，融资约束越低，这一效果在 1% 显著性水平下显著。表 8 第 (2) 列则发现政策关联程度越高的企业，现金流水平也越高。因此，我们从直接度量融资约束程度以及从现金流水平这两个角度，都观察到政策关联程度高的企业融资能力更强，因此也更有将资源用于提高业绩的可能。

表 8 政策关联缓解融资约束有效性检验

变量	(1)	(2)	(3)	(4)
	SA	SA	CASH	CASH
政策关联度	-0.0759*** (0.0185)	-0.0704*** (0.0175)	0.0344** (0.0145)	0.0336** (0.0144)
财务杠杆		0.131** (0.0659)		-0.343*** (0.0797)
资金状况		0.0521*** (0.00611)		-0.0375*** (0.00519)
企业规模		-0.0510*** (0.00225)		-0.0151*** (0.00175)
员工规模		-0.000492 (0.00176)		0.00674*** (0.00138)
风险厌恶		0.113* (0.0657)		-0.331*** (0.0795)
截距项	-3.169*** (0.0125)	-2.705*** (0.0736)	0.0443*** (0.00983)	0.537*** (0.0830)
R <sup>2</sup>	0.969	0.973	0.413	0.419
观测值	24310	24310	24310	24310

注：同表 2。

#### 四、稳健性检验

本文从三个角度来检验实证结果的稳健性：一是更换企业业绩水平指标为总资产收

益率；二是为了应对 Word2Vec 计算类似于“黑箱”的问题而采用更为直观的词典法来计算政策关联度；三是为了避免词典构造中的主观性，选取最宽泛的词典作为政策词典。

#### 1. 以总资产收益率作为业绩水平

前文以净资产收益率作为公司业绩的衡量指标，该指标重在度量股东资金税后净利润的能力，而总资产收益率则重在测量公司运用全部资产产生税前利润的能力。本小节以 ROA 为因变量，估计了政策回归的基准模型、分样本模型以及扣除财政补贴和税收返还后计算的 ROA 与政策关联度之间的关系。我们发现结果总体稳健。为节省篇幅，表 9 报告了未扣除政府补贴和税收返还后以 ROA 为因变量时的回归结果。虽然政策关联度对总资产收益率的关联程度有所下降（如全样本系数为 0.052，全样本下对 ROE 的回归中则是 0.0782），但总体上仍然显著为正，并且分样本中仍然可以观察到对民营企业、中小企业和东部地区企业有显著效果。这一结果与使用 ROE 的回归结果（见表 7）类似。

表 9 以总资产收益率为被解释变量的回归结果

变量	(1)	(1)	(2)	(3)	(4)	(5)	(6)
	全样本	国有企业	民营企业	大企业	中小企业	东部地区企业	中西部地区企业
政策关联度	0.0520*** (0.0114)	0.00905 (0.0145)	0.0742*** (0.0185)	0.0207 (0.0162)	0.0633*** (0.0143)	0.0656*** (0.0137)	0.0247 (0.0205)
财务杠杆	-0.275*** (0.0499)	234.8*** (57.22)	-0.287*** (0.0510)	-0.407 (0.298)	-0.267*** (0.0500)	-0.196 (0.453)	-0.225*** (0.0520)
资金状况	0.0504*** (0.00456)	0.0702*** (0.00746)	0.0353*** (0.00651)	0.0516*** (0.0113)	0.0498*** (0.00503)	0.0400*** (0.00545)	0.0669*** (0.00819)
企业规模	0.00856*** (0.00152)	0.00898*** (0.00218)	0.00764*** (0.00224)	0.00741** (0.00321)	0.00902*** (0.00172)	0.00582*** (0.00186)	0.0144*** (0.00268)
员工规模	0.00163 (0.00119)	-0.00113 (0.00159)	0.00323* (0.00191)	0.00379* (0.00210)	0.000900 (0.00140)	0.00543*** (0.00131)	-0.00576** (0.00237)
风险厌恶	-0.151*** (0.0495)	235.0*** (57.22)	-0.182*** (0.0503)	-0.260 (0.299)	-0.149*** (0.0495)	-0.0756 (0.453)	-0.0945* (0.0510)
截距项	0.0739 (0.0524)	-235.0*** (57.22)	0.104* (0.0563)	0.215 (0.300)	0.0620 (0.0532)	0.00310 (0.454)	0.00832 (0.0580)
R <sup>2</sup>	0.519	0.538	0.518	0.596	0.506	0.517	0.518
观测值	24310	8879	13002	4864	19446	16282	8028

