

doi:10.19920/j.cnki.jmsc.2025.03.001

# 基于文本分析的首席经济学家信心指数的构建与宏观预测<sup>①</sup>

张涛<sup>1</sup>, 汤珂<sup>1</sup>, 谢海华<sup>2</sup>

(1. 清华大学社会科学学院经济学研究所, 北京 100084; 2. 北京雁栖湖应用数学研究院, 北京 101408)

**摘要:** 经济学家的综合研判和信心程度在经济波动和不确定性加剧的今天更加具有参考价值,然而现有的经济学家信心指数频率较低、编制成本较高、时效性不强,因此没有受到足够的重视.本研究借助“首席经济学家论坛”在线文本数据、通过前沿的自然语言处理技术 TextRank + FinBERT 方法构建了月频的首席经济学家信心指数及子指数(CECI).研究发现,首席经济学家信心指数和统计局的季度经济学家信心指数走势一致,但是时效性更强、更新频率更高;相比其他主体信心,首席经济学家信心作为景气指标效果更好;在宏观预测方面,首席经济学家信心指数系列能显著改进对主要宏观变量的样本外预测效果.本研究是使用人工智能方法构建景气指标并进行宏观预测的有益尝试,本文的方法论可以应用到其他高频景气指数的构建上.

**关键词:** 经济景气指数; 经济学家信心; 自然语言处理; 宏观预测

**中图分类号:** F222.1; C813 **文献标识码:** A **文章编号:** 1007-9807(2025)03-0001-18

## 0 引言

我国正面临后疫情时代的经济结构调整和国内外风险等挑战.及时监测和预测经济周期波动有助于更精准的宏观调控,防范系统性风险,促进高质量发展.此外,经济主体信心的监测与引导也是当前形势下宏观调控的重要抓手.在2022年党的二十大后的中央经济工作会议上,习近平总书记强调,“爬坡过坎,关键是提振信心.”<sup>②</sup>,因此,在当前复杂经济形势下,经济学家群体的信心水平是一个重要的观测变量,首先,专家的综合判断比统计指标可能更具信息含量和稳定性,尤其在经济拐点处,经济专家的判断更准确<sup>[1]</sup>.其次,经济学家的信心和情绪变动具有代表性,有利于政府监测和引导市场主体信心.

在官方景气指数体系中,经济学家信心属于

景气动向调查指标.但是官方的经济学家信心指数由于问卷调查成本高、发布频率低(季度)、滞后时间长,在学界和业界的应用都较少.因此,能否利用大数据和人工智能技术来优化传统的经济学家信心指数,成为值得探讨的问题.本研究选取了经济学家时评文本这一独特的文本数据源,使用 TextRank + FinBERT 自然语言处理方法,构造了在线的首席经济学家信心指数系列(Chief Economist Confidence Indexes, CECI).包括1个总信心指数和产出、房地产、国际、风险、金融、政策、物价这7类子信心指数.并实证检验了指数作为景气指标和改进宏观预测等方面的表现.

本文经过对指数系列的评测主要有以下发现:第一,首席经济学家信心指数和统计局的季度经济学家信心指数走势高度一致,因此兼顾了低成本、高时效、高频率和代表性.第二,首席经济学

① 收稿日期: 2023-02-16; 修订日期: 2024-09-27.

基金项目: 国家自然科学基金资助项目(72192802; 72342008).

作者简介: 张涛(1998—),男,湖南娄底人,博士生. Email: zhangtao\_econ@163.com

② 网址: <http://cpc.people.com.cn/n1/2022/12/18/c64387-32588970.html>

家信心指数系列的景气领先性和景气指示性均较好,尤其是总信心指数和产出信心指数,优于采购经理信心指数系列,更是远好于消费者和投资者的信心指标。第三,首席经济学家信心指数系列具有增量信息,可以改进对主要宏观变量的预测效果,包括规模以上工业增加值、社会消费品零售总额、房地产开发新增固定资产、进出口金额、美元兑人民币、M2、CPI这7个重要宏观经济变量。

本研究可能有两方面的边际贡献。第一,在景气指数的构建方法上,基于经济主体信心调查构造景气指数的理论进展较为缓慢,目前仍主要依托于问卷调查,时效性较差、编制成本较高,因此国内外的景气动向调查主要由政府主导推进<sup>[2-4]</sup>,削弱了其实用价值。本研究提出的自然语言处理方法则为信心指数构建提供了新的方法论,有利于降低指数编制成本和提高指数发布的时效性。第二,本研究对文本大数据和自然语言处理技术的经济金融应用提供了新的视角和案例,目前基于文本大数据的经济金融研究包括基于财经报纸文本的经济政策不确定性指数构建<sup>[5,6]</sup>、基于网络财经媒体文本的股票预测性检验<sup>[7-9]</sup>、基于上市公司年报或公告文本的公司金融研究等<sup>[10-14]</sup>。将文本分析技术应用于经济主体信心指数编制的研究还几乎没有,本研究是对这一领域的有益补充。

## 1 文献综述

### 1.1 经济景气指数的类型与编制方法

20世纪初,经济周期波动的量化测度开始兴起,从20世纪初的哈佛指数到20世纪50年代美国NBER扩散指数<sup>[15]</sup>,再到基于状态空间模型的SWI景气指数<sup>[16]</sup>。基本思想是基于多维的统计指标和统计模型合成刻画经济状态的景气指数。随着可用指标的增多,Aruoba等<sup>[17]</sup>利用状态空间模型构建了美国的经济周期高频指数,美联储进一步运用混频的多维数据和混频动态因子模型进行GDP的实时宏观预测<sup>[18]</sup>。2020年新冠疫情之后,各国对实际经济活动的短期波动关注度升高。

Lewis等<sup>[19]</sup>使用美国周度经济指标集和动态因子模型构建了一个每周经济指数(WEI)跟踪高频经济变化。Wegmüller等<sup>[20]</sup>则用类似的方法将多个周度指标合成了综合指数,衡量瑞士经济的实际活动水平。除了基于统计指标合成的景气指数,另一类重要的景气监测方法是景气动向调查。景气动向调查方法通过问卷调查直接获取经济主体的预期和信心,典型的采访对象包括消费者、投资者、采购经理、企业家和经济专家等。德国慕尼黑ifo经济研究所从1949年开始了世界上第一次系统的企业信心问卷调查,从1983年开始进行国际性的经济学家信心调查,涉及全球数十个国家和地区。该指数目前在国际宏观经济指标预测和经济周期分析等领域发挥着重要作用<sup>[4]</sup>。

### 1.2 中国的经济景气指数研究

中国的经济景气研究吸收了国际先进经验,结合中国特色,取得了持续的进步。在经济预测模型<sup>[21]</sup>和基于统计指标的景气指数构建方面,产生了较多的学术进展。陈磊和高铁梅<sup>[22]</sup>编制了SWI型景气指数,并应用于实际经济分析与预测。刘汉和刘金全<sup>[23]</sup>使用混频数据进行经济总量实时预报与短期预测。郑挺国和王霞<sup>[24]</sup>设计了基于混频数据测度的中国经济周期一致指数。费兆齐和刘康<sup>[25]</sup>利用金融数据编制了日度经济先行指数对宏观经济进行高频监测研究。

另一方面,在景气动向调查和经济主体信心指数方面,理论进展较缓、政府主导的实践推进较多。我国现在运行的基于景气调查的指数主要包括消费者信心指数、采购经理指数(PMI)、经济学家信心指数,涉及主体均为经济各部门重要参与者或者研究者<sup>[2,3]</sup>。1997年开始,国家统计局中国经济景气监测中心开始建立全国范围内的消费者信心调查,每个月进行一次。2004年,汇丰中国制造业采购经理指数开始发布。2005年,中国国家统计局与中国物流与采购联合会正式发布中国制造业PMI,并把PMI调查列为国家统计局的正式调查制度,中国非制造业PMI也于2008年开始发布。在实践中,PMI指标系列是目前应用最为广泛的景气调查指数,是金融机构和政府部门的重要决策依据。在经济学家信心指数方面,中国经济

景气监测中心借鉴了 ifo 的成功经验,于 2005 年开始正式进行中国经济学家信心指数调查,数据以季度频率更新<sup>[3]</sup>。但是官方的经济学家信心指数因为发布频率低、滞后时间长,在学界和业界的应用都较少。

