

竞争战略，盈余质量与生命周期 ——基于文本分析的实证研究^{*}

胡楠，邱芳娟，梁鹏

摘要：本文首次采用基于 Word2Vec 的自然语言处理方法构建了中国情景下的竞争战略指标，并在此基础上对竞争战略、盈余质量与生命周期之间的关系进行了实证检验。从企业特性上讲，非国有企业、高新技术企业、中小企业和东部企业更青睐于差异化战略。同时，处于生命周期上升期的企业差异化战略实施程度相对较高。进一步研究表明，实施差异化战略的企业盈余持续性较高，盈余波动性较低，且上述关系在成熟期和震荡期的企业中更为显著。基于文本分析方法的竞争战略指标构建在一定程度上解决了传统竞争战略衡量方式存在的偏差和潜在内生性问题，并为竞争战略、盈余质量与生命周期之间的关系提供了新的证据支持。本文为中国情景下竞争战略的更好衡量提供了新的视角，并对目前竞争战略的应用情况以及经济意义进行了探讨和分析，为中国经济政策制定和营商环境建设提供了证据支持，并对企业利益相关者的决策制定有重要的参考意义。

关键词：竞争战略；盈余质量；生命周期；Word2Vec；自然语言处理
中图分类号：C93 文献标识码：A

Competitive Strategy, Earnings Quality, and Life Cycle: Evidence from Large Sample based on Text Analysis

Abstract: To overcome shortcomings of traditional measures of competitive strategy, we construct a textual measure of competitive strategy under the Chinese context, based on a natural language processing technology called Word2Vec. We find that non-state-owned firms, small- and medium-sized firms, and high-tech firms as well as firms in eastern China are more likely to practice differentiation strategy than the remaining half. Moreover, differentiation strategy is also a preferred choice for introduction, growth and mature firms. We further find strong evidence that compared to cost leadership strategy, firms with differentiation strategy have more persistent and less volatile future earnings. And such relation is more pronounced for mature and shake firms. Our text-based measure of competitive strategy provides a new perspective for future research of the competitive strategy in the Chinese context. We also provide evidence on supporting the construction of economic policy and business environment in China.

Keywords: Competitive Strategy; Earnings Quality; Life Cycle; Word2Vec; Natural Language Processing

^{*} 本研究得到国家自然科学基金“企业描述性创新信息披露的影响因素和经济后果：基于上市公司年报文本分析的视角”（基金号：71772150）的资助。

竞争战略，盈余质量与生命周期 ——基于文本分析的实证研究

胡楠，邱芳娟，梁鹏

摘要：本文首次采用基于 Word2Vec 的自然语言处理方法构建了中国情景下的竞争战略指标，并在此基础上对竞争战略、盈余质量与生命周期之间的关系进行了实证检验。从企业特性上讲，非国有企业、高新技术企业、中小企业和东部企业更青睐于差异化战略。同时，处于生命周期上升期的企业差异化战略实施程度相对较高。进一步研究表明，实施差异化战略的企业盈余持续性较高，盈余波动性较低，且上述关系在成熟期和震荡期的企业中更为显著。基于文本分析方法的竞争战略指标构建在一定程度上解决了传统竞争战略衡量方式存在的偏差和潜在内生性问题，并为竞争战略、盈余质量与生命周期之间的关系提供了新的证据支持。本文为中国情景下竞争战略的更好衡量提供了新的视角，并对目前竞争战略的应用情况以及经济意义进行了探讨和分析，为中国经济政策制定和营商环境建设提供了证据支持，并对企业利益相关者的决策制定有重要的参考意义。

关键词：竞争战略；盈余质量；生命周期；Word2Vec；自然语言处理
中图分类号：C93 文献标识码：A

一 引言

战略是企业未来的长远规划。战略决定企业一切行为和决策(Chandler, 1962; Rumelt et al., 1991)。因此，战略问题一直是学术界和产业界的研究热点。自 2003 年开始，中国证监会也逐步规范上市公司的年报披露并加强上市公司对战略相关信息的披露要求，由此可见战略对于企业生存和发展的重要性。Porter(1980)根据企业获得竞争优势的方式将战略分为差异化战略和成本领先战略。该竞争战略因为涵盖范围广泛且界定清晰被视为经典战略分类。随后，国内外学者积极验证竞争战略的有效性，其中探讨竞争战略与企业财务绩效之间关系的文章最为广泛（Hambrick, 1983; Prescott, 1986; Kim, 2004; Zott, 2008; Acquaah, 2008; Claver, 2012; Banker, 2014）。在中国情景下，Li 和 Li（2008）认为成本领先战略能给中国企业带来更好的财务绩效表现。雷辉等（2015）和周兵等（2018）从盈余持续性的角度解释了竞争战略对企业持续性竞争优势的影响，但是结论并不一致。雷辉等（2015）认为实施差异化战略的企业有较强的财务绩效滞后性和持续性，而周兵等（2018）却发现实施成本领先战略的企业在盈余持续性方面表现优秀。其次，现有文献并未考虑企业特性，如生命周期，对竞争战略与盈余持续性之间关系的重要影响。竞争战略的正确选择是企业提高竞争优势和国际影响力有力保障，也是中国实现强国战略的重要基础，因此，进一步研究竞争战略与盈余持续性之间的关系有一定的现实意义。

在现有文献中，竞争战略的衡量主要通过调查问卷，PIMS 数据库和公开披露的财务数据（Dess and Davis, 1984; Hambrick, 1982; Banker et al., 2014），但是以上数据都存在一定的衡量偏差。其中，调查问卷和 PIMS 数据库容易受到被调查人和数据分析者主观评价以及问卷回复率的影响；财务数据是多种因素共同作用的结果，因此数据噪音较大，其次，以财务数据为衡量基础的竞争战略相关研究存在较强的内生性问题（Banker, 2016, 2017）。

随着自然语言处理技术的快速发展，有大量国内外文献研究表明，相比于会计数字信息，文本中蕴含着大量增量信息（王克敏等，2018；姜付秀等，2017；Miller and Skinner 2015；Davis et al., 2012；

Loughran and McDonald, 2011; Li, 2010; Muslu et al., 2015; 薛爽等, 2010; 林乐、谢德仁, 2017; 马黎珺等, 2019; 曾庆生等, 2018; Heinle et al., 2017)。Banker(2016, 2017)首次提出了文本分析方法下的竞争战略衡量, 但是该文本分析方法仍旧沿用传统的词典法, 传统词典法下的竞争战略测量未考虑语境环境, 因此对竞争战略的衡量不够全面、客观。其次, Banker(2016, 2017)是基于美国情景提出的文本分析方法, 考虑到中英文表达的特点差异较大, 该方法不能直接借用来衡量中国情境下的竞争战略。因此, 如何在中国情境下更好的衡量竞争战略是一个亟待解决的问题。

基于以上分析, 本文选取 2003-2016 年中国沪深两市 A 股上市公司年度财务报告作为研究样本, 采用基于 Word2Vec 的自然语言处理方法构建了竞争战略指标。竞争战略指标构建主要通过以下四个步骤: (1) 阅读竞争战略相关文献, 选取描述不同竞争战略特点的种子词; (2) 利用基于 Word2Vec 的自然语言处理方法构建词典, 本文选取相似程度最高的前 10 个相似词以构建本文词集; (3) 计算竞争战略词集的词频和句频; (4) 构建竞争战略指标。在此基础上, 本文对不同性质的公司进行了战略对比, 并进一步分析了竞争战略与盈余质量之间的关系以及生命周期对以上关系的调节作用。研究发现, 相比较于其它企业, 非国有企业, 高新技术企业, 中小型企业 and 东部企业更青睐于差异化战略。从生命周期角度来看, 处于生命周期上升期的企业使用差异化战略占较高比例。同时, 实证结果表明, 相较于实施成本领先战略的企业, 实施差异化战略的企业盈余持续性更高, 盈余波动性更低, 以上结果对处于生命周期成熟期和振荡期的企业影响更为显著。同时, 为解决潜在内生性问题, 本文选取同一行业同一年度且资产规模相近的企业进行匹配, 实证结果稳健。最后, 本文以 2014 年“大众创业, 万众创新”为背景分析了战略转型的实际意义。结果表明, 在“大众创业, 万众创新”的背景下, 企业的战略转型(从成本领先战略转为差异化战略)增强了企业的盈余持续性, 验证了“大众创业, 万众创新”政策的有效性。

本文的创新主要体现在以下 3 个方面: (1) 相较于传统竞争战略衡量方式, 本文基于 Word2Vec 文本分析方法构建的竞争战略指标更加客观, 为中国情境下的战略管理研究提供了新的数据构建参考; (2) 本文对不同性质的企业进行了竞争战略分析。其中, 非国有企业, 高新技术企业, 中小企业和东部企业的差异化战略实施程度更高, 为中国政府的经济政策以及营商环境改进提供了数据支持; (3) 本文验证了竞争战略与盈余持续性之间的关系, 并实证检验了生命周期对以上关系的调节作用。结论丰富了中国战略管理领域研究。

本文的后续安排如下: 第二部分为文献综述与理论分析; 第三部分为研究设计, 重点介绍了竞争战略指标的构建和相关有效性检验, 并阐述了模型设计和其它变量的选择; 第四部分为实证结果分析, 在此基础上进行了内生性处理; 最后结论总结全文。

二 文献综述与理论分析

(一) 竞争战略

战略是企业决策和行为的“指南针”(Chandler, 1962)。好的战略能够帮助企业有效分配资源, 降低经营不确定性。同时, 通过对战略描述的解读, 利益相关者能够建立对公司定位的正确认知。