注：同表 2。

#### 2. 以词典法作为计算政策关联度的方法

由于政策关联度是本文关键解释变量，而前文展示了三种不同的政策关联度计算方法。虽然 TF-IDF 因其加权方式不适用于国务院政府工作报告这样的文件，图 1 显示采用词典法等权政策关联度虽然比 Word2Vec 政策关联度在趋势上高度相关，因为忽视上下文



而波动更大,但是计算方法也更为直观。在本小节中我们采用词典法等权来估计模型。表 10 给出了基于词典法等权构造政策关联度后的回归结果。其中列 (1) 和列 (2) 以 *ROE* 为因变量,列 (3) 和列 (4) 以 *ROA* 为因变量的总效果和净效果,其政策关联度系数均在 1% 或者 10% 的水平上显著为正。表 10 说明,采用词典法等权的估计结果和采用 Word2Vec 一致。

表 10 基于词典法等权基本回归

变量	(1)	(2)	(3)	(4)
	<i>ROE</i>	<i>ROE</i>	<i>ROA</i>	<i>ROA</i>
政策关联度	0.179*** (0.0601)	0.110* (0.0570)	0.109*** (0.0222)	0.0613*** (0.0204)
财务杠杆		-0.847*** (0.160)		-0.277*** (0.0511)
资金状况		0.0659*** (0.0127)		0.0507*** (0.00456)
企业规模		0.0122*** (0.00448)		0.00815*** (0.00153)
员工规模		0.00199 (0.00350)		0.00169 (0.00119)
风险厌恶		-0.680*** (0.159)		-0.153*** (0.0507)
截距项	0.0320* (0.0177)	0.602*** (0.166)	0.0194*** (0.00654)	0.0976* (0.0534)
$R^2$	0.295	0.312	0.468	0.519
观测值	24310	24310	24310	24310

注:同表 2。

### 3. 更为宽泛的词典

政策关联度的高低除了取决于计算方法外,也取决于我们使用的词典。在构建政策词典时我们运用了对中国政策的理解来筛选词语,但是这一做法增加了主观性。为了防止词典主观性带来计算偏误,我们选用更为宽泛的词典来做稳健性检验。具体来说,对分完词的政府工作报告,在剔除的、地、了等停顿词,去除数字后不再做处理,直接以剩余的词构建政策类词典,在此基础上采用同样的方法构建政策关联程度指标,分别以 *ROE* 和 *ROA* 为被解释变量检验本文的基本结果是否依然显著(见表 11)。这一结果表明,虽然用更宽泛的词典可能增加一些政府工作报告和企业年报都有并反映政策信息的词汇,但是由于这两类文件的读者存在很大差异,这两类文本的风格差异也很大,因此增加的一些非政策词汇并未从本质上改变指标,也说明本文的发现并非由于可以挑选词典导致。