### 1.3 文本大数据的经济金融应用

理论研究和文献计量学研究均显示,大数据技术改变了经济研究和经济预测的研究范式<sup>[26, 27]</sup>。包括从模型驱动到数据驱动、从结构化数据到非结构化数据、从低频到高频等。因此,近年来包括文本大数据在内的非结构化数据在经济金融领域获得了广泛应用。目前经济金融领域基于文本大数据的研究涉及不同文本来源和研究问题。一类是基于财经报纸文本构建的经济政策不确定性指数。经济不确定性在理论上对宏观经济系统具有重要影响<sup>[28]</sup>,但是却很难用传统指标衡量,得益于文本大数据和计量技术的进步, Baker 等<sup>[5]</sup>用美国的报纸新闻文本构建了月度的经济政策不确定性指数; Huang 和 Luk<sup>[6]</sup>以中国报纸文本为基础构建了中国经济政策不确定性月度指数,后续实证研究表明经济政策不确定性指数对金融机构系统性风险等经济变量存在显著影响<sup>[29]</sup>。一类是基于网络财经媒体文本和股票预测的研究。Antweiler 和 Frank<sup>[7]</sup>基于雅虎发布的 150 多万条消息,发现留言板情绪可以预测股票市场的收益和波动。部慧等<sup>[8]</sup>基于东方财富网股吧帖文与朴素贝叶斯方法,发现基于股评的投资者情绪对股票收益率和交易量有当期影响。罗琦等<sup>[9]</sup>通过百度指数平台搜索量数据构建投资者现金股利情绪指数和高送转情绪指数,并考察了 A 股上市公司的股利迎合行为及股价崩盘风险。还有一类是基于上市公司年报文本或公告文本的公司金融研究。Li<sup>[10]</sup>使用美国上市公司年报和季报文件的管理层讨论与分析构建了管理层语调指标,发现语调能预测公司未来的盈利。Jiang 等<sup>[11]</sup>使用上市公司财报和电话会议记录等文本数据衡量经理人情 绪,发现经理人情 绪能显著反向预测未来市场收益率。王克敏等<sup>[12]</sup>研究了管理者出于自利动机的

年报文本信息复杂性披露策略。胡楠等<sup>[13]</sup>基于年报文本构建了管理者短视主义指标。成程等<sup>[14]</sup>通过对 A 股上市公司 320 多万篇公告的文本信息进行词频分析,检验了上市公司发展供应链金融业务对企业价值的影响。尽管文本大数据在经济金融方面的应用方兴未艾,但尽本文所知,将文本分析技术应用于经济主体信心指数编制的研究还几乎没有。

## 2 首席经济学家信心指数的构建

### 2.1 传统经济学家信心指数的编制方法

经济学家信心指数是由国家统计局景气监测中心设计和维护的重要指标,通过问卷的形式调查国内有代表性、有影响力的宏观经济学家(一般为 100 位),了解他们对经济运行状况、未来发展趋势等的判断。典型问题为“关于对当前经济形势的判断”、“关于对未来 6 个月的发展趋势判断”、“关于当前经济发展中存在的问题”等。问卷收集之后,对每个问题的回答赋分(特殊问题除外),最后根据所有专家对于某一问题的判断得分进行平均即可得到该问题的得分。问卷分别采集即期经济景气指数和预期经济景气指数,最后算术平均得到经济学家信心指数<sup>③</sup>。官方的经济学家信心指数是季度层面的指标,通过《中国经济景气月报》进行发布,当季数据发布时间一般滞后一个月左右。

### 2.2 首席经济学家信心指数系列的编制方法

#### 2.2.1 数据来源

本研究使用的文本数据来自于“首席经济学家论坛”网站,该论坛中的文本素材对于反映经济学家信心具有较好的代表性。论坛成员由专职中国经济分析的国内外一流金融机构的首席经济学家所组成。常任成员 93 位,专家姓名和所在机构的可视化词云可分别参见图 1 的左侧小图和中间小图,包括国务院参事、中国人民银行参事等政府经济学家和数十位知名金融机构的市场经济学家,市场机构包括知名的券商、基金、银行保险、资

③ 详情可见 <https://data.cnki.net/StatisticFocus/Article?id=30>

管和海外投行等. 在中国学术论坛影响力排行榜中,中国首席经济学家论坛以 377 万个百度结果搜索量,影响力全国前二. 论坛的主要业务范围包括“跟踪全球和国内经济的动向”和“甄别可能影响经济、贸易、金融发展的各种问题,搜集和发布相关信息,形成有关报告”. 时评文章内容关键词的可视化词云可见图 1 的右侧小图,涉及经济运行的不同侧面. 作为机构的首席经济学家,及时发

布最新最准确的观点极大地影响和决定其工作质量和工作评价,出于对个人声誉与机构声誉的维护,首席经济学家的观点质量有所保障. 相比起噪音较大的财经媒体报道,这些专家的言论更加专业和前瞻;相比起理论学者,这些经济学家活跃在市场一线、经济嗅觉更敏锐. 因此,在各人群中首席经济学家对经济形势的判断具有最好的权威性.

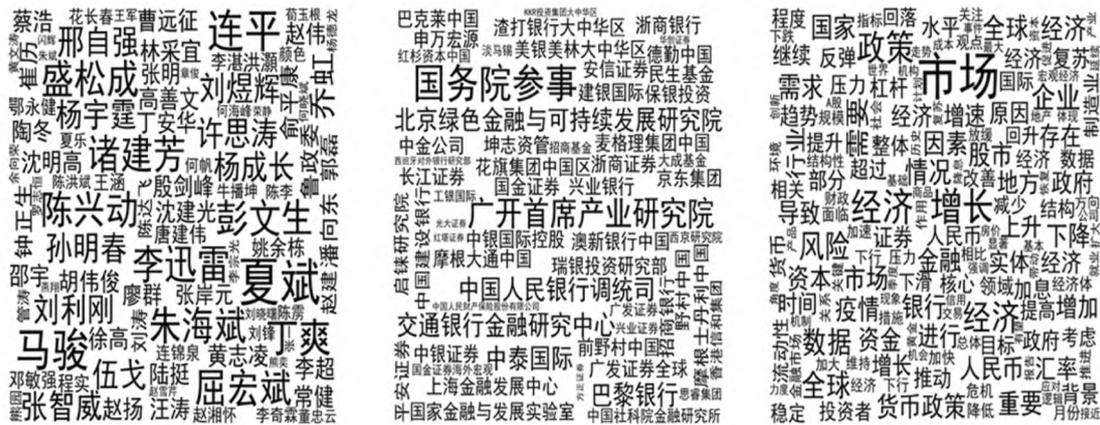


图 1 可视化词云: 首席经济学家姓名(左侧)、首席经济学家所在机构(中间)和文本关键词(右侧)

Fig. 1 Visualization word clouds: Names of chief economists (left), institutions of chief economists (center) and text keywords (right)

表 1 文章分类及部分对应关键词

Table 1 Article classification and some corresponding keywords

类别	词语
产出	商品(105)、利润(105)、社会(107)、零售总额(107)、新常态(113)、基本面(116)、新出口(116)、总需求(118)、产能(121)、需求端(122)、支出(123)、产能过剩(123)、高增长(124)
金融	企业债券(101)、负债(104)、银行存款(104)、融资平台(106)、政策利率(107)、高收益(110)、外汇占款(111)、金融改革(111)、估值(121)、牛市(122)、上海(123)、市场情绪(125)
风险	监管部门(108)、下行压力(114)、防风险(116)、风险偏好(126)、信用风险(148)、加杠杆(202)、金融监管(206)、监管(243)、波动(284)、危机(339)、债务(384)、金融风险(422)、不确定性(515)
国际	欧盟(101)、外汇市场(105)、全球化(110)、欧元(113)、人民币升值(125)、外部环境(143)、欧元区(147)、关税(167)、贸易摩擦(177)、英国(178)、人民币贬值(189)、人民币国际化(192)
物价	通胀预期(115)、食品价格(116)、石油价格(125)、价格指数(153)、通胀压力(160)、通缩(197)、通胀(485)、价格(735)、原油价格(103)、原油(104)、非食品价格(105)、食品(112)、高通胀(116)
政策	银行间市场(102)、货币当局(103)、逆周期(104)、宏观调控(113)、信贷增速(114)、力度(116)、操作(119)、基建(120)、创新(120)、公共服务(120)、工具(121)、宽松政策(141)、收紧(141)
房地产	房价上涨(103)、房地产调控(120)、房地产销售(150)、房地产投资(344)、房价(373)、房地产市场(471)、房地产(653)、房地产行业(133)

注: 括号中数字为词频. 由于篇幅限制,表中只列出了少部分关键词,完整表格留存备案.