(Mintzberg, 1987)。因此, 长期以来, 战略是产业界和学术界的研究热点。Porter(1980)在前人文献的基础上提出了竞争战略的概念。该战略认为, 企业生存的前提是拥有竞争优势, 而竞争优势主要来源于对不同竞争战略的选择。根据企业行为的独特性, 竞争战略包括成本领先战略和差异化战略。相较于其它战略分类(Miles, 1978; Mintzberg, 1978), Porter(1980)的战略分类更为清晰, 适用性更为广泛且文献基础较为牢固(Hambrick et al., 1982)。因此, 学术界和产业界都对该战略研究表示出极大的热情。(Dess et al., 1980; Kim et al., 2004; Jusoh and Parnell, 2008; Banker et al., 2014; Li and Li, 2008; Claver et al., 2012)。在相关研究中, 竞争战略的衡量主要通过以下方式: (1) 调查问卷(Dess and Devis, 1984); (2) PIMS 数据库(Hambrick, 1982); (3) 相关财务数据(Banker et al., 2014)。但是由于衡量方式本身的局限性, 竞争战略的衡量存在一定偏差。如调查问卷数据的客观性会受到管理层主观认知和问卷回复率的不利影响。PIMS 数据库虽然是第三方数据库, 但是该数据库的数据来源仍旧是调查问卷。同

时, PIMS 数据库还会进一步受到第三方数据分析人员主观认知的不利影响。财务数据的公开为竞争战略大样本研究提供了数据支持, 但是基于财务数据的竞争战略衡量存在以下问题: (1) 数据噪音较大。财务数据本身是多种因素相互影响的结果, 因此, 是否可以使用财务数据去衡量竞争战略值得商榷; (2) 数据内生性强。在现有竞争战略大样本研究中, 基于财务数据的竞争战略与财务绩效之间的关系本身就存在很强的内生性问题, 实验结果的稳健性因此受到一定的限制。基于此, 客观且有效衡量竞争战略是保证相关研究结论可靠性的重要前提。

机器学习等技术在会计与金融领域的兴起为竞争战略的再衡量提供了新的视角。Banker 首次提出了基于文本分析的竞争战略衡量 (Banker, 2016; Banker and Xinjie, 2017)。但是该衡量方式仍延续使用了传统词典法。基于词典法的文本分析未考虑文本语境环境, 易缺乏全面性, 公开性和客观性 (Loughran and McDonald, 2011)。本文首先对 Banker (2016, 2017) 中提及的文本分析方法进行了优化, 并提出了基于 Word2Vec 的竞争衡量方式, 然后使用基于 Word2Vec 的文本衡量方式对竞争战略与盈余质量之间的关系进行了再验证。相较于传统文本分析方法, 基于 Word2Vec 的自然语言处理方法能根据词汇的上下文内容, 更加综合客观和全面地反映出词汇在语义、句法等方面的特征 (Mikolov et al., 2013; Turney, 2013; Taddy, 2015; Shin et al., 2016)。

(二) 竞争战略与盈余质量

盈余信息披露是企业利益相关者制定决策的重要依据。高质量盈余为不同决策制定者提供了更多的相关信息 (Dechow et al., 2010)。作为盈余质量的两个不同的衡量指标, 盈余持续性与盈余波动性的概念, 影响因素和经济后果被广泛探讨 (Kormendi and Lipe, 1987; Lev and Thiagarajan, 1993; Sloan, 1996; Dichev et al., 2009; 宋建波等, 2012; 方红星、张志平, 2013; 肖华、张国清, 2013; 白天玺、黄志忠, 2015 等)。而竞争战略与盈余质量之间的关系并未得到一致结论。王嵩 (2014) 和周兵等 (2018) 认为差异化战略存在滞后效应, 因此实施差异化战略的企业更有动机实施差异化管理, 从而降低企业的盈余持续性。而 Banker 等 (2014) 和雷辉等 (2015) 则指出, 实施差异化战略企业的盈余波动性较强, 但是考虑到研发投入和产品创新结果的不确定性, 实施差异化战略企业的盈余波动性也相对较强。虽然以上学者对竞争战略与盈余质量之间的关系进行了一定程度的探讨, 但是研究结论并未一致。同时, 以上文章都以财务数据作为竞争战略的衡量指标, 结论存在一定的内生性问题。因此, 本文使用基于 Word2Vec 的文本分析方法对中国企业的竞争战略进行重新衡量并对竞争战略与盈余质量之间的关系进行实证验证。

(三) 竞争战略与生命周期

生命周期是对企业在不同阶段行为改变过程的描述。在不同阶段, 企业的外部环境、资源和组织结构拥有特质性和规律性 (Quinn and Cameron, 1983; Miller and Friesen, 1984; Dickinson, 2011)。企业的治理结构、融资方式和风险承担水平在不同的生命周期阶段都有明显差异 (李云鹤等, 2016; 马宁、王雷, 2018)。Mintzberg (1978) 指出处于不同阶段的企业会选择不同的竞争战略。因此, 企业的生命周期阶段会影响到企业竞争战略的选择, 而竞争战略与生命周期的匹配会增强或者削弱竞争战略对盈余持续性的影响。

(四) 理论分析与假设

本文从资源基础理论和盈余管理理论的角度分析竞争战略与盈余质量之间的关系。

资源基础理论认为企业是各种资源的集合体, 资源的不可模仿性, 异质性和稀缺性决定了企业是否具有持续性竞争优势 (Barney, 1997)。持久的竞争优势才能产生持久的价值创造, 而盈余持续性是持久价值创造的外在表现 (Collins and Kothari, 1989; Asthana and Zhang, 2006)。因此, 拥有不可模仿, 高异质性和高稀缺性资源的企业盈余持续性较高。而差异化战略关注产品创新, 增强自身的品牌辨识度和独特性, 而成本领先战略则主要通过过程创新来提升运营效率和降低成本从而获得竞争优势。随着现代科技的迅速发展, 由过程创新带来的竞争优势会很容易被同行业模仿。因此, 从资源基础理论的角度, 实施差异化战略的企业自身资源不易被模仿且行业门槛提升, 盈余持续性较强。

同时，企业的盈余持续性受到企业盈余管理行为的负面影响。盈余管理是指管理层出于寻租的目的对财务报告信息进行人为干预 (Schipper, 1989)，企业盈余的价值相关性越高，企业进行盈余管理的动机越强。实施差异化战略的企业主要通过不断的产品创新巩固自身优势，前期产品研发费用投入较高，绩效的滞后性较强 (雷辉等, 2015)。Lev 和 Sougiannis (1996) 发现，相较于传统财务报告收益，调整研发支出之后的会计收益价值相关性更强。因此，相较于实施成本领先战略的企业，实施差异化战略的企业的盈余价值相关性较低，因而实施盈余管理的动机较低，企业的盈余持续性较高。另一方面，考虑到盈余持续性与盈余波动性之间的负相关关系。本文提出以下假设：

H1：实施差异化战略企业的盈余持续性更好，盈余波动性更低。

其次，按照生命周期理论，企业在不同生命周期阶段的风险承担水平不同 (马宁、王雷, 2018)。对于处于起步期和成长期的企业来说，企业的融资约束较大 (黄宏斌等, 2016)，风险承担水平较低 (马宁、王雷, 2018)，差异化战略的高收益也因此被高风险效应所抵消；而成熟期和衰退期的企业风险承担水平明显提升，且企业的市场份额较大，因此差异化战略所带来的收益明显更为明显。当企业市场份额下降，利润率降低，企业会选择降低投资，满足现金流需求或者通过高投资进行战略转型，此时差异化战略的盈余持续性影响不明显。基于此，本文提出以下假设：

H2：对于属于成熟期和衰退期的企业，差异化战略的盈余持续性正向影响更为显著；

H3：对于初创期，成长期和衰退期的企业，差异化战略的盈余持续性正向影响不显著；

三 研究设计

(一) 样本和数据来源

本文选取 2003-2016 年中国沪深两市 A 股上市公司年度财务报告作为研究样本。2003-2016 年间共计 29,727 个原始公司年度样本，然后删除：(1) 3,328 个金融行业、ST 以及*ST 公司的样本；(2) 1,926 个 ROA 数据缺失的样本；(3) 1,446 个战略披露数据缺失的样本；(4) 1,022 个控制变量缺失的样本，最终得到 22,005 个行业年度样本数据。

公司竞争战略的文本数据来源于 WinGo 财经文本数据平台，该平台是中国首家财经文本智能研究平台，具有专业的财经领域中文文本数据分析技术。财务数据均来源于 CSMAR 数据库。为避免极端值的影响，本文对所有连续变量进行了上下 1% 的缩尾处理。为克服异方差、序列相关问题等的潜在影响，本文对所有回归都做了行业和年度聚类调整。

(二) 竞争战略的文本度量

本文首次采用“种子词+Word2Vec 相似词扩充”方法对企业竞争战略进行度量。在文本分析领域，词频是对概念进行度量的常见方式 (王雄元等, 2017)。Weber (1990) 指出，词频代表某文档对该词汇的重视程度。因此，本文利用竞争战略词汇总词数在年报文本总词频中的占比来衡量企业的竞争战略偏好。相较于单纯考虑竞争战略词汇的总词数来衡量词频，竞争战略词汇总词数在年报文本总词数的占比能有效避免文档大小对检验结果的影响 (Loughran and McDonald, 2011)。

1. 竞争战略种子词的选取

为构建竞争战略词集，本文首先对竞争战略进行概念梳理，提取描述相关竞争战略特性的词汇。根据 Porter (1980) 的竞争战略理论，成本领先战略的实现主要通过严格的成本和费用控制，引入自动化机器设备提升生产效率等方式，差异化战略则更加重视产品和服务的独特性和差异化。Dess 和 Davis (1984) 将成本领先战略和差异化战略的特点进一步细化和完善。他们认为，企业通过运营效率、竞争性定价、原材料采购和生产过程创新来实现成本和费用的控制，从而实现成本领先战略；而实施差异化战略的企业则更加关注新产品研发，品牌定位，营销技术创新等方面。因此，通过梳理前期经典战略文献 (Porter, 1980; Dess and Davis, 1984; Kotha and Orne, 1989; Devaraj et al., 2004; Banker et al., 2014)，本文对竞争战略的相关词汇进行总结，然后利用谷歌翻译、有道词典以及金山词霸等三个权威翻译软件进行翻译，