表 11 新词典下政策关联指标与企业业绩

变量	(1)	(2)	(3)	(4)
	ROE	ROE	ROA	ROA
政策关联度	0.0577*** (0.0134)	0.0388*** (0.0133)	0.0305*** (0.00509)	0.0168*** (0.00484)
财务杠杆		-0.846*** (0.161)		-0.277*** (0.0510)
资金状况		0.0652*** (0.0127)		0.0504*** (0.00457)
企业规模		0.0119*** (0.00451)		0.00811*** (0.00154)
员工规模		0.00202 (0.00350)		0.00173 (0.00119)
风险厌恶		-0.680*** (0.159)		-0.152*** (0.0506)
截距项	0.0668*** (0.00565)	0.626*** (0.166)	0.0419*** (0.00208)	0.110** (0.0533)
R <sup>2</sup>	0.295	0.312	0.469	0.519
观测值	24310	24310	24310	24310

注：同表 2。

当然，要识别企业政策关联程度对绩效的影响还需要进一步考虑是否存在遗漏变量或者反向因果，即政策是否会根据企业绩效作出调整。对于遗漏变量问题，我们参考已有的文献，尽可能合理地选取影响企业绩效的控制变量，同时本文的回归均同时控制了时间固定效应和个体效应模型，这样既能控制影响公司绩效且短期内不随时间发生变化的个体因素，也控制了诸如宏观经济环境等随时间变化但不随公司个体变化的时间因素。另外，由于国务院政府工作报告是每年全国人民代表大会发布的全国性的公文，因为单个企业的绩效而调整政府工作报告的可能性极低，因此我们认为对于政策关联度这一指标而言存在反向因果的可能性较低。

## 五、结论与建议

政策研究离不开政府工作报告，它集中体现了政府对过去一年以及未来一段时期内各个产业的政策信息，是研究政策方向不可多得的重要素材。但如何通过实证研究来匡算政府工作报告的政策影响程度并挖掘影响的具体机制存在一个重要挑战，就是政策不易量化，尤其是研究政府工作报告这种以文本为载体的信息。近年来，文本大数据分析和机器学习在经济学和金融学中得到了广泛的运用，为量化政策提供了新的思路和方法。我们认为，政府工作报告是政策的权威载体，企业年报文本含有政策支持的产业信息，因此企业年报和政府工作报告两者的文本数据必然存在文本相似，即政策关联，如何测算两者的关

联或者相似程度是关键。鉴于此,本文利用计算机科学的文本大数据分析技术和机器学习技术对企业年报和政府工作报告的政策关联程度进行测算,以政府工作报告和企业年报为文本数据来源,使用自然语言处理技术测算企业年报和政府工作报告的政策关联程度,并从企业经营等多个角度实证分析政策关联对企业的影响,为量化政策以及计算机科学的自然语言处理技术在经济学领域的运用提供了新视角。同时,研究结论这对评估政策有效性有着重要的理论和现实意义。

本文主要发现有以下六点。第一,2012年以后企业经营活动与政府政策的关联度显著上升。第二,对政策更敏感的企业,更可能是民营企业、中小企业,或东部地区企业;这类企业往往资金状况良好,财务杠杆率低;属于轻资产企业,并且更敢于冒风险。第三,高政策关联度显著提升了企业的绩效水平,这一结论在使用滞后一期和滞后两期的政府工作报告计算的政策关联指标依然显著,这一发现无论是对净资产收益率还是对总资产收益率均成立。第四,考虑到不同规模、不同所有制、不同地区的企业存在差异,采用分组回归的方式,检验不同维度下政策关联程度对净资产收益率的影响发现,政策关联度高的民营企业、中小企业和东部地区企业,其绩效提升效应更为显著,对国有企业、大企业和中西部地区企业绩效提升效应不显著。第五,进一步剖析其作用机理可知,政策关联度可以通过两个途径影响企业绩效:其一,政策关联程度高的企业获得更多来自政府的补贴和税收优惠,这直接提升了企业绩效水平,且无论企业所有制、规模和所在地区,只要政策关联度高,就会有机会获得更多来自政府的补贴和税收优惠;其二,政策关联程度越高的企业融资约束越低,可以获得更多短期融资,进而驱动了企业绩效的改善。第六,扣除了政府补贴和税收优惠后,政策关联度高的企业仍然有更好的绩效。