本研究使用爬虫的方法爬取论坛上所有文章,时间范围从 2014 年 3 月到 2022 年 12 月,一共有 8 302 篇文章. 每个月平均采样约 78 篇高质量的分析文章. 第一步,对文本库进行关键词提

取. 对于每篇文章,采用 CnKPRank 算法<sup>[30]</sup>提取能够代表文章主旨内容和思想的关键词及词组,共有 893 个. 代表性词语包括“产出”、“结构调整”、“通胀”、“货币政策”等. 这些关键词及词组

是“首席经济学家论坛”文章的具有代表性和概括性的词汇,是后续的文章分类的基础.第二步,根据经济含义对关键词进行聚类.考虑到宏观经济的复杂性,本研究通过多轮小组讨论和专家建议,将关键词分为7大类别,分别是“产出”、“金融”、“风险”、“国际”、“物价”、“政策”、“房地产”(见表1文章分类及部分对应关键词).这7个类别基本涵盖了所有文章,具有代表性和完备性.而每篇文章根据关键词类型可能进入多个类别,保证了每个类别具有足够的文章数量.剩下没有确定类别的关键词被标注为“无”,它们在所有出现次数100以上的关键词及词组中占比57.1%.

这些分类对应着重要的经济运行侧面,也是构建指数集的基础,本文根据词语筛选出对应类

别的文章,并进一步计算总指数和子指数.相比单一指数而言,这种做法能更加充分准确地挖掘文本集包含的不同侧面信息.

### 2.2.2 文章摘要提取以及情感分析

#### 1) 基于TextRank算法的文章摘要提取

本文采用TextRank方法<sup>[31]</sup>对原始文本进行预处理,构造出每篇文章的摘要.TextRank是无监督的方法,无需事先准备语料进行模型训练,因此它是一个十分简洁并且有效的文本分析方法.采用构造出的文章摘要而不是全文的主要原因在于,首席经济学家论坛的文章一般较长,而且含有一些冗余的信息,提取文章摘要可以将文章重要信息提炼出来,以提升后续情感分析的准确度和效率.

#### 算法1: 基于TextRank的文章摘要提取

输入: 文章  $A$

1. 文章  $A$  分句  $\rightarrow \{S_1, S_2, \dots, S_N\}$
  2. 基于“算法2”计算语句相似度:
    - for 每一个语句组合  $(S_i, S_j)$ :
      - | 计算  $Similarity(S_i, S_j)$
  3. 构造有权图  $G = (V, E)$ :
    - a. 构造节点集  $V$ :
      - for 文章  $A$  中的每个语句  $S_k$ :
        - | 构造节点  $v_k$ , 权重  $W(v_k) = 1$
    - b. 构造连线集  $E$ :
      - for 每一个语句组合  $(S_i, S_j)$ :
        - | if  $Similarity(S_i, S_j) > 0.35$ :
          - | 构造连线  $e_{ij}$ , 权重  $W(e_{ij}) = Similarity(S_i, S_j)$
  4. 迭代计算节点权重
    - for  $i = 1$  to  $N$ :
      - |  $L(v_i) \leftarrow v_i$  的邻接节点集
      - |  $W(v_i) = (1-d) + d \times \sum_{v_j \in L(v_i)} \frac{W(e_{ij})}{\sum_{v_k \in L(v_j)} W(e_{jk})} W(v_j)$
      - until  $W(v_i)$  变化幅度小于  $\lambda$
  5. 基于节点权重对语句进行排序
  6. 获取节点权重前五的语句:  $S_{r_1}, S_{r_2}, S_{r_3}, S_{r_4}, S_{r_5}$
- 输出: 摘要  $M = \{S_{r_1}, S_{r_2}, S_{r_3}, S_{r_4}, S_{r_5}\}$

#### 算法2: 语句相似度计算

输入: 语句  $S_i$ , 语句  $S_j$

1.  $S_i$  分词  $\rightarrow \{C_{i1}, C_{i2}, \dots, C_{iX}\}$
  2.  $S_j$  分词  $\rightarrow \{C_{j1}, C_{j2}, \dots, C_{jY}\}$
  3. 基于word2vec计算词语向量:
    - for  $z = 1$  to  $X$ 
      - |  $E(C_{iz}) \leftarrow \text{word2vec}(C_{iz})$
    - for  $z = 1$  to  $Y$ 
      - |  $E(C_{jz}) \leftarrow \text{word2vec}(C_{jz})$
  4. 计算语句向量:
    - $E(S_i) = \frac{E(C_{i1}) + E(C_{i2}) + \dots + E(C_{iX})}{X} = \{e_{i1}, \dots, e_{iX}\}$
    - $E(S_j) = \frac{E(C_{j1}) + E(C_{j2}) + \dots + E(C_{jY})}{Y} = \{e_{j1}, \dots, e_{jY}\}$
  5. 计算语句相似度:
    - $Similarity(S_i, S_j) = \frac{\sum_{k=1}^T e_{ik} \times e_{jk}}{\sqrt{\sum_{k=1}^T (e_{ik})^2} \times \sqrt{\sum_{k=1}^T (e_{jk})^2}}$
- 输出:  $Similarity(S_i, S_j)$

图2 算法伪代码: 基于TextRank的文章摘要提取(左侧)和语句相似度计算(右侧)

Fig. 2 Algorithm pseudocode: Article summary extraction based on TextRank (left) and sentence similarity calculation (right)

基于TextRank方法获取文章的摘要的流程(见图2左侧“算法1: 基于TextRank的文章摘要提取”)如下:

① 将文章划分为多个语句,例如:文章被划分为  $\{S_1, S_2, \dots, S_N\}$ , 其中  $N$  是文章的语句数量,  $S_i (1 \leq i \leq N)$  表示一个语句.

② 计算每一对语句之间的相似度. 对语句  $S_i$  和  $S_j$  进行相似度计算的流程(见图 2 右侧“算法 2: 语句相似度计算”)如下:

a) 将语句进行分词, 例如: 语句  $S_i$  分词为  $\{C_{i1}, C_{i2}, \dots, C_{iX}\}$ , 其中  $X$  是分词结果中词语的数量,  $C_{ik}$  ( $1 \leq k \leq X$ ) 表示  $S_i$  分词结果中的一个词语. 不同语句的分词结果中词语的数量可能不同.

b) 采用 word2vec 方法获取词语的向量, 然后将所有词语的向量进行累加并求平均, 得到该语句的向量

$$E(S_i) = \frac{E(C_{i1}) + E(C_{i2}) + \dots + E(C_{iX})}{X} \quad (1)$$

式中  $E(C_{ij})$  是词语  $C_{ij}$  的词向量,  $E(S_i)$  是语句  $S_i$  的向量.  $E(S_i) = \{e_{i1}, e_{i2}, \dots, e_{iT}\}$ , 其中  $T$  是语句向量的维度,  $e_{ik}$  是语句向量的第  $k$  个分量.

c) 采用余弦相似度方法计算语句之间的相似度

$$Similarity(S_i, S_j) = \frac{\sum_{k=1}^T e_{ik} \times e_{jk}}{\sqrt{\sum_{k=1}^T (e_{ik})^2} \times \sqrt{\sum_{k=1}^T (e_{jk})^2}} \quad (2)$$

$Similarity(S_i, S_j)$  表示语句  $S_i$  和  $S_j$  之间的相似度. 相似度值在 0 到 1 之间, 数值越大则语句相似程度越高.