得到每个英文词汇的多个中文汉语翻译结果。在充分考虑中文的特征属性以及竞争战略的内在含义的情况下,对翻译结果进行了初步筛检:(1)删除词义明显不符合竞争战略内涵的词汇。如,在词汇翻译结果中,成本领先战略下的“efficient”的翻译结果之一为“有效”,但是“有效”更倾向于中性词,英文环境下的“efficient”在差异化战略中更倾向于“高效”,强调“效率”,因此,本文保留“efficient”的翻译结果“高效”和“效率”。(2)通过文本阅读对语义模糊的词汇进行进一步语义确认。如,差异化战略下的“innovation in marketing technique”一词可翻译为“销售技巧”和“营销创新”。“销售技巧”的提升在一定程度上可以代表企业对于品牌的重视,但是也可能理解为中性词。为此,本文对包含“销售技巧”的词汇的所有年报进行阅读,最终发现,提到“销售技巧”的文档主要涉及人员简介和员工培训,并不能有效反映企业的竞争战略状态,因此,本文将“销售技巧”从竞争战略翻译结果词汇中删除,保留翻译结果“营销创新”。

在经过上述步骤之后,为避免竞争战略词汇翻译的主观性,本文对相关竞争战略词汇初步结果进行交叉验证,对词汇进行进一步筛检和补充。最终,本文将修订后的战略词集继续发至2位战略学术和业界专家进行审验和反馈,最终形成36个基于文献的竞争战略词汇种子词集,其中,21个差异化战略词汇,15个成本领先战略词汇。

2. 相似词扩充

在事物描述过程中,同一特征往往会使用语义相似的不同词汇集合。因此,在确定竞争战略的种子词词集之后,本文对种子词词集进行了相似词扩充。目前的学术研究中,扩充词集的方法主要有:(1)词典法(Price et al., 2012; 肖华、张国清, 2013; Davis et al., 2015);(2)有监督的机器学习法(Feng, 2010; Huang and Li, 2011)。以上两种传统方法存在以下缺陷:首先,词典法往往会忽略所研究的特定文本语境,找到的词汇不能全面、准确、客观地衡量出对应的文本特征(Loughran and McDonald, 2011);其次,在大样本研究环境下,有监督的机器学习法需要手工编码,手工编码的效率较低,另一方面人工判别存在较强的主观性,会导致最终词集存在偏误的情况(Zhang et al., 2015)。本文使用了无监督的基于Word2Vec的统计语言模型对竞争战略种子词进行相似词扩充。Word2Vec模型属于神经网络语言模型,由Mikolov等(2013)提出,并在各种自然语言处理任务中取得了相当大的成功(Mikolov et al., 2013; Turney, 2013; Taddy, 2015; Shin et al., 2016)。它根据上下文内容将词语表示成密集的实值向量,这些向量一起形成了词向量空间,空间上的词向量之间的距离度量即为词语之间的语义相似性。距离越近,相似度越大,则这两个词的词义越相近、句法越相似。其中,本文采用的是使用最广泛的余弦距离进行语义相似性的测量。Word2vec模型不仅克服了稀疏向量表征造成的维数灾难;同时,也很好的刻画了词与词之间语义相似性,克服了词汇鸿沟。本文选取每个战略种子词相似度最高的前10个词作为该词的相似词词集初步结果,然后删除掉相似词词集中的重复词汇,如成本领先战略词集中,“成本管控”既是“成本管理”的相似词,又是“成本控制”的相似词,因此,在成本领先战略的最终词集结果中只保留一次即可。此外,本文进一步对财务报告原文进行抽样对比,发现竞争战略词集与财务报告中竞争战略描述内容具有很强的一致性。最终,本文基于Word2Vec算法得到的竞争战略词集最终包含290个词汇,其中差异化战略词汇165个,成本领先战略词汇125个。

通过神经网络语言模型训练得到的相似词推荐结果,克服了传统人工方式的缺陷,能综合、客观地反映出词语在语义、句法等方面的特征。此外,Word2Vec技术在训练词汇方面具有高效性并且在英文语料(Mikolov et al., 2013)和中文语料(Zhang et al., 2015)处理过程中都表现优异。在本文的研究中,以成本领先战略中的“成本管理”和差异化战略中的“品牌建设”为例,“成本管理”基于Word2Vec模型选取出的相似度前10的相似词分别为“成本控制”、“成本管控”、“预算管理”、“全面预算管理”、“精细化管理”、“目标成本”、“费用控制”、“降本增效”、“过程控制”、“精益管理”,相似词与种子词的相似度波动区间为0.642~0.854;“品牌建设”的相似词为“品牌宣传”、“渠道建设”、“品牌”、“品牌战略”、“品牌管理”“产品创新”、“企业品牌”、“营销创新”、“市场营销”、“品牌营销”,相似度波动区间为0.612~0.738。

（三）竞争战略文本度量的有效性

鉴于本文首次使用基于 Word2Vec 文本分析方法对企业的竞争战略状态进行衡量，本文对其进行了有效性检验。Scholtes (Scholtes et al., 2011) 指出衡量工具的有效性包括内容效度，结构效度和效标效度。因此，本文从以上三个角度对企业竞争战略的文本度量的有效性进行检验。

1. 内容效度

内容效度具体指通过基于 Word2Vec 的文本分析技术所提取的竞争战略词集的全面性以及与企业竞争战略偏好的相关性。词集构建是文本分析方法的基础。为检验竞争战略词集的有效性，本文采取了专家验证法和抽样法。首先，本文邀请 2 位学术界和业界专家对竞争战略词集的有效性进行分析，分析结果肯定了竞争战略词集对企业竞争战略描述的相关性和全面性。其次，本文分别抽取了成本领先战略词频和差异化战略词频最高的前 5 家公司的年度财务报告，并认真研读其公司财报和公司官网等渠道所披露的竞争战略信息，将公司在其他渠道树立的形象和文本分析方法下确定的竞争战略定位进行比较，结果发现本文所提取的关键词能够准确衡量企业的竞争战略定位，利用该文本分析方法所确定的竞争战略侧重和公司的对外展示形象保持一致。此外，战略通常被描述为企业的长远计划，因此为进一步验证竞争战略的长期性，本文将每个战略的词频说占比划分为三分位，然后取位于上三分位的样本进行检验。如果文本度量下的竞争战略是长期的，则每年处于第三分位的公司数目不会发生变化。根据表 1，在随后的 3 年中，57.5% 的差异化战略企业和 58.72% 的成本领先战略企业仍处于第三分位，由此验证了竞争战略的文本度量体现了战略的长期性，也支撑了文本度量方式的内容效度。

表 1 竞争战略的长期性检验

竞争战略	Q1	Q2	Q3	N
Panel A: 差异化战略				
Year0			100.00	685
Year1	1.83	21.16	77.02	657
Year2	4.92	27.23	67.85	650
Year3	8.43	28.93	62.64	629
Year4	9.43	33.08	57.50	647
Panel B: 成本领先战略				
Year0			100.00	685
Year1	1.06	22.93	76.02	663
Year2	4.53	28.44	67.03	640
Year3	5.85	30.96	63.19	633
Year4	9.97	31.31	58.72	642

2. 结构效度与效标效度

竞争战略的结构效度主要观测基于 Word2Vec 的竞争战略衡量方式是否能够有效预测假设。为此，本文通过对现有经典文献的复制来分析竞争战略文本度量的结构效度。我们的文献复制结果与历史文献保持一致，从而证明了基于 Word2Vec 的竞争战略衡量方式的结构效度较为稳健。

效标效度主要观察不同工具之间的相关性。在本文，效标效度测量竞争战略文本度量方式与传统竞争战略衡量方式之间的关系。为此，本文选取总资产周转率、和毛利率分别作为衡量企业成本领先战略和差异化战略的传统指标 (Balsam et al., 2011)，变量定义见表 4。表 2、3 为文本分析法下的竞争战略指标与传统竞争战略财务指标相关性检验结果。相关性结果显示，本文所使用的竞争战略衡量指标与传统的竞争战略财务指标显著正相关。其中，成本领先战略与总资产周转率相关系数为 0.128；差异化战略与毛利率相关系数为 0.189，且都在 1% 水平上统计显著。该结果进一步验证了本文使用的基于深度学习的竞争战略衡量工具的效标效度。

综上所述，基于 Word2Vec 的竞争战略衡量方式在内容效度、结构效度和效标效度方面均具有显著的有效性，为后续的实证研究提供了有效的数据和方法支撑。

表 2 成本领先战略效标效度检验

	COST	RAR
<i>COST</i>		0.204***
<i>RAR</i>	0.128***	

注：***，**，*分别表示 1%、5%、10%的显著水平，下同

表 3 差异化战略效标效度检验

	DIFF	GPR
<i>DIFF</i>		0.187***
<i>GPR</i>	0.189***	

(四) 模型设计和其他变量

为检验竞争战略与企业盈余持续性之间关系，本文采用 Sloan(1996)模型进行验证：

$$E_{i,t+j} = \alpha + \beta_1 \times E_{i,t} + \beta_2 \times Strategy_{i,t} \times E_{i,t} + \beta_3 \times Strategy_{i,t} + \beta_4 \times Controls_{i,t} + \sum YearDummies + \sum IndustryDummies + \varepsilon_{i,t} \quad (1)$$