上述发现表明,为获取政策相关优惠而采取策略性创新并不是政策有效的唯一途径,事实上政策在不同阶段都存在有效的空间。从短期看,融资支持、税收优惠、政府补贴等政策支持会直接改善企业现金流状况;从中期看,企业有优化自身业务以在未来获取更多政策支持的动机,这点从民营企业、中小型企业 and 东部地区企业的绩效对政策关联度更敏感可以看出;从长期看,紧跟政策的企业释放了行业有较好发展前景的信号,通过充分利用政府政策,企业可以改善信息不对称、降低企业投资不确定性、降低交易成本,进而形成竞争优势。

综合政策关联指标测算和实证分析结果可知,政策关联度高能显著提升企业的绩效水平,但对不同规模、不同所有制、不同地区的企业存在差异,为了更好地提升政府政策实施的效率以及提升政策对企业的推导效用,本文从以下几个维度提出政策建议。第一,相关政策的实施仍然存在改进空间。例如,虽然本文发现政策关联度高的中小企业在税收返还方面没有得到更多的优惠。如果政府可以在政策制定中结合产业发展需求,并适当地向民营企业 and 中小企业倾斜,同时强化融资、税收、财政补贴等措施在资源优化配置中的作用,就能更好地实现政策的调控目标。第二,政策关联对企业的绩效影响存在显著的地域差异,对东部地区的企业绩效提升效应显著,反映了政府政策在地理空间存在差异,因此对地方政策,政府应该适当关注政策对中西部地区企业的支持,以平衡政策资源分配。第三,政府政策带有导向作用,且政府政策的执行往往配套融资、税收、财政补贴等方面的支持,因此对企业而言,则需要关注国家政策发展的方向,需要积极地利用政策带来的便利和资源,根据政府政策对自身业务范围做出有益的调整,实现绩效水平的提升。

## 参 考 文 献

- [1] Aghion P., Cai J., Dewatripont M., Du L., Harrison A., Legros P., 2015, *Industrial Policy and Competition* [J], American Economic Journal: Macroeconomics, 7 (4), 1~32.
- [2] Chen Y., Huang Z., Shen Y., 2018, *Textual Investor Sentiment and Stock Market: Evidence from China* [R], Working Paper.
- [3] Hoberg G., Phillips G., 2016, *Text-Based Network Industries and Endogenous Product Differentiation* [J], Journal of Political Economy, 124 (5), 1423~1465.
- [4] Hadlock C. J., Pierce J. R., 2010, *New Evidence on Measuring Financial Constraints: Moving beyond the KZ Index* [J], Review of Financial Studies, 23 (5), 1909~1940.
- [5] Hassan T. A., Hollander S., Van Lent L., Tahoun A., 2019, *Firm-Level Political Risk: Measurement and Effects* [R], Quarterly Journal of Economics, 134 (4), 2135~2202.
- [6] Jegadeesh N., Wu D., 2013, *Word Power: A New Approach for Content Analysis* [J], Journal of Financial Economics, 110 (3), 712~729.
- [7] Kaplan S. N., Zingales L., 1995, *Do Financing Constraints Explain Why Investment Is Correlated with Cash Flow?* [R], NBER Working Papers, No. 5267.
- [8] Krugman P. R., 1983, *Targeted Industrial Policies: Theory and Evidence* [J], Proceedings-Economic Policy Symposium-Jackson Hole, Federal Reserve Bank of Kansas City, 123~176.
- [9] Lamont O., Polk C., Saaá-Requejo J., 2001, *Financial Constraints and Stock Returns* [J], Review of Financial Studies, 14 (2), 529~554.
- [10] Loughran T., McDonald B., 2011, *When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks* [J], Journal of Finance, 66 (1), 35~65.
- [11] Madylova A., Oguducu S. G., 2009, *A Taxonomy Based Semantic Similarity of Documents Using the Cosine Measure* [C], in 2009 24th International Symposium on Computer and Information Sciences.
- [12] Manela A., Moreira A., 2017, *News Implied Volatility and Disaster Concerns* [J], Journal of Financial Economics, 123 (1), 137~162.
- [13] Maudos J. N., De Guevara J. F., 2004, *Factors Explaining the Interest Margin in the Banking Sectors of the European Union* [J], Journal of Banking & Finance, 28 (9), 2259~2281.
- [14] Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean J., 2013, *Distributed Representations of Words and Phrases and Their Compositionality* [R], in Advances in Neural Information Processing Systems, October 2013.
- [15] Pennington J., Socher R., Manning C. D., 2014, *Glove: Global vectors for word representation* [C], Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing.
- [16] Peters M., Schneider M., Griesshaber T., Hoffmann V. H., 2012, *The Impact of Technology-Push and Demand-Pull Policies on Technical Change-Does the Locus of Policies Matter?* [J], Research Policy, 41 (8), 1296~1308.
- [17] Whited T. M., Wu G., 2006, *Financial Constraints Risk* [J], Review of Financial Studies, 19 (2), 531~559.
- [18] 蔡庆丰、田霖：《产业政策与企业跨行业并购：市场导向还是政策套利》[J]，《中国工业经济》2019年第1期。
- [19] 陈冬华、李真、新夫：《产业政策与公司融资——来自中国的经验证据》[C]，2010年中国会计与财务研究国际研讨会。
- [20] 戴璐、宋迪：《高管股权激励合约业绩目标的强制设计对公司管理绩效的影响》[J]，《中国工业经济》2018年第4期。
- [21] 邓可斌、曾海舰：《中国企业的融资约束：特征现象与成因检验》[J]，《经济研究》2014年第