③ 构建图结构  $G = (V, E)$

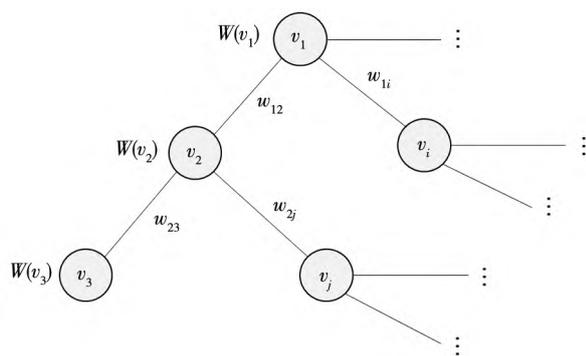


图 3 基于文本语句以及语句相似度构建的有权图结构示例  
Fig. 3 An example of a weighted graph structure constructed based on text sentences and sentence similarity

图 3 是基于文本构建的图的结构示例, 图  $G$  是一个无向有权图.  $W = \{W(v_1), W(v_2), \dots,$

$W(v_N)\}$  表示节点的权重, 即语句对于文章的重要性, 初始值都为 1.  $w_{ij}$  表示节点  $v_i$  和  $v_j$  的连线的权重, 它的值为  $v_i$  和  $v_j$  对应的语句的相似度. TextRank 算法主要针对节点权重(即  $W$  值)进行迭代计算, 计算公式如下

$$W(v_i) = (1-d) + d \times \sum_{v_j \in L(v_i)} \frac{w_{ji}}{\sum_{v_k \in L(v_j)} w_{jk}} W(v_j) \quad (3)$$

其中  $L(v_i)$  表示  $v_i$  的邻接节点集合.  $d$  是阻尼系数(damping factor), 数值在 0 ~ 1 之间. TextRank 算法运用式(3)不断地迭代更新每个节点的权重, 直到达到收敛条件.

④ 经过 TextRank 迭代计算之后, 权重排名靠前的  $n$  个节点对应的语句被选为关键语句并组成摘要.  $n$  的取值可以基于语句数量的限制或者文字数量的限制.

### 2) 基于 FinBERT 模型的文本情感分析

本文采用 FinBERT 方法对文本进行情感分析. FinBERT (BERT for Financial Text Mining) 基于金融领域的语料训练, 并采用多任务方式进行学习. 中文 FinBERT<sup>④</sup> 的训练数据来自中文财经新闻, 金融研报及上市公司公告和金融类百科词条等. 相比 BERT、BERT-wwm 和 RoBERTa-wwm-ext 等通用领域的预训练模型, 中文 FinBERT 在多个金融领域的具体任务中获得了显著的性能提升, 将 F1 值提高了 2% 至 5.7%. 情感类型分为三种: 正向, 中性和负向. 基于 FinBERT 模型的情感分析结果是三种情感的评分, 数值在 0 ~ 1 之间, 三种情感评分的总和为 1, 信心指数算法基于正向情感评分值来进行计算.

图 4 是基于 FinBERT 的财经文章情感分析的模型图. FinBERT 接收文章摘要作为输入并对其进行编码. 编码分成三部分:

- a) 词语编码: 对摘要文章进行分词, 并采用 word2vec 方法给每个词语进行初始编码.
- b) 语句编码: 对输入的摘要的语句使用不同符号进行编码.
- c) 位置编码: 表示词语和符号在输入文本中

④ 网址: <https://github.com/valuesimplex/FinBERT>.

的位置.

每个位置的三种编码经过累加之后,输入到 Transformer 编码器进行编码. [CLS]位置(即输入序列的第一个位置)产生的 Transformer 编码  $\vec{x}$  输入到分类器产生分类结果. 分类器是 FFNN(Feed-

forward Neural Network) + softmax 的结构.  $\vec{x}$  是一个维度为  $1 \times 512$  的向量,经过 FFNN 和 softmax 的输出结果为:  $\vec{y} = \{y_1, y_2, y_3\}$ ,这是一个维度为  $1 \times 3$  的向量,其中  $y_1, y_2, y_3$  分别表示三种情感(正向,中性和负向)的概率.

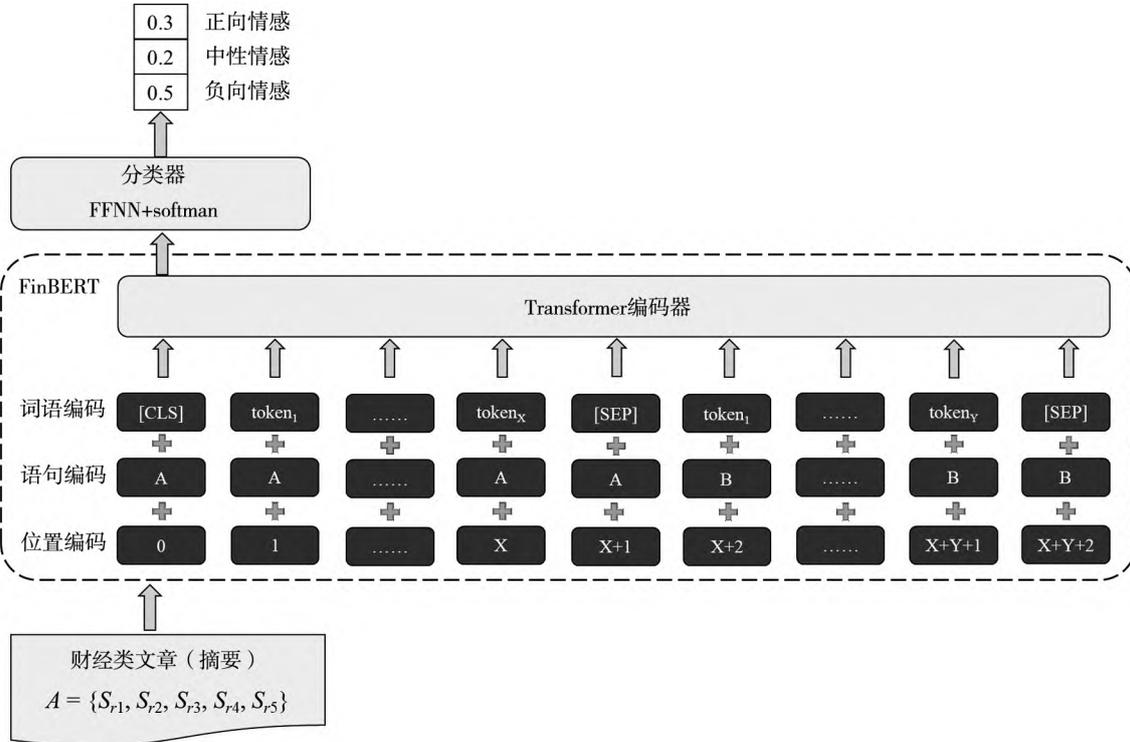


图4 基于 FinBERT 的财经文章情感分析

Fig.4 Sentiment analysis of financial articles based on FinBERT

2.2.3 指数的编制和计算流程

首席经济学家信心指数编制和计算流程如

图5所示. 主要分为文本所属类别的划分和信心指数的计算这两部分,具体如下:

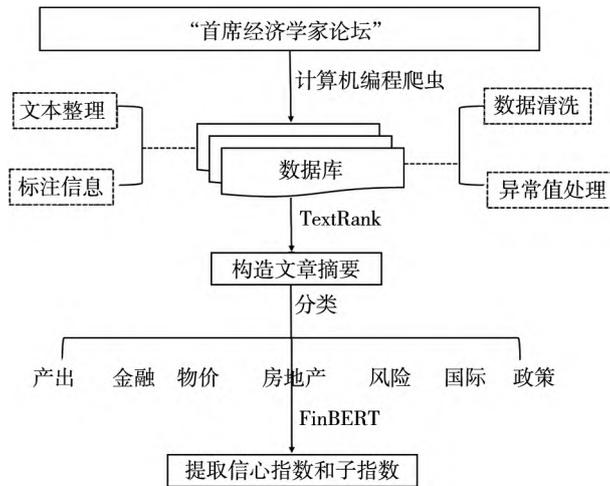


图5 首席经济学家信心指数的编制和计算流程

Fig.5 The construction and calculation process of The Chief Economist Confidence Indexes

### 1) 文本所属类别的划分

①提取每篇文章的关键词.采用中文关键词抽取方法<sup>[30]</sup>,提取每篇文章的关键词并计算出词语权重,取权重排名在前30的词语作为文章关键词.