模型(1)采用的是多元回归分析，其中， $E_{i,t+j}$ 为目标企业*i*在*t+j*年的盈余，*j*的取值为1~4， $E_{i,t}$ 为企业*i*在*t*年的盈余。参照方红星等(2013)和肖华等(2013)，本文采用ROA作为企业盈余的主要代理变量。自变量 $Strategy_{i,t}$ 是竞争战略的衡量变量。该变量为哑变量。当企业*i*在*t*年的差异化战略与成本领先战略之比大于当年总样本取值中值(差异化战略/成本领先战略)，则该变量为1，即企业在当年实施的是差异化战略；反之该变量为0，即企业在当年实施的是成本领先战略。此外，参照前人研究，本文在回归分析中控制了公司规模(*SIZE*)、资产负债率(*LEVERAGE*)、公司成长性(*GROWTH*)、第一大股东持股比例(*PERCENT*)、国有股股东持股比例(*NATIONAL*)、高管持股比例(*MANAGE*)、公司CEO和董事长是否两职合一(*CEO*)、公司上市年限(*AGE*)、股票年度回报率(*RETURN*)、股票收益波动率(*VOLATILITY*)、审计师质量(*BIG4*)、公司会计年度内是否亏损(*LOSS*)等因素的影响，并控制了年份、行业固定效应。变量的具体定义和说明见下表4。

在检验竞争战略的盈余波动性时，本文参考先前文献(Cheng, 2008; 卢闯等, 2011; 孙健等, 2016)，采用模型(2)验证竞争战略与盈余波动性之间的关系：

$$EVOL_{i,t} = \alpha + \gamma_1 \times Strategy_{i,t} + \gamma_2 \times ASIZE_{i,t} + \gamma_3 \times AROA_{i,t} + \gamma_4 \times AROA_{i,t-1} + \gamma_5 \times AROA_{i,t-2} + \gamma_6 \times ALEVERAGE_{i,t} + \gamma_7 \times AGROWTH_{i,t} + \gamma_8 \times AAGE_{i,t} + \sum YearDummies + \sum IndustryDummies + \varepsilon_{i,t} \quad (2)$$

在模型(2)中， $EVOL_{i,t}$ 变量为企业*i*在*t*年时每个5年的ROA的标准差，其他变量均为每个变量5年的平均值。

表 4 主要变量定义与说明

变量	含义	定义
战略文本指标		
<i>NCOST</i>	成本领先战略总词频	财务报告中成本领先战略相关词集的总词频
<i>NDIFF</i>	差异化战略总词频	财务报告中差异化战略相关词集的总词频
<i>COST</i>	成本领先战略词频占比	年报平均每100字出现成本领先战略相关词集的总次数
<i>DIFF</i>	差异化战略词频占比	年报平均每100字出现差异化战略相关词集的总次数

<i>COST_SEN</i>	成本领先战略句频占比	财务报告中成本领先战略相关词集的总句频占年报总句频的比例
<i>DIFF_SEN</i>	差异化战略句频占比	财务报告中差异化战略相关词集的总句频占年报总句频的比例
<i>RATIO</i>	竞争战略比	差异化战略词频占比/成本领先战略词频占比
<i>Strategy</i>	竞争战略类型	$RATIO_{i,t} > RATIO_{i,t-1}$ 当年中值, 取值为1; 否则为0
因变量		
<i>ROA</i>	总资产收益率	净利润/期末期初资产均值*100%
<i>EVOL</i>	盈余波动性	每个5年中样本公司ROA的标准差
控制变量		
<i>SIZE</i>	公司规模	总资产的自然对数值
<i>LEVERAGE</i>	资产负债率	总负债/总资产*100%
<i>GROWTH</i>	公司成长性	(本期、上期主营业务收入之差/上期主营业务收入*100%)
<i>PERCENT</i>	第一大股东持股比例	第一大股东所持股份数占总股数的百分比
<i>NATIONAL</i>	国有股股东持股比例	国有股股东所持股份数占总股数的百分比
<i>MANAGE</i>	高管持股比例	高管所持股份数占总股数的百分比
<i>CEO</i>	两职合一	CEO和董事长由不同的人担任, 则取值为1; 否则为0
<i>LN(AGE+1)</i>	上市年限	Ln(公司上市的年数+1)
<i>RETURN</i>	股票年度回报率	考虑现金红利再投资的年个股回报率
<i>VOLATILITY</i>	股票收益波动率	十二个月的考虑现金红利再投资的月个股回报率的标准差
<i>BIG4</i>	审计师质量	公司审计师属于国际四大则取值为1, 否则为0
<i>LOSS</i>	当年是否亏损	当期净利润为负, 取值为1, 否则为0
<i>GPR</i>	毛利率	(营业收入-营业成本)/营业收入*100%
<i>RAR</i>	总资产周转率	营业收入/期末期初资产均值*100%

四 实证结果与分析

(一) 描述性统计

表5为描述性统计结果, 通过描述性统计可以看出, 成本领先战略词频明显高于差异化战略词频, 该数据表明了企业运营对成本的重视。成本构成是企业考虑竞争战略时的普遍考虑方向, 大多数企业通过成本优化手段来提高运营效率, 最终实现企业收益最大化。然而, 从竞争战略词汇词频最小值和最大值的取值跨度来看, 差异化战略词汇占比的波动较高, 说明实施差异化战略的企业特性有所差异, 也进一步验证了差异化战略更强调企业或者产品自身的独特性。其次, 本文使用了竞争战略句频占比进行了稳健性检验, 成本领先战略和差异化战略的句频占比平均值分别为0.075和0.055。

表5 描述性统计结果

Variable	N	Mean	Median	Std. Dev	Min	Max
<i>NCOST</i>	22,005	75.804	71.000	33.318	15.000	183.000
<i>NDIFF</i>	22,005	59.289	50.000	41.653	6.000	200.000
<i>COST</i>	22,005	0.394	0.381	0.127	0.142	0.754
<i>DIFF</i>	22,005	0.293	0.262	0.164	0.048	0.798
<i>COST_SEN</i>	22,005	0.075	0.071	0.027	0.027	0.158
<i>DIFF_SEN</i>	22,005	0.055	0.048	0.032	0.009	0.162
<i>ROA</i>	22,005	0.039	0.036	0.061	-0.230	0.221

<i>SIZE</i>	22,005	21.847	21.698	1.242	18.951	25.549
<i>LEVERAGE</i>	22,005	0.467	0.471	0.215	0.050	0.999
<i>GROWTH</i>	22,005	0.219	0.125	0.556	-0.680	3.770
<i>PERCENT</i>	22,005	36.630	34.550	15.522	9.080	75.100
<i>NATIONAL</i>	22,005	0.125	0.000	0.211	0.000	0.729
<i>MANAGE</i>	22,005	0.058	0.000	0.139	0.000	0.614
<i>CEO</i>	22,005	0.016	0.000	0.127	0.000	1.000
<i>LN(AGE+1)</i>	22,005	2.110	2.197	0.684	0.000	3.135
<i>RETURN</i>	22,005	0.305	0.074	0.783	-0.710	3.344
<i>VOLATILITY</i>	22,005	0.141	0.127	0.065	0.042	0.384
<i>BIG4</i>	22,005	0.061	0.000	0.239	0.000	1.000
<i>LOSS</i>	22,005	0.097	0.000	0.296	0.000	1.000

（二）竞争战略年度与行业趋势分析

为研究竞争战略的发展趋势，本文对成本领先战略词频和差异化战略词频的时间分布和行业分布分别进行了统计。图 3 为相关竞争战略词频的时间趋势。整体而言，随着时间的推进，成本领先战略和差异化战略的词频都处于上升趋势，这与证监会对企业的披露要求不断严格有关。同时，企业的相关竞争战略词汇词频也出现了大幅度的增加。此外，图 3 也揭示了外界环境变化对竞争战略方向的影响。如企业的成本领先战略词频在 2008 出现了峰值。在全球金融危机时期，企业将更多的注意力投入到节省成本以求继续生存，随着金融危机的缓解，成本领先战略词频开始下滑，差异化战略词频开始上升。2014 年，国家大力实施“万众创新”等创新激励措施，企业的差异化战略词频大幅度上升，差异化战略词频占比和成本领先战略词频占比之间的差距逐渐缩小，到 2016 年，两者之间的词频占比基本持平，说明企业对创新的重视程度增加。图 4 为竞争战略词频的行业分布。数据表明，制造业更强调其成本优势，而住宿餐饮业、信息传输、软件和信息技术服务业和文化、体育、娱乐等行业则更注重差异化。该趋势与行业特性相符合。在我国，制造业的进入门槛比较低，大多数企业通过价格优势获得市场份额，而通信技术等行业更迭速度快，对企业的科研技术能力要求比较高，因此更注重自身的差异性。