2 期。

[22] 鞠晓生、卢荻、虞义华：《融资约束，营运资本管理与企业创新可持续性》[J]，《经济研究》2013 年第 1 期。

[23] 林毅夫：《产业政策与我国经济的发展：新结构经济学的视角》[J]，《复旦学报（社会科学版）》2017 年第 2 期。

[24] 林毅夫、向为、余森杰：《区域型产业政策与企业生产率》[J]，《经济学（季刊）》2018 年第 2 期。

[25] 宋凌云、王贤彬：《重点产业政策、资源重置与产业生产率》[J]，《管理世界》2013 年第 12 期。

[26] 谭劲松、冯飞鹏、徐伟航：《产业政策与企业研发投入》[J]，《会计研究》2017 年第 10 期。

[27] 韩乾、洪永森：《国家产业政策、资产价格与投资者行为》[J]，《经济研究》2014 年第 12 期。

[28] 韩永辉、黄亮雄、王贤彬：《产业政策推动地方产业结构升级了吗？——基于发展型地方政府的理论解释与实证检验》[J]，《经济研究》2017 年第 8 期。

[29] 刘莉亚、何彦林、王照飞、程天笑：《融资约束会影响中国企业对外直接投资吗？——基于微观视角的理论和实证分析》[J]，《金融研究》2015 年第 8 期。

[30] 陆国庆、王舟、张春宇：《中国战略性新兴产业政府创新补贴的绩效研究》[J]，《经济研究》2014 年第 7 期。

[31] 钱雪松、康瑾、唐英伦、曹夏平：《产业政策、资本配置效率与企业全要素生产率——基于中国 2009 年十大产业振兴规划自然实验的经验研究》[J]，《中国工业经济》2018 年第 8 期。