②基于关键词对文章进行分类.根据表1的类别与词语的对应关系,统计出文章关键词在7种类别的分布情况,并依此确定文章类别(可以有多个).文章类别划分的标准是:如果存在一个或者多个类别的关键词等于或者大于2个,则把这些类别作为文章的分类结果;否则,如果所有类别的关键词均等于1或0,则该文章的分类结果为“无类别”;另外,如果计算总类指数,则所有文本都归为“总类”.

### 2) 计算总类或子类信心指数

①基于TextRank算法构建每篇文章的摘要.对于每篇从“首席经济学家论坛”获取的文章,使用“2.2.2节的1).基于TextRank算法的文章摘要提取”描述的方法计算每个语句的权重,然后取权重前5的语句组成摘要.如果上述摘要的字数大于500,则截取前500个字.

②基于FinBERT模型分析文章摘要的情感.

使用“2.2.2节的2).基于FinBERT模型的文本情感分析”描述的方法,以文章摘要作为输入,由模型计算出该文章的情感得分.

③计算经济学家的信心指数.对一个月内的文章进行类别划分.计算总类/某个子类的信心指数时,将属于该类别的所有文章按照前述方法进行情感分析,并求出情感得分的平均值,即是该类别的首席经济学家信心指数.

## 3 指数的初步分析与对比

### 3.1 变量介绍

本研究所涉及主要数据指标(见表2)包括首席经济学家信心指数、官方信心指数、官方统计指标三大类.首席经济学家信心指数包括1个总信心指数和产出、房地产、国际、风险、金融、政策、物价这7类子信心指数.从描述性统计结果来看,各指数经过标准化之后,均值和标准差一致,但首席经济学家信心的总指数和金融子指数最小值相比其余指数较低,风险指数的最大值相比其余指数较高.说明这几个指数曾在某些时点偏离均值程度较大.

表2 主要变量描述性统计

Table 2 Descriptive statistics of main variables

变量	含义	观测量	最小值	最大值	均值	标准差
首席经济学家信心指数						
<i>CECI</i>	首席经济学家信心.总指数	106	-2.19	3.63	0.00	1.00
<i>CECI_output</i>	首席经济学家信心.产出	106	-1.79	3.57	0.00	1.00
<i>CECI_housing</i>	首席经济学家信心.房地产	106	-1.84	2.60	0.00	1.00
<i>CECI_intl</i>	首席经济学家信心.国际	106	-1.55	3.01	0.00	1.00
<i>CECI_risk</i>	首席经济学家信心.风险	106	-1.79	4.17	0.00	1.00
<i>CECI_finance</i>	首席经济学家信心.金融	106	-2.03	3.10	0.00	1.00
<i>CECI_policy</i>	首席经济学家信心.政策	106	-1.93	3.58	0.00	1.00
<i>CECI_price</i>	首席经济学家信心.物价	106	-1.43	3.96	0.00	1.00
官方信心指数						
<i>PMI</i>	制造业采购经理人指数	106	35.70	52.40	50.31	1.73
<i>PMI_exp</i>	采购经理人指数.新出口订单	106	28.70	52.00	48.26	3.32
<i>con_conf</i>	消费者信心指数	106	85.50	127.00	112.61	11.03
<i>con_exp</i>	消费者预期指数	106	84.20	131.00	115.61	11.66
<i>inv_conf</i>	投资者信心指数	106	41.30	71.20	56.71	5.98
<i>OECI</i>	经济学家信心指数*	36	60.20	151.50	90.15	22.27
<i>OECI_now</i>	经济学家信心指数.即期*	36	15.60	131.30	74.36	28.18
<i>OECI_exp</i>	经济学家信心指数.预期*	36	55.90	180.40	105.96	35.91

续表 2

Table 2 Continues

变量	含义	观测量	最小值	最大值	均值	标准差
官方统计指标						
<i>GDP</i>	GDP, 不变价, 当季同比*	36	-6.90	18.70	6.10	3.44
<i>Coincident</i>	宏观经济景气指数, 一致指数	106	86.37	116.52	99.18	5.33
<i>IndOutput</i>	规模以上工业增加值, 当月同比	106	-25.87	52.34	6.23	6.41
<i>RealEstate</i>	房地产开发新增固定资产, 累计同比	106	-21.10	45.60	2.36	12.82
<i>Trade</i>	进出口金额, 当月同比	106	-22.41	67.79	5.54	13.99
<i>Exchange</i>	期末汇率, 美元兑人民币, 同比	106	-11.32	11.78	1.03	5.14
<i>M2</i>	广义货币 M2, 同比	106	8.00	14.70	10.34	1.88
<i>CON</i>	社会消费品零售总额, 当月同比	106	-20.50	34.20	7.67	7.07
<i>CPI</i>	CPI, 当月同比	106	-0.50	5.40	1.93	0.97

注：数据集的起止时点为 2014/03 到 2022/12，除了带 \* 号的变量为季度变量，其余均为月度变量。首席经济学家信心指数系列来自于本文的计算，已经标准化。其他数据来自于 Wind 数据库，官方信心指数和宏观经济景气一致指数直接使用原始水平值。官方统计指标有同比形式的直接用同比，无同比形式的用累计同比，单位为 %。

官方信心指数涵盖了消费者、投资者、采购经理、经济学家四大类经济主体，可用于不同经济主体信心的景气指示性对比实验。前三类为月度频率发布，涉及消费者的信心指数为消费者信心指数和消费者预期指数。涉及投资者信心的指标为投资者信心指数。涉及采购经理信心的指标包括制造业采购经理人指数 (PMI)、采购经理人指数新出口订单。涉及经济学家信心的官方指标为季度频率发布，包括即期信心、预期信心和总信心指数。从描述性统计结果来看，PMI 和 PMI\_exp 波动性较小，而几个季频的官方经济学家信心指数波动较大。

官方统计指标包括季度 GDP 同比和月度宏观经济一致指数，分别用于季度变量相关性分析和景气指示性分析。除此之外的月度统计指标则作为目标变量用于预测评估实验。这些宏观变量包括规模以上工业增加值、社会消费品零售总额、房地产开发新增固定资产、进出口金额、美元兑人民币、广义货币、消费者物价指数这 7 个主要宏观经济变量。官方统计指标来自于 Wind 数据库，有同比形式的直接用同比，无同比形式的用累计同比。这些指标是政府和业界分析预测的主要对象，全面地覆盖了宏观经济运行的产出、消费、投资、进出口、货币金融、物价等主要侧面。从描述性统

计结果来看，进出口金额同比和房地产开发新增固定资产累计同比波动最大，CPI 同比和 M2 同比波动较小。M2 同比的均值最大，超过 10%，社会消费品零售总额同比和规模以上工业增加值同比的均值次之，期末汇率同比和 CPI 同比的均值最小，处于 1%~2% 之间。

### 3.2 首席经济学家信心指数和官方指数的一致性

高频指标应该既具有和对应官方低频指标变动趋势的一致性，又能反应高频变动特征<sup>[32]</sup>，因此首先考察高频经济学家信心指数和官方指数的一致性，将其与传统的季度经济学家信心指数作对比。

图 6 展示了两个指数标准化后的走势图，可以看出，首席经济学家信心总指数和官方指数保持了总体变化的一致性。官方指数缺乏月度细节，经过插值之后比较平缓，而在线指数采样频率更高，因此波动性也更大，更能体现经济波动的细节。例如，在 2020 年初疫情冲击之下，官方指数没有明显地响应，而高频指数却出现了明显下降，准确地反应了宏观经济形势的变化。这说明高频指数不仅具有代表性，而且还具有很好的时效性，能够及时传递经济变动的信息。接下来通过核心指标之间的相关性分析来做出更准确的定量判断。