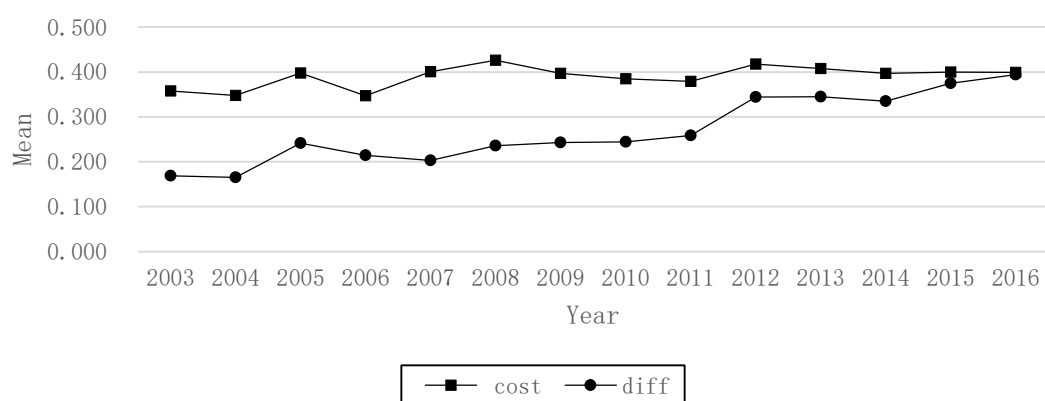


图 1 竞争战略词频的时间趋势分析

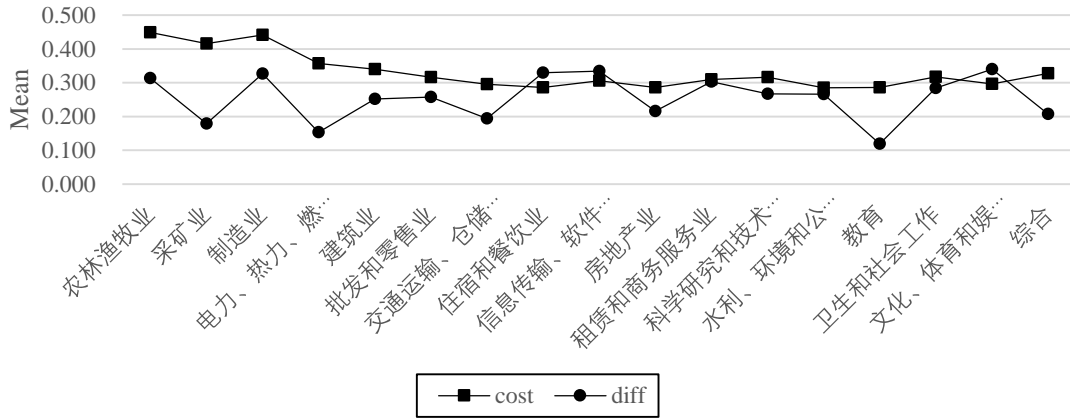


图 2 竞争战略词频的行业分布分析

(三) 竞争战略与企业特征分析

本文将不同特征的企业进行分组，并对比了不同特征企业的竞争战略实施情况，具体情况如表 6 所示。从表中可以看出，相较于国有企业，非国有企业的差异化战略实施情况较高。的市场资源较少，因此，为了获得竞争优势，非国有企业更倾向于使用差异化战略。该现象吻合我国先前调查数据，即民营企业贡献了 70% 以上的技术创新。从企业资产规模来看，中小型企业差异化程度较高。大型企业的市场份额和利润率较为稳定，因此，对于大型企业来说，企业的投资更为保守，企业以保持现状为目标。而中小型企业的冒险精神较高，希望以差异化和创新提高市场份额。同时，由于东部企业的营商环境较高，私有财产保护意识更为强烈，因此，相比于西部企业，东部企业的差异化战略实施程度较高。最后，因为行业属性问题，高新技术企业比其它企业的差异化程度更高。以上分析符合现实企业的基本状况，也进一步验证了所构建的竞争指标的有效性。

表 6 竞争战略与企业特征

Panel A: 国有企业与非国有企业						
国有企业		非国有企业		差异检验		
N	Mean	N	Mean	Mean	t	
11287	0.438	10717	0.564	-0.126***	-18.85	
Panel B: 高新技术产业和其他企业						
高新技术企业		其他企业		差异检验		
N	Mean	N	Mean	Mean	t	
1161	0.770	20844	0.485	0.285***	19.09	
Panel C: 大型企业和中小型企业						
大型企业		中小型企业		差异检验		
N	Mean	N	Mean	Mean	t	
11004	0.479	11001	0.520	-0.041***	-6.07	
Panel D: 东部企业与西部企业						
东部企业		西部企业		差异检验		
N	Mean	N	Mean	Mean	t	
13745	0.523	3263	0.397	0.126***	13.03	

（四）竞争战略与生命周期

企业的发展遵从着不可逆规律，而企业生命周期的划分则体现了企业不同阶段的内外部运营环境特点。为此，本文参考 Dickinson (2011) 的生命周期划分方式，将企业的生命周期划分为起步期，成长期，成熟期，震荡期和衰退期，并进行了描述性分析以及分组实证以检验生命周期对企业竞争战略与盈余质量之间的关系的作用。如图所示，当企业处于起步期，成长期，成熟期时，企业更倾向于使用差异化战略，而当企业处于震荡期或者衰退期时，企业则希望通过低成本战略重新挽回市场。

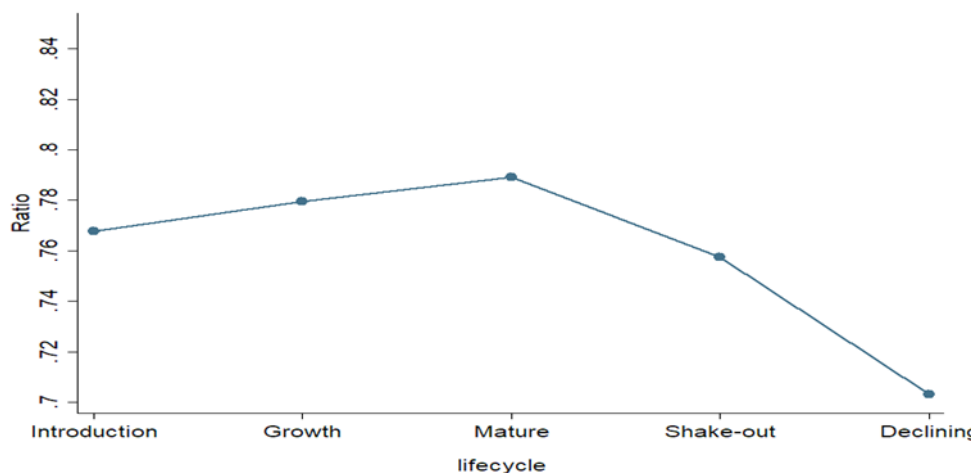


图3 不同生命周期的竞争战略应用

（五）实证结果分析

1. 竞争战略与盈余持续性

为了检验竞争战略与企业盈余持续性之间的关系，本文对模型（1）进行了多元回归分析检验，具体结果见表7。从表中可以看出， $(t+1) - (t+4)$ 年竞争战略变量 Strategy 与盈余变量 ROA 之间交乘项系数显著为正，验证了差异化战略对企业盈余持续性有正向影响的假设。即实施差异化战略企业的盈余持续性更强。

表7 竞争战略与盈余持续性

	(1) ROA(t+1)	(2) ROA(t+2)	(3) ROA(t+3)	(4) ROA(t+4)
ROA(t)	0.509*** (22.46)	0.374*** (15.39)	0.308*** (13.72)	0.234*** (10.46)
STRATEGY(t) *ROA(t)	0.159*** (6.92)	0.179*** (8.62)	0.157*** (6.63)	0.154*** (7.27)
STRATEGY(t)	-0.002 (-1.33)	-0.002** (-2.16)	-0.001 (-0.45)	0.001 (0.49)
SIZE(t)	-0.001 (-1.02)	-0.003*** (-4.19)	-0.004*** (-6.12)	-0.005*** (-7.49)
LEVERAGE(t)	-0.017*** (-4.12)	-0.015*** (-3.82)	-0.011*** (-2.66)	-0.009** (-2.13)
GROWTH(t)	0.001 (0.68)	-0.002* (-1.79)	-0.001 (-1.00)	-0.002* (-1.80)
PERCENT(t)	0.000*** (4.37)	0.000*** (4.56)	0.000*** (4.33)	0.000*** (3.71)
NATIONAL(t)	0.001	-0.003	-0.006**	-0.009***

	(0.22)	(-1.01)	(-2.02)	(-3.47)
MANAGE(t)	0.006	0.008*	0.006	0.002
	(1.51)	(1.80)	(0.98)	(0.22)
CEO(t)	-0.003	-0.001	-0.000	-0.001
	(-1.17)	(-0.36)	(-0.06)	(-0.14)
AGE(t)	-0.001*	-0.000	0.000	0.001
	(-1.86)	(-0.59)	(0.27)	(0.58)
RETURN(t)	0.012***	0.008***	0.006***	0.007***
	(5.61)	(3.82)	(2.69)	(3.20)
VOLATILITY(t)	-0.065***	-0.071***	-0.083***	-0.092***
	(-4.79)	(-4.78)	(-6.29)	(-6.23)
BIG4(t)	0.007***	0.009***	0.008***	0.010***
	(5.97)	(7.21)	(4.80)	(5.17)
LOSS(t)	0.032***	0.028***	0.025***	0.024***
	(14.84)	(11.72)	(9.82)	(9.99)
CONSTANT	0.029**	0.065***	0.108***	0.139***
	(2.13)	(5.39)	(8.56)	(10.90)
Year Fixed Effects	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES
Observations	19338	16890	14564	12284
Adj-R ²	0.345	0.223	0.160	0.118

2. 内生性处理

此外，考虑到竞争战略与企业盈余之间可能存在的互为因果关系，本文进一步对样本进行匹配，匹配原则为：（1）匹配企业来自于同一样本同一行业；（2）互为匹配企业的资产规模最为接近。根据匹配结果，本文对变量 $treat_t$ 赋值。当被匹配企业当年实施的是差异化战略而匹配企业当年实施的是成本领先战略，则哑变量 $treat_t$ 的取值为 1；其他情况下，哑变量 $treat_t$ 的取值为 0。具体的实证结果见表 8。表 8 结果显示，哑变量 $treat_t$ 与企业盈余变量 $ROA_{i,t}$ 之间的交乘项显著为正，进一步验证了假设（1），即实施差异化战略的企业的盈余持续性更高。