[32] 杨德明、刘泳文：《“互联网+”为什么加出了业绩》[J]，《中国工业经济》2018 年第 5 期。

[33] 余明桂、范蕊、钟慧洁：《中国产业政策与企业技术创新》[J]，《中国工业经济》2016 年第 12 期。

[34] 沈艳、陈赟、黄卓：《文本大数据分析在经济学和金融学中的运用：一个文献综述》[J]，《经济学（季刊）》2019 年第 4 期。

[35] 伊志宏、杨圣之、陈钦源：《分析师能降低股价同步性吗——基于研究报告文本分析的实证研究》[J]，《中国工业经济》2019 年第 1 期。

[36] 谢宗彦、黎峻、周纯洁：《基于 word2vec 的酒店评论情感分类研究》[J]，《北京联合大学学报》2018 年第 4 期。

[37] 田星、郑瑾、张祖平：《基于词向量的 Jaccard 相似度算法》[J]，《计算机科学》2018 年第 7 期。

[38] 杨继生、阳建辉：《行政垄断、政治庇佑与国有企业的超额成本》[J]，《经济研究》2015 年第 4 期。

[39] 谭之博、赵岳：《企业规模与融资来源的实证研究——基于小企业银行融资抑制的视角》[J]，《金融研究》2012 年第 3 期。

[40] 吴超鹏、薛南枝、张琦、吴世农：《家族主义文化、“去家族化”治理改革与公司绩效》[J]，《经济研究》2019 年第 2 期。

[41] 王克敏、王华杰、李栋栋、戴杏云：《年报文本信息复杂性与管理者自利——来自中国上市公司的证据》[J]，《管理世界》2018 年第 12 期。

[42] 王克敏、刘静、李晓溪：《产业政策、政府支持与公司投资效率研究》[J]，《管理世界》2017 年第 3 期。

[43] 王克敏、杨国超、刘静、李晓溪：《IPO 资源争夺、政府补助与公司业绩研究》[J]，《管理世界》2015 年第 9 期。

[44] 曾庆生、周波、张程、陈信元：《年报语调与内部人交易：“表里如一”还是“口是心非”？》[J]，《管理世界》2018 年第 9 期。

[45] 张杰、郑文平、翟福昕：《融资约束影响企业资本劳动比吗？——中国的经验证据》[J]，《经济学（季刊）》2016 年第 2 期。

## Research on the Impact of Industrial Policy Correlation on Company Performance

Qin Fei<sup>1</sup> Shen Yan<sup>2</sup>

(1. School of Mathematics and Statistics, Beijing Technology and Business University;

2. National School of Development, Peking University)

**Research Objectives :** This research constructs an indicator to measure the degree of correlation between the corporate and government, and evaluates the impact of policy correlation on corporate performance and the internal mechanism. **Research Methods :** Based on text data from the Chinese government work reports and corporate annual reports from 2009 to 2018, this paper employs data text analysis technology to measure the policy correlation between the corporate annual report and the government work report text, and estimates the impact of policy correlation on corporate performance. **Research Findings :** First, the corporate's policy correlation has increased significantly after 2012. Second, companies with higher policy correlation are characterized with private, small or medium-sized, eastern located, light-asset, low financial leverage, good funding, and riskier. Third, higher policy correlation leads to the better performance of these enterprises, especially for the private, small or medium-sized and eastern enterprises. Fourth, high policy correlation enterprises are less financing constraints and more government-subsidized and tax incentives. Fifth, after deducting government subsidies and tax incentives, there is still a positive relationship between policy correlation and corporate performance, which means that strategic innovation for obtaining policy-related benefits is not the only way for policies to be effective. **Research Innovations :** This paper creatively combines text data analysis technology, and uses a similarity algorithm to measure the degree of policy correlation between annual reports and government work reports and provides a new perspective for quantitative policies. **Research Value :** The method in this paper provides a reference value for the use of text data analysis technology to study economic issues, and the research conclusions provide suggestions for the government to optimize policy implementation methods and improve policy implementation efficiency.

**Key Words :** Policy Correlation; Corporate Performance; Text Data Analysis; Natural Language Processing

**JEL Classification :** G30

(责任编辑: 焦云霞)