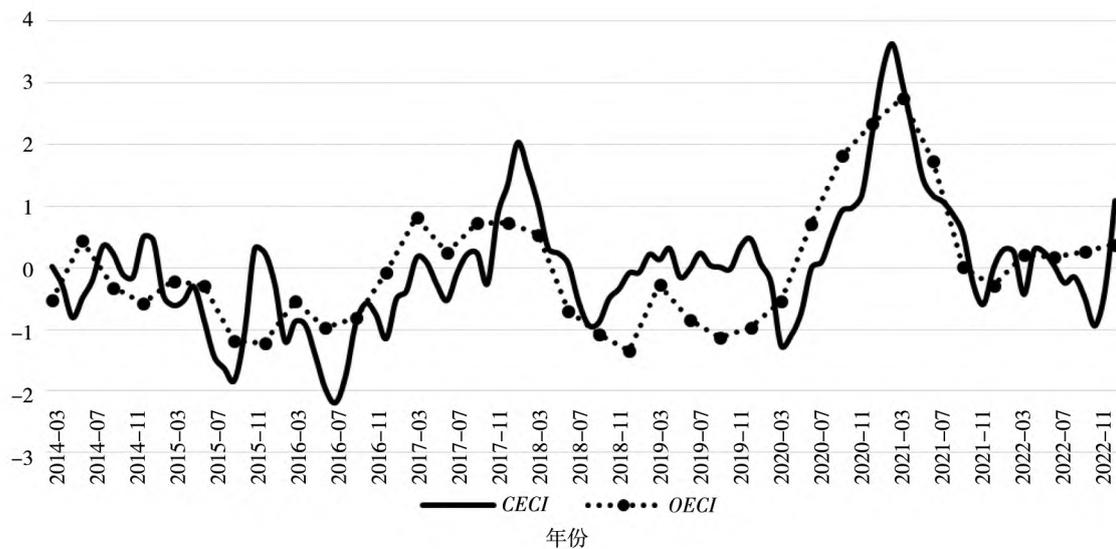


图6 首席经济学家信心指数和官方经济学家信心指数的对比

Fig. 6 Comparison of The Chief Economist Confidence Index and The Official Economist Confidence Index

注: OECI(官方经济学家信心指数)以季度频率更新,本文将其在月度层面进行线性插值从而画图. CECI(首席经济学家信心总指数)以月度频率更新.两个指数经过标准化以便于画图对比,横轴为月度日期,纵轴单位为标准差.

### 3.3 核心指标的相关性分析

从表3的第1列可以发现,首席经济学家信心指数系列中,总体信心和产出信心的相关性最大,说明总指数主要反映了产出方面的信心.而第5列显示,产出信心指数和总体信心是与官方信心相关性最大的两个指标,相关系数分别达到0.78和0.71.和GDP相关性最大的两个首席经济学家信心指数分别为政策信心和产出信心,相

关系数达到0.58和0.54.另外产出信心还和领先一期的GDP具有10%水平上显著的相关性(0.28).说明首席经济学家产出信心在预测经济总产出方面表现较好.对于官方信心的两个子指数,即期景气指数和预期景气指数来说,研究发现首席经济学家信心指数系列和官方即期景气指数的相关性基本均超过预期指数,具有更大且非常显著的相关系数,这说明高频指数主要反映的是即期信心.

表3 核心季度变量的相关性分析

Table 3 Correlation analysis of core quarterly variables

变量	CECI	GDP(-1)	GDP	GDP(+1)	OECI	OECI <sub>now</sub>	OECI <sub>exp</sub>
CECI	1.00	0.08	0.44 ***	0.25	0.71 ***	0.66 ***	0.41 **
CECI <sub>output</sub>	0.91 ***	0.06	0.54 ***	0.28 *	0.78 ***	0.70 ***	0.47 ***
CECI <sub>housing</sub>	0.44 ***	0.22	0.40 **	0.02	0.13	0.38 **	-0.13
CECI <sub>intl</sub>	0.76 ***	0.13	0.53 ***	0.23	0.64 ***	0.65 ***	0.33 *
CECI <sub>risk</sub>	0.59 ***	-0.22	0.37 **	0.06	0.57 ***	0.33 *	0.49 ***
CECI <sub>finance</sub>	0.76 ***	0.14	0.42 **	0.17	0.45 ***	0.63 ***	0.09
CECI <sub>policy</sub>	0.79 ***	0.06	0.58 ***	0.15	0.44 ***	0.46 ***	0.22
CECI <sub>price</sub>	0.50 ***	0.26	0.52 ***	0.14	0.45 ***	0.37 **	0.30 *

注:为了和其他季度指标对齐,首席经济学家信心指数在此处采用季度平均取值.“GDP”,“GDP(-1)”和“GDP(+1)”分别表示GDP季度同比,以及其滞后一期和领先一期.表中数字为皮尔森相关性系数,\*\*\*、\*\*、\*分别表示在1%、5%、10%水平下显著.

## 4 首席经济学家信心指数的进一步评测

### 4.1 不同经济主体信心的景气指示性

在经济统计实践层面,不同经济主体的信心是经济景气监测体系的重要组成部分.本文想要验证的是,通过人工智能技术高频化的首席经济学家信心能否有效地补充现有的景气指标体系.

从理论上来说,合格的景气指标应当具有以下特点<sup>[3, 33]</sup>:第一,经济重要性或综合性.首席经济学家虽然不直接参与经济活动,但是作为学者在分析判断经济形势上更加全面、准确、客观,在综合性方面占优;第二,统计连续与稳定性.即统计数据必须长期可追踪可信任,且景气对应关系需要具有相对稳定性.在这方面经济主体的信心指标具有相对优势,因为经济主体在形成信心和预期时会考虑包括经济结构变动等显性和隐性的

可得信息,经济学家群体因为具有更高的专业知识和理性水平尤其如此;第三,数据更新的及时性.这是在线指标最有优势的地方,传统的统计方法往往涉及到大范围的数据搜集或者问卷调查,时间成本和金钱成本较高,大数据时代的一个特点是显著降低了各项经济成本<sup>[34]</sup>,这使得基于在线大数据的指标能够真正做到低成本高频更新;第四,景气对应性,即统计上的相关性,这一点也是最重要的判断标准,可以通过统计方法判定,主流做法是时差相关分析.时差相关分析是从大量候选指标中挑选合适景气指标的标准做法<sup>[3]</sup>.首先选定反映经济景气的基准循环指标,然后计算每个候选指标和基准指标取不同延迟阶数时的互相关系数.其中绝对值最大的相关系数代表候选指标和基准指标之间的最优时差相关系数,相应的阶数代表时差阶数,反映了候选指标先行(时差阶数为负)、一致(时差阶数为零)或者滞后(时差阶数为正)的景气指示关系.

表4 不同主体信心与经济一致指数的时差相关分析

Table 4 Time-lagged correlation analysis between confidence of different entities and the economic coincident index

变量	-3	-2	-1	0	1	2	3	时差阶数
<i>CECI</i>	0.46	0.54	0.59	0.61	0.58	0.50	0.39	0
<i>CECI_output</i>	0.46	0.58	0.65	0.68	0.66	0.59	0.50	0
<i>CECI_housing</i>	0.44	0.46	0.44	0.37	0.31	0.23	0.15	-2
<i>CECI_intl</i>	0.45	0.52	0.56	0.54	0.49	0.41	0.33	-1
<i>CECI_risk</i>	0.22	0.32	0.38	0.41	0.40	0.35	0.30	0
<i>CECI_finance</i>	0.57	0.63	0.65	0.63	0.55	0.43	0.30	-1
<i>CECI_policy</i>	0.35	0.45	0.50	0.51	0.43	0.31	0.18	0
<i>CECI_price</i>	0.33	0.41	0.44	0.43	0.38	0.31	0.24	-1
<i>PMI</i>	0.24	0.32	0.40	0.46	0.47	0.44	0.40	1
<i>PMI_exp</i>	0.40	0.49	0.52	0.52	0.47	0.41	0.34	-1
<i>con_conf</i>	0.40	0.40	0.39	0.37	0.33	0.27	0.22	-3
<i>con_exp</i>	0.40	0.40	0.39	0.37	0.33	0.27	0.21	-2
<i>inv_conf</i>	-0.18	-0.15	-0.13	-0.09	-0.06	-0.05	-0.05	-3

注:各指标的含义可见表2,第2列~第8列代表时差相关系数,第9列代表时差阶数.