表 8 竞争战略与盈余持续性-基于匹配样本

	(1)	(2)	(3)	(4)
	ROA(t+1)	ROA(t+2)	ROA(t+3)	ROA(t+4)
ROA(t)	0.563***	0.436***	0.366***	0.290***
	(27.41)	(18.45)	(15.46)	(12.26)
TREAT(t) *ROA(t)	0.103***	0.106***	0.075***	0.074***
	(3.97)	(3.41)	(2.81)	(3.16)
TREAT(t)	-0.002	-0.002	-0.001	0.002
	(-1.15)	(-1.08)	(-0.30)	(1.20)
Controls	YES	YES	YES	YES
Constant	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES
Observations	19338	16890	14564	12284
Adj-R ²	0.340	0.216	0.153	0.111

其次，本文以“大众创业，万众创新”为背景，探讨企业竞争战略的转变（成本领先战略转为差异化战略）是否会提升企业的盈余持续性。李克强总理于 2014 年 9 月在夏季达沃斯论坛提出“大众创业，万众创新”的理念。该政策旨在通过创新驱动国家产业升级，实现国家的强国战略。在这一政策引导下，中小企业积极响应，纷纷进行战略转型。为探究战略转型（成本领先战略转为差异化战略）对企业盈余持续性的影响，本文选取 2013 年实施成本领先战略的企业，并对该企业的战略转型进行度量。即 2014 年以后，企业由成本领先战略转为差异化战略，变量 change 赋值为 1，否则，变量 change 赋值为 0。具体实证结果如表 9 所示。表 9 第（1）列包括去除 2014 年的全样本研究；第（2）列包括 2014 年前后各两年的样本数据。表 9 结果表明，当企业从成本领先战略转为差异化战略时，企业的盈余持续性显著增强。

表 9 竞争战略变化与盈余持续性

	(1)	(2)
	ROA(t+1)	ROA(t+1)
ROA(t)	0.003*	0.032*
	(1.66)	(1.77)
CHANGE(t) *POST2014(t) *ROA(t)	0.159***	0.142***
	(3.89)	(3.31)
Controls	YES	YES
Constant	YES	YES
Year Fixed Effects	YES	YES
Industry Fixed Effects	YES	YES
Observations	9,470	3,261
Adj-R ²	0.153	0.167

3. 竞争战略与盈余波动性

为验证竞争战略对盈余波动性的影响，本文对模型（2）进行了回归检验分析。从表 10 中可以得出，竞争战略与盈余波动性呈负显著关系。即相较于采用成本领先战略的企业，采用差异化战略企业的盈余波动性较低。该结果验证了先前文献（Graham et al., 2005; Dichevet al., 2009; Frankelet al., 2009; 白天玺、黄志忠, 2015）的研究结论。但是雷辉（2015）也提出，企业实施差异化战略后其绩效水平的滞后期比成本领先战略的滞后期更长。因此，虽然差异化战略的盈余质量较好，但是对于追求较短滞后期的企业来说，成本领先战略是更合适的竞争战略选择。从而解释了为何成本领先战略产生的盈余持续性较短，波动性较强但是仍然是部分企业的战略选择。

表 10 竞争战略与盈余波动性

	(1)
	EVOL
STRATEGY(t)	-0.005***
	(-5.36)
ASIZE(t)	-0.012***
	(-19.51)
AROA(t)	-0.089**
	(-2.48)
AROA(t-1)	0.058**
	(2.53)
AROA(t-2)	-0.007*

	(-1.74)
ALEVERAGE (t)	0.057*** (9.93)
AGROWTH (t)	0.011*** (8.66)
AAGE (t)	0.000*** (5.12)
CONSTANT	0.265*** (24.62)
Year Fixed Effects	YES
Industry Fixed Effects	YES
Observations	11,885
Adj-R ²	0.194

4. 不同生命周期下的竞争战略盈余持续性表现

表 11 结果显示，对处于成熟期-震荡期的企业来说，差异化战略选择对盈余持续性有正向显著影响。可能的解释是处于成熟期的企业本身已有的盈余持续性已经表现优异。而对于起步期，成长期和震荡的企业来说，企业的差异化竞争战略选择会帮助企业降低市场的进入门槛或者扭转局面，但是由于起步期和成长期企业的风险承受能力较低，所以竞争战略对盈余持续性的影响相对较弱。而处于衰退期企业的盈余持续性则不受到竞争战略选择的影响，可能是因为处于衰退期企业的整体经营较为混乱或者进入清算阶段，因此，竞争战略的选择对于企业来说并没有显著性影响。因此，与当前企业的普遍选择不同的是，差异化战略对带给衰退期企业更为显著的正向影响。因此，处于衰退期的企业更应该积极创新而不是低成本竞争。

表 11 不同生命周期的竞争战略与盈余持续性

	Dependent Variable=ROA (t+1)				
	Introduction	Growth	Mature	Shake	Declining
STRATEGY (t)*					
ROA (t)	0.054 (0.85)	0.013 (0.33)	0.127*** (4.07)	0.112** (2.32)	0.032 (0.42)
STRATEGY (t)	0.000 (0.07)	0.003 (1.15)	-0.001 (-0.44)	0.001 (0.38)	0.000 (0.10)
Controls	YES	YES	YES	YES	YES
Constant	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES
Observations	2516	6191	6973	2542	1116
Adj-R ²	0.244	0.359	0.454	0.239	0.176

（六）稳健性检验

为进一步检验模型的稳健性，本文对模型（1）和（2）的相关变量进行了多维度测量。（1）聚焦于财务报告中的经营情况讨论与分析部分的竞争战略描述；（2）参照前人研究，使用净资产收益率（ROE）作为企业盈余的替代衡量指标；（3）考虑到竞争战略是企业的长期行为，利用相关竞争战略词汇词频的累计平均值来衡量企业的竞争战略水平；（4）采用句频作为竞争战略文本的衡量方式。实证结果依然稳健；（5）控制了传统的财务数据。限于篇幅，此处不报告稳健性回归结果表格。

五 研究结论

竞争战略是企业获得竞争优势的重要手段，但是目前的竞争战略衡量方式存在一定的偏差和潜在内生性问题。基于此，本文选取 2003-2016 年中国沪深两市 A 股上市公司年度财务报告作为研究样本，利用基于 Word2Vec 的文本分析方法构建了竞争战略指标。在此基础上，本文分析了竞争战略在不同特征企业或者不同生命周期的应用。研究发现，非国有企业，高科技企业，中小型企业 and 东部企业更青睐于差异化战略。本文进一步对竞争战略与盈余质量之间的关系以及生命周期对上述关系的调节作用进行了实证检验。研究表明，实施差异化战略企业的盈余持续性相对较高，盈余波动性较小，且上述关系对于成熟期和衰退期的企业影响更为显著。

本文在理论和实践方面都有一定意义。在理论方面，本文在中文情境下构建了基于文本分析的竞争战略指标，一定程度克服了现有竞争战略衡量方式所带来的偏差和潜在内生性问题，为中国战略管理领域研究提供了新的衡量视角。其次，本文实证检验了竞争战略、盈余质量与生命周期之间的关系，丰富了竞争战略的相关研究。在实践方面，本文实证验证了民营企业 and 中小企业企业对科技创新的贡献，以及营商环境较好的东部沿海企业更有动力去创新。因此，本文启示政府应加快私有产权保护立法以及营商环境优化，进一步激发企业的创新潜能。同时，本文进一步验证了“大众创新，万众创业”政策的有效性并为相关者决策提供了重要参考。

参考文献:

- (1) 白天玺、黄志忠:《盈余波动、盈余持续性与盈余公告后漂移》,《当代会计评论》,2015年第1期。
- (2) 方红星、张志平:《内部控制对盈余持续性的影响及其市场反应——来自A股非金融类上市公司的经验证据》,《管理评论》,2013年第12期。
- (3) 黄宏斌、翟淑萍、陈静楠:《企业生命周期、融资方式与融资约束——基于投资者情绪调节效应的研究》,《金融研究》,2016年第7期。
- (4) 姜付秀、王运通、田园、吴恺:《多个大股东与企业融资约束——基于文本分析的经验证据》,《管理世界》,2017年第12期。
- (5) 雷辉、王亚男、聂珊珊、欧阳丽萍:《基于财务绩效综合指数的竞争战略绩效时滞效应研究》,《会计研究》,2015年第5期。
- (6) 李云鹤、李湛、唐松莲:《企业生命周期、公司治理与公司资本配置效率》,《南开管理评论》,2011年第3期。
- (7) 林乐、谢德仁:《分析师荐股更新利用管理层语调吗?——基于业绩说明会的文本分析》,《管理世界》,2017年第11期。
- (8) 刘睿智、胥朝阳:《竞争战略、企业绩效与持续竞争优势——来自中国上市公司的经验证据》,《科研管理》,2008年第6期。
- (9) 卢闯、刘俊勇、孙健、杨棉之:《控股股东掏空动机与多元化的盈余波动效应》,《南开管理评论》,2011年第5期。
- (10) 马宁、王雷:《企业生命周期、竞争战略与风险承担》,《当代财经》,2018年第5期。
- (11) 马黎珺、伊志宏、张澈:《廉价交谈还是言之有据?——分析师报告文本的信息含量研究》,《管理世界》,2019年第7期。
- (12) 宋建波、高升好、关馨姣:《机构投资者持股能提高上市公司盈余持续性吗?——基于中国A股上市公司的经验证据》,《中国软科学》,2012年第2期。
- (13) 孙健、文雯、袁蓉丽、石玉凝:《上市公司委托理财与盈余波动性》,《中国软科学》,2016年第6期。
- (14) 孙谦:《盈余持续性研究综述及启示》,《厦门大学学报(哲学社会科学版)》,2010年第1期。
- (15) 孙世攀、徐霞、汪冬梅:《盈余持续性与公司治理实证研究——来自沪深上市公司的经验证据》,《科技与管理》,2011年第1期。
- (16) 王雄元、李岩琼、肖恣:《年报风险信息披露有助于提高分析师预测准确度吗?》,《会计研究》,2017年第10期。
- (17) 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》,《管理世界》,2018年第12期。
- (18) 肖华、张国清:《内部控制质量、盈余持续性与公司价值》,《会计研究》,2013年第5期。
- (19) 薛爽、肖泽忠、潘妙丽:《管理层讨论与分析是否提供了有用信息?——基于亏损上市公司的实证探索》,《管理世界》,2010年第5期。
- (20) 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”?》,《管理世界》,2018年第9期。
- (21) 周兵、黄芳、任政亮:《企业竞争战略与盈余持续性》,《中国软科学》,2018年第3期。
- (22) Acquaaah, M. and M. Yasai-Ardekani, 2008, "Does the implementation of a combination competitive strategy yield incremental performance benefits? A new perspective from a transition economy in Sub-Saharan Africa.", *Journal of Business Research*, 61(4), pp. 346-354.
- (23) Asdemir, O., G. D. Fernando and A. Tripathy, 2013, "Market perception of firm strategy.", *Managerial Finance*, 39(2), pp. 90-115.
- (24) Asthana, S. C. and Y. Zhang, 2006, "Effect of R&D investments on persistence of abnormal earnings.", *Review of Accounting & Finance*, 5, pp.124-139.