为了探究首席经济学家信心指数和其他经济主体信心作为景气指标的相对优劣,本研究的基准循环变量选为宏观经济一致指数<sup>⑤</sup>(conincident index).候选指标除了首席经济学家信心总指标及7个子指标,还含有消费者、投资者和采购经理三大经济主体的信心指标集.

分析结果如表4所示.首席经济学家信心指

数系列的所有指标时差阶数都为零或者为负,说明在景气指示方面具有一致性或者领先性.具有一致性的指标包括总指数和产出、风险、政策等几个子指数,具有领先性的指标包括房地产、国际、金融和物价等几个子指数,均领先1期~2期.此外,以上所有指标在0期~2期的不同领先期对应的时差相关系数差异不大,具有稳定性.从时差

⑤ 《中国经济景气月报》发布的经济一致指数是权威的月度经济监测景气指标,可以在月度频率反映宏观经济总体运行状况.

相关系数大小的角度出发,景气指示能力最强的是首席经济学家信心产出子指数,其次是首席经济学家信心总指数,和基准指标的相关系数分别为0.68和0.61.这一结果符合预期,因为基准指标为经济一致指数,主要反映经济总量产出,因此和产出子指数相关性更大.

涉及采购经理信心的指标包括:制造业采购经理人指数和新出口订单PMI,其中新出口订单PMI的时差阶数为领先1个月,时差相关系数为0.52,PMI的时差阶数为滞后1个月,时差相关系数为0.47.涉及消费者信心的指标包括:消费者信心指数、消费者预期指数.从评估结果中来看,消费者信心指数和预期指数的时差阶数分别为-3和-2,但是时差相关系数只有0.4.结果说明消费者信心指数系列的领先性较好,但是景气指示准确度欠佳.涉及投资者信心的指标为投资者信心指数.结果显示,投资者信心指数的景气指示性较差,时差相关系数均为负,且绝对值较小.

经过不同指标的时差阶数和时差相关系数对比,可以看出首席经济学家信心指数系列的景气领先性和景气指示性均较好,尤其是总指数和产出子指数,优于采购经理信心指数系列,更是远好于消费者和投资者的信心指数系列.因此首席经济学家信心指数系列可以作为有效的景气指标,扩充我国的经济景气监测体系.

#### 4.2 首席经济学家信心指数与宏观变量

首席经济学家信心指数系列是多元指标集,反映了经济专家对宏观经济不同侧面的观点.通过不同的信心子指数和不同宏观变量进行一对一的格兰杰因果检验,可以考察二者之间的样本内预测关系,也可以检验指数经济含义的合理性.

代表性宏观变量包括规模以上工业增加值、社会消费品零售总额、房地产开发新增固定资产、进出口金额、美元兑人民币、M2、CPI这7个主要宏观经济变量,指标形式为同比或累计同比,统计特征参见表2.从经济逻辑来看,这些指标全面地覆盖到了宏观经济运行的产出、消费、投资、进出口、货币金融、物价等主要侧面;从实践经验来看,翻阅诸多券商研报可以发现这些指标是政府和业界分析预测的主要对象.

本文通过格兰杰因果检验判断不同的首席经济学家信心指数对宏观变量的样本内预测作用.如果检验结果显著地拒绝原假设,说明在控制了宏观变量自身的滞后性之后,首席经济学家信心指数的滞后项仍然有助于预测宏观变量,体现了增量信息.根据表5格兰杰因果检验,全样本上首席经济学家信心指数集对宏观变量的预测具有增量信息.每个宏观变量都可以被总指数或者多个子指数所预测,规模以上工业增加值可以被“首席经济学家信心.产出”所预测;房地产开发新增固定资产可以被“首席经济学家信心.房地产”所预测;进出口金额和美元兑人民币都可以被“首席经济学家信心.国际”所预测;广义货币M2可以被“首席经济学家信心.金融”和“首席经济学家信心.政策”所预测;社会消费品零售总额可以被“首席经济学家信心.产出”和“首席经济学家信心.物价”所预测;CPI可以被“首席经济学家信心.物价”所预测.这些主要的对应关系都符合经济逻辑,说明信心指数确实科学地反映了其代表的经济侧面.此外,表5显示宏观变量还可能被经济对应关系之外的信心子指数预测,这是因为宏观经济具有系统性和复杂性,不同侧面的经济学家信心可能携带了共同的信息.

表5 首席经济学家信心指数对宏观变量的格兰杰因果检验

Table 5 Granger Causality Test of The Chief Economist Confidence Indexes on macroeconomic variables

变量	<i>IndOutput</i>	<i>RealEstate</i>	<i>Trade</i>	<i>Exchange</i>	<i>M2</i>	<i>CON</i>	<i>CPI</i>
<i>CECI</i>	4.85 ***	1.48	2.81 ***	1.29	2.35	1.72 *	1.28
<i>CECI_output</i>	5.49 ***	2.03	4.10 ***	3.64 **	3.41 **	4.41 ***	1.90 *
<i>CECI_housing</i>	1.77	1.91 **	1.31	1.47	1.29	1.60	0.45
<i>CECI_intl</i>	1.67	1.23	1.90 *	3.07 ***	2.47 **	1.27	2.46 **
<i>CECI_risk</i>	1.55	1.70	1.61	2.14 **	1.04	1.31	1.86 *
<i>CECI_finance</i>	1.88 *	1.60	2.37 **	0.63	2.32 **	2.68 ***	1.85 *
<i>CECI_policy</i>	2.48 **	1.96 *	2.34 **	0.81	2.96 **	2.48 ***	4.10
<i>CECI_price</i>	2.43 **	3.27 ***	2.60 **	1.27	0.54	2.57 **	1.76 *

注:表中数字代表经济学家信心指标对宏观变量的单向格兰杰因果检验F统计量,\*\*\*、\*\*、\*分别表示在1%、5%、10%水平下显著.

### 4.3 首席经济学家信心指数的样本外预测表现

首席经济学家信心指数和宏观变量之间的格兰杰因果检验考察了全样本上的变量对应关系和样本内预测能力,本研究进一步模拟真实预测场景,进行详尽的样本外预测检验以确认首席经济学家信心指数的信息增量.预测目标参照上节仍然选取为7个主要宏观经济变量.预测步长设置为提前0期~6期,涵盖了短期、中期到长期预测的不同期限.值得注意的是,提前0期代表现在预测(nowcasting),在月底时宏观变量的统计数据还没有发布,一般要到下月中旬才会发布,而本月的首席经济学家信心指数由于高频实时更新已经发布,因此可以用于预测当月的宏观变量.

为了验证首席经济学家信心指数在样本外宏观预测方面的信息增量,最关键的设定是对不同预测信息集的选取<sup>[35,36]</sup>.预测变量信息集一方面由滞后阶数 $k$ 决定,在月底进行预测时,宏观变量由于发布滞后,可用信息集包括1阶到 $k$ 阶滞后项,首席经济学家信心指数的可用信息集则包括当期到 $k$ 阶滞后项<sup>⑥</sup>.另一方面由覆盖的指标决定,基准模型中预测变量集只包括该宏观变量自身信息,对照模型既包括该宏观变量自身信息,还包括了不同数量的首席经济学家信心指数.对照模型1,包含宏观变量自身滞后+首席经济学家

信心的核心信心指数集(Xcore);对照模型2,包含宏观变量自身滞后+首席经济学家信心的子指数集(Xsmall);对照模型3,包含宏观变量自身滞后+首席经济学家信心的全指数集(Xall).核心信心指数集(Xcore)涉及产出、物价、政策3个信心子指数,这3个指标既反映了最重要的宏观经济侧面,而且在预测宏观变量的格兰杰因果检验中具有相对最好的效果.子指数集(Xsmall)涵盖了7个信心子指数,将所有侧面的信心指数都包括其中.全指数集(Xall)则尽可能使用所有指标,囊括了7个信心子指数和1个信心总指数.

考虑到可能的模型参数时变问题,本文采用滚动窗口<sup>⑦</sup>预测方法,测试集为2015-03—2022-12,每一步预测之后将一个新的观测样本加入训练集,并剔除最早的一个观测样本,再进行拟合和下一步预测.预测方式为适合高维自变量集的岭回归(Ridge)<sup>⑧</sup>.模型表现用样本外均方根误差RMSE来表示,并计算同时包含宏观变量自身信息和首席经济学家信心指标集的对照模型之于仅包含宏观变量自身信息的基准模型的相对误差.计算公式为:相对RMSE=对照模型RMSE/基准模型RMSE,结果小于1说明对照模型改进了基准模型的预测效果,数值越小,对照模型表现越好.实验结果如表6所示.