- (25) Balsam, S., G. D. Fernando and A. Tripathy, 2011, "The impact of firm strategy on performance measures used in executive compensation.", *Journal of Business Research*, 64(2), pp. 187-193.
- (26) Banker, R. D., Y. Liang and X. Ma, 2016, "A textual measure of strategy." A textual measure of strategy, Working paper.
- (27) Banker, R. D., R. Mashruwala and A. Tripathy, 2014, "Does a differentiation strategy lead to more sustainable financial performance than a cost leadership strategy?" , *Management Decision* , 52(5), pp. 872-896.
- (28) Banker, R. D. and M. Xinjie, 2017, " Business strategy and accounting research: A textual-analysis approach." Working paper, Temple University.
- (29) Barney, J. B., 1997, "Gaining and Sustaining Competitive Advantage.", Upper Saddle River, NJ: *Prentice-Hall*, Inc 222(s 3–4), pp. 361-369.
- (30) Chandler, A. D., 1962, " Strategy and Structure: Chapters in the History of the American Industrial Enterprise.", Cambridge, Mass., *MIT Press*.
- (31) Cheng, S., 2008, "Board size and the variability of corporate performance.", *Journal of Financial Economics*, 87(1), pp.157-176.
- (32) Claver-Cortés, E., E. M. Pertusa-Ortega and J. F. Molina-Azorín, 2012, "Characteristics of organizational structure relating to hybrid competitive strategy: Implications for performance.", *Journal of Business Research*, 65(7), pp. 993-1002.
- (33) Collins, D. W. and S. P. Kothari, 1989, "An analysis of intertemporal and cross-sectional determinants of earnings response coefficients .", *Journal of Accounting & Economics*, 11(2), pp.143-181.
- (34) Davis, A. K., W. Ge, D. Matsumoto and J. L. Zhang, 2015, "The effect of manager-specific optimism on the tone of earnings conference calls.", *Review of Accounting Studies*, 20(2), pp.639-673.
- (35) Dechow, P., Ge, W., and Schrand, C., 2010, "Understanding earnings quality: A review of the proxies, their determinants and their consequences.", *Journal of Accounting and Economics*, 50(2-3), pp.344-401.
- (36) Dechow, P. M., S. P. Kothari and R. L. Watts, 2006, "The relation between earnings and cash flows." *Social Science Electronic Publishing*, 25(2), pp.133-168.
- (37) Dess, G. G. and P. S. Davis, 1984, "Porter'(1980) Generic Strategies as Determinants of Strategic Group Membership and Organizational Performance." , *Academy of Management Journal*, 27(3), pp. 467-488.
- (38) Devaraj, S., D. G. Hollingworth and R. G. Schroeder, 2004, "Generic manufacturing strategies and plant performance.", *Journal of Operations Management* , 22(3), pp.313-333.
- (39) Dickinson, V., 2011. " Cash Flow Patterns as a Proxy for Firm Life Cycle. ", *The Accounting Review*, 86(6), pp.1969-1994.
- (40) Dichev, I. D., and Tang, V. W., 2009, "Earnings volatility and earnings predictability.", *Journal of Accounting and Economics*, 47(1-2), pp. 0-181.
- (41) Dickinson, V., 2011, "Cash Flow Patterns as a Proxy for Firm Life Cycle." , *Accounting Review*, 86(6), pp.1969-1994.
- (42) Feng, L. I., 2010, "The Information Content of Forward-Looking Statements in Corporate Filings—A Naïve Bayesian Machine Learning Approach.", *Journal of Accounting Research*, 48(5), pp.1049–1102.
- (43) Francis, J., R. Lafond and K. Schipper, 2004, "Costs of Equity and Earnings Attributes.", *Accounting Review*, 79(4), pp. 967-1010.
- (44) Frankel, R., and Litov, L., 2009. " Earnings persistence "., *Journal of Accounting and Economics*, 47(1-2), pp.182-190.
- (45) Graham, J. R., Harvey, C. R., and Rajgopal, S. , 2005., " The economic implications of corporate financial reporting." , *Journal of Accounting and Economics*, 40(1-3), pp.3-73.

- (46) Hambrick, D. C., 1983, "High Profit Strategies in Mature Capital Goods Industries: A Contingency Approach." *Academy of Management Journal*, 26(4), pp. 687-707.
- (47) Hambrick, D. C., I. C. Macmillan and D. L. Day, 1982, "Strategic Attributes and Performance in the BCG Matrix--A PIMS-Based Analysis of Industrial Product Businesses.", *Academy of Management Journal*, 25(3), pp.510-531.
- (48) Heinle, M. S., Smith, K. C., and Verrecchia, R. E., 2017, "Risk-factor disclosure and asset prices.", *The Accounting Review*, 93(2), pp.191-208.
- (49) Huang, K. W. and Z. Li, 2011., "A multilabel text classification algorithm for labeling risk factors in SEC form 10-K.", *Social Science Electronic Publishing*, vol.2, pp.18.
- (50) Jusoh, R., and Parnell, J. A., 2008, "Competitive strategy and performance measurement in the Malaysian context: An exploratory study.", *Management Decision*, 46(1), pp.5-31.
- (51) Jr, R. B. R. and J. A. P. Ii, 2010, "Planned patterns of strategic behavior and their relationship to business-unit performance.", *Strategic Management Journal*, 9(1), pp. 43-60.
- (52) Kim, E., D. I. Nam and J. L. Stimpert, 2004, "Testing the applicability of Porter's generic strategies in the digital age: a study of Korean cyber malls.", *Journal of Business Strategies*, 21(1).
- (53) Kormendi, R. and R. Lipe, 1987, "Earnings Innovations, Earnings Persistence, and Stock Returns.", *Journal of Business*, 60(3), pp. 323-345.
- (54) Kotha, S. and D. Orne, 1989, "Generic manufacturing strategies: A conceptual synthesis." , *Strategic Management Journal*, 10(3), pp. 211-231.
- (55) Lev, B. and T. Sougiannis, 1996, "The capitalization, amortization, and value-relevance of R&D.", *Industrial Health*, 53(6), pp. 498-504.
- (56) Lev, B. and S. R. Thiagarajan, 1993, "Fundamental Information Analysis.", *Journal of Accounting Research*, 31(2), pp.190-215.
- (57) Li, C. B., and Li, J. J., 2008, "Achieving superior financial performance in China: differentiation, cost leadership, or both?." *Journal of International Marketing*, 16(3), pp.1-22.
- (58) Loughran, T. and B. Mcdonald, 2011, "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks.", *Journal of Finance*, 66(1), pp. 35-65.
- (59) Muslu, V., Radhakrishnan, S., Subramanyam, K. R., and Lim, D., 2014, "Forward-looking MD&A disclosures and the information environment." *Management Science*, 61(5), pp.931-948.
- (60) Mikolov, T., K. Chen, G. Corrado and J. Dean, 2013, "Efficient Estimation of Word Representations in Vector Space.", *Computer Science*.
- (61) Miles, R. E., C. C. Snow, A. D. Meyer and C. H. Jr, 1978, "Organizational strategy, structure, and process.", *Academy of Management Review*, 3(3), pp.546-562.
- (62) Miller, D. and P. H. Friesen. 1984. "A Longitudinal Study of the Corporate Life Cycle." *Management Science*, 30(10), pp. 1161-1183.
- (63) Miller, D. and P. H. Friesen, 1986, "Porter's (1980) generic strategies and performance: An empirical examination with American data: II. Performance implications.", *Organization Studies*, 7(1), pp. 37-55.
- (64) Miller, G. S. and D. J. Skinner, 2015, "The evolving disclosure landscape: how changes in technology, the media, and capital markets are affecting disclosure.", *Journal of Accounting Research*, 53(2), pp. 221-239.
- (65) Mintzberg, H., 1978, "Patterns in Strategy Formation.", *Management Science*, 24(9), pp. 934-948.
- (66) Mintzberg, H., 1987, "The strategy concept I: Five Ps for strategy.", *California Management Review*, 30(1), pp.11-24.
- (67) Nandakumar, M. K., A. Ghobadian and N. O'Regan, 2011, "Generic strategies and performance – evidence from manufacturing firms.", *International Journal of Productivity & Performance Management*, 60(3), pp.222-251.

- (68) Porter, M. E., 1980, "Competitive Strategy", *Free Press*.
- (69) Prescott, J. E., 1986, "Environments as Moderators of the Relationship between Strategy and Performance.", *Academy of Management Journal*, 29(2), pp.329-346.
- (70) Price, M. K., J. S. Doran, D. R. Peterson and B. A. Bliss, 2012, "Earnings conference calls and stock returns: The incremental informativeness of textual tone.", *Journal of Banking & Finance*, 36(4), pp.992-1011.
- (71) Quinn, R. E. and K. Cameron, 1983, "Organizational Life Cycles and Shifting Criteria of Effectiveness: Some Preliminary Evidence.", *Management Science*, 29(1), pp.33-51.
- (72) Rumelt, R. P., Schendel, D., and Teece, D. J., 1991, "Strategic management and economics.", *Strategic management journal*, 12(S2), pp.5-29.
- (73) Schipper, K., 1989, "Commentary on Earnings Management.", *Accounting Horizons*, 3.
- (74) Scholtes, V. A., C. B. Terwee and R. W. Poolman, 2011, "What makes a measurement instrument valid and reliable?", *Injury-international Journal of the Care of the Injured*, 42(3), pp. 236-240.
- (75) Shin, D., S. He, G. M. Lee, A. B. Whinston, S. Cetintas and K. C. Lee, 2016, "Content Complexity, Similarity, and Consistency in Social Media: A Deep Learning Approach.", *Social Science Electronic Publishing*.
- (76) Sloan, R. G., 1996, "Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings?", *Accounting Review*, 71(3), pp. 289-315.
- (77) Taddy, M., 2015, "Document Classification by Inversion of Distributed Language Representations.", *Computer Science*.
- (78) Turney, P. D., 2013, "Distributional semantics beyond words: Supervised learning of analogy and paraphrase.", *Computer Science*.
- (79) Weber, R. P., 1990, "Basic content analysis", *Sage Publications*.
- (80) Zhang, D., H. Xu, Z. Su and Y. Xu, 2015, "Chinese comments sentiment classification based on word2vec and SVM perf.", *Expert Systems with Applications*, 42(4), pp. 1857-1863.
- (81) Zott, C. and R. Amit, 2008, "The fit between product market strategy and business model: Implications for firm performance.", *Strategic Management Journal*, 29(1), pp. 1-26.

附录:

附录 1 Word2vec 神经网络语言模型简介

Word2vec 是对神经网络语言模型隐藏层到输出层以及输出层上的归一化运算量巨大提出的优化模型。其包含 CBOW 和 Skip-gram 两种模型。CBOW 模型是根据当前词的上下文来预测当前词, Skip-gram 模型是根据当前词来预测其他词出现在它上下文的概率, 即根据当前词来预测其上下文的一个词。为了减少投影层到输出层以及输出层上的归一化计算的运算量, CBOW 模型和 Skip-gram 模型在训练过程中又分别采用了两种技术, 称为 Hierarchical Softmax (HS) 和 Negative Sampling (NEGS)。

我们采用 CBOW+HS 模型进行训练。CBOW 模型包含输入层, 隐藏层以及输出层。神经网络的输入层是当前词的上下文向量, 通常定义一个窗口, 取其上下窗口内的单词; 隐藏层是对输入层词向量的求和或者平均; 输出层的维度为语料词汇量的大小。Hierarchical Softmax 是在投影层和输出层之间以语料中词汇为叶子节点, 词频为结点权重构建了一颗 Huffman 树。假设语料词汇量大小为 V , 那么 Huffman 树的深度为 $V-1$, 所以采用 Hierarchical Softmax 技术的模型是一个深度神经网络, 这也是 word2vec 被成为深度学习算法的原因。Huffman 树是一颗二叉树, 所以在树的每个结点上都相当于是一个二分类问题。Hierarchical Softmax 将原始投影层到输出层的 V 分类问题近似为 $\log(V)$ 个二分类问题, 其优势在于, 当新进来一个样本时, 随机梯度法每一次迭代时, 神经网络不再像原来一样更新全部的权重, 而是只更新连接目标词汇的 Huffman 路径上的权重, 这样大大加快了随机梯度法收敛的过程。在算生成的目标词汇的概率时, 只需要将路径上各个结点二分类后对应的概率相乘。神经网络结构如下图所示:

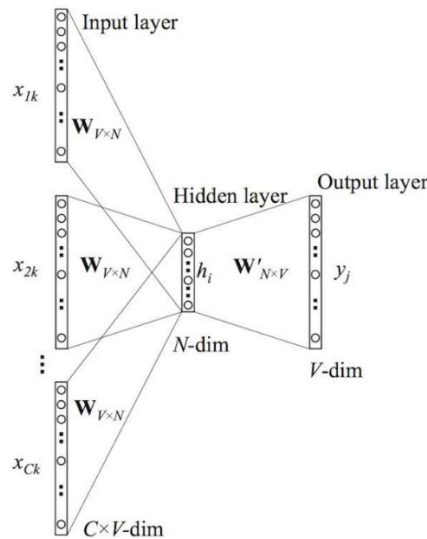


图 1 CBOW 神经网络模型结构图

在上图中, 输入层是由 one-hot 编码的输入上下文 $\{x_1, \dots, x_C\}$ 组成, 其中窗口大小为 C , 词汇表大小为 V 。隐藏层是 N 维的向量。最后输出层是也被 one-hot 编码的输出单词 y 。被 one-hot 编码的输入向量通过一个 $V \times N$ 维的权重矩阵 W 连接到隐藏层; 隐藏层通过一个 $N \times V$ 的权重矩阵 W' 连接到输出层。

接下来, 我们假设已知输入与输出权重矩阵的大小。

第一步, 计算隐藏层 h 的输出, 即输入向量的加权平均。如下:

$$h = \frac{1}{C} W \left(\sum_{i=1}^C x_i \right) \tag{1}$$

第二步, 计算在输出层每个结点的输入。如下:

$$u_j = v_{wj}'^T h \tag{2}$$

其中, v_j' 是词对应的输出向量, $v_{wj}'^T$ 是输出矩阵 W' 的第 j 列。

最后，计算输出层的输出 y_j ，如下：

$$y_{c,j} = P(w_{y,j} | w_1, \dots, w_c) = \frac{\exp(u_j)}{\sum_{j=1}^V \exp(u' j)} \quad (3)$$

这样就得到了通过第 j 个词的上下文 C 个词来预测该词的概率目标函数，通过最大化此目标函数 (3)，即可得到该词的词向量。

附录 2 竞争战略词集示例

差异化战略部分词集	独特，唯一，新产品，品牌，营销创新，多样化，特色化...
成本领先战略部分词集	成本控制，经营效率，精细化管理，降本增效，规模化...

附录 3 公司年报竞争战略信息披露示例

Panel A 差异化战略

公司名称	战略词频	年报信息披露示例
广誉远中药	327	“2016 年公司……继续以全产业链打造广誉远高品质中药战略为指导……不断创新营销，强化市场拓展和渗透，完善、深化营销渠道，加速产能提升和产品结构优化升级，进一步提升运营效率，促进公司持续、健康、快速发展。”
金发拉比孕婴童用品	326	“公司始终秉承“绿色自然，安全舒适”的理念，制定高于国家和行业标准的企业标准，建立严密有效的品质管控体系，把严格的质量检测渗透到每个生产物流环节，确保产品有高标准亲肤舒适性和触肤安全性。”
联建光电	319	“公司致力于打造数据驱动的智能整合营销服务集团……依据自身的核心竞争力和长期发展规划，一方面开展一系列的宣传推广活动和小间距 LED 显示应用的持续开发，保持了数字显示业务的领先地位；另一方面，通过投资并购的方式整合营销传播服务行业，为客户品牌管理提供全生命周期的服务。”
利欧集团	312	“公司坚定不移执行向互联网产业转型的战略……构建了同时覆盖品牌客户和效果类客户、覆盖全部数字媒体、提供全品类服务的业务体系，并从业务和资本两个层面不断丰富公司的业务生态，稳步扩大业务体量，持续提升盈利能力。”
深圳歌力思服饰	307	“歌力思……鼓励品牌内生增长；同时着延伸品类设计，体现品牌附加值。公司未来将持续推动国际资源整合，拓展歌力思高级时装品牌集团的“版图”。不断延伸和扩大其国内和国际的影响力，融会国际资源，以国际化视野来向中国消费者展现自己的世界观，带来真正属于中国创造的国际时尚。”

Panel B 成本领先战略

公司名称	战略词频	公司理念
湖北兴发化工	396	“公司通过立足科学发展，创新发展理念，转变发展方式，走专业化、集约化、精细化、信息化的道路，大力推进技术创新和循环经济，持续推动节能减排，为公司可持续健康发展提供有力保障。”
安徽广信农化	357	“公司一贯秉持‘广结中外、信昭天下、精耕细作、永续经营’的发展理念，以及安全至上和环保当先的经营思想，把握市场导向，以务实开拓的创业精神引领企业不断稳定发展，永攀高峰。”
新奥生态控股	339	“公司面对复杂的宏观经济形势和新产业发展态势，坚持构建低成本核心竞争力，审慎应对各种风险和挑战……通过采取强化精益管理持续降低生产成本激发现有业务潜能……。”
北海国发海洋生物产业	323	“公司将积极把握国家战略机遇，顺应行业发展趋势，紧紧围绕医疗大健康产业这一主线，通过内强外延双轮驱动，精耕细作，创新发展……努力实现公司综合一流、可持续、快速健康发展。”
中国石油化工	320	“公司坚持以供给侧结构性改革为主线，充分发挥一体化优势，扬长补短，不断增加对社会石油石化产品服务的有效供给，实现了经济效益和资产效率的较大提升……压减高成本原油产量，有效控制了生产成本，持续改善能源结构……。”