表6 样本外预测的相对误差

Table 6 Relative errors of out-of-sample predictions

预测目标	加入指数集	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$
IndOutput	Xcore	0.94	1.04	0.99	0.94	0.91	0.86**	0.88*
	Xsmall	1.03	1.03	1.00	0.95*	0.91	0.87**	0.85**
	Xall	1.06	1.04	1.00	0.92*	0.90*	0.86**	0.69
RealEstate	Xcore	0.94	0.73**	0.55	0.76*	0.89**	0.91**	0.77*
	Xsmall	0.90*	0.77**	0.53	0.72*	0.84**	0.78**	0.76*
	Xall	0.91*	0.74**	0.54	0.74*	0.85***	0.79***	0.74*
Trade	Xcore	0.84**	0.87**	0.71	0.81	0.89*	0.87**	0.85**
	Xsmall	0.82**	0.68	0.69	0.79*	0.85*	0.79**	0.81**
	Xall	0.86**	0.69	0.71	0.78*	0.85*	0.82**	0.81**
Exchange	Xcore	1.19	0.99	0.93	0.95	0.89	0.99	1.01
	Xsmall	1.04	0.86*	0.80**	0.84*	0.81*	0.90	0.92
	Xall	1.01	0.84**	0.80**	0.83**	0.79*	0.89	0.90

⑥ 取滞后项时会损失样本量,为了在充分利用滞后信息和尽量保留可用样本之间取得平衡,将 $k$ 设置为6.

⑦ 滚动窗口长度设置为12个月.

⑧ 使用10折验证法选取调节参数 $\lambda$ .

续表6

Table 6 Continues

预测目标	加入指数集	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$
<i>M2</i>	Xcore	0.96	0.90*	0.86*	0.88	0.98	1.07	1.17
	Xsmall	0.88**	0.81***	0.78***	0.80**	0.91	1.03	1.13
	Xall	0.87**	0.81***	0.80**	0.78*	0.93	0.99	1.12
<i>CON</i>	Xcore	0.61	0.50	0.48	0.65*	0.63	1.03	1.00
	Xsmall	0.68*	0.71*	0.58	0.59*	0.46	1.01	1.07
	Xall	0.65*	0.74*	0.51	0.60*	0.91	0.98	0.92
<i>CPI</i>	Xcore	1.14	0.89	0.82**	0.81**	0.85**	0.97	1.12
	Xsmall	1.05	0.82**	0.75**	0.70**	0.73**	0.86	0.98
	Xall	1.05	0.81**	0.74**	0.73**	0.75**	0.87	0.99

注：该表的数字表示不同预测目标、不同预测步长( $h$ )、加入不同信心指数集的场景下,不同对照模型之于基准模型的相对预测误差.自变量集中宏观变量和首席经济学家信心指数的最大滞后阶数 $k=6$ .结果小于1说明对照模型改进了基准模型的预测效果,数值越小,对照模型表现越好.使用DM-test检验预测性能的改进是否统计显著,\*\*\*、\*\*、\*分别表示在1%、5%、10%水平下显著.

加入不同数量的首席经济学家信心指数集之后,不同宏观变量在不同提前期上的预测效果被显著地改善.大部分情况下,全指数集的预测改进效果最好,侧面论证了首席经济学家信心指数集的信息含量较为丰富.规模以上工业增加值在提前3个月~6个月具有最好的预测改善效果,核心指数集在提前5个月~6个月的远期预测效果最好,全指数集在提前3个月~5个月的中远期预测效果最好.预测误差改善程度大部分超过10%,说明首席经济学家信心指数对规模以上工业增加值主要是起到中远期预测作用.房地产开发新增固定资产的预测误差在0个月~6个月的不同提前期都有改善.子指数集和全指数集的改进预测效果比核心指数集更好,在提前1月预测和提前3个月~6个月预测显著改进20%~30%的预测误差,在当月预测显著改进10%左右的预测误差.进出口金额在0个月~1个月的临近预测和4个月~6个月的远期预测改进效果最好.核心指数集在临近预测表现更优,子指数集和全指数集在远期预测效果更优,预测误差改进程度分布在10%~20%之间.美元兑人民币汇率在提

前1个月~4个月的中期预测改进效果较好,且全指数集的效果最好,预测误差改进20%左右.*M2*在提前0个月~3个月的近期预测效果较好,全指数集效果最好,预测误差改进20%左右.社会消费品零售总额在0个月~1个月的近期预测和提前3个月的中期预测效果较好,全指数集效果最好,预测误差改进程度分布在30%~40%之间.*CPI*在提前1个月~4个月的中期预测效果较好,全指数集效果最好,预测误差改进程度分布在20%~30%之间.

为了保证结果的可信度,本文进一步进行稳健性检验,加入全信心指数集(Xall)后,将预测模型从Ridge变为Lasso和elasticNet,或将滞后阶数从6减少到3.表7结果显示,更换模型和改变滞后阶数之后,规模以上工业增加值的预测改善有所弱化,美元兑人民币汇率的预测改善却有所增强,但是主要结果仍然稳健.以上结果说明,在模拟现实场景的样本外预测中,即使控制同样阶数的宏观变量滞后项,加入首席经济学家信心指数集仍然能够显著增加有效信息、改善预测效果.

表7 样本外预测的相对误差:稳健性

Table 7 Relative errors of out-of-sample predictions: Robustness

预测目标	模型和参数	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$
<i>IndOutput</i>	Lasso( $k=6$ )	1.25	1.23	1.20	1.27	1.03	0.97	0.64
	elasticNet( $k=6$ )	1.19	1.07	1.15	1.25	1.03	0.96**	0.66
	Ridge( $k=3$ )	0.98	1.03	1.04	1.04	1.05	1.01	0.94*

续表7

Table 7 Continues

预测目标	模型和参数	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$
RealEstate	Lasso( $k=6$ )	0.93	0.44*	0.50*	0.44	0.42	0.63**	0.75**
	elasticNet( $k=6$ )	0.93	0.44*	0.45**	0.86*	0.76**	0.72***	0.72**
	Ridge( $k=3$ )	1.06	0.90*	0.68*	0.63**	0.67**	0.70**	0.70*
Trade	Lasso( $k=6$ )	0.93	0.66	0.74	0.94	0.83**	0.84**	0.80**
	elasticNet( $k=6$ )	0.91	0.62	0.68	0.78	0.80**	0.88**	0.75**
	Ridge( $k=3$ )	0.86	0.86**	0.85	0.76	0.88	0.82*	0.93
Exchange	Lasso( $k=6$ )	0.94	0.81**	0.62***	0.63***	0.51**	0.62***	0.66**
	elasticNet( $k=6$ )	0.91	0.77***	0.61***	0.62***	0.52**	0.64***	0.64***
	Ridge( $k=3$ )	1.28	0.91	0.87	0.85*	0.78**	0.83***	0.84***
M2	Lasso( $k=6$ )	0.86**	0.70***	0.67***	0.77***	0.86	0.95	1.01
	elasticNet( $k=6$ )	0.89*	0.74***	0.67***	0.78***	0.86*	0.92	0.96
	Ridge( $k=3$ )	1.03	0.83**	0.82***	0.77***	0.78***	0.77**	0.93
CON	Lasso( $k=6$ )	0.76**	1.08	0.72	0.90	1.16	1.53	1.32
	elasticNet( $k=6$ )	0.66	1.15	0.72	0.86*	1.31	1.22	2.12
	Ridge( $k=3$ )	0.82**	0.71**	0.81*	0.53	0.43	0.92	1.01
CPI	Lasso( $k=6$ )	1.05	0.67**	0.68**	0.61**	0.65**	0.77	0.76
	elasticNet( $k=6$ )	0.93	0.70**	0.64**	0.61**	0.65**	0.76*	0.78
	Ridge( $k=3$ )	1.32	0.97	0.86**	0.78**	0.78**	0.74**	0.76**

注：该表的数字表示不同预测目标、不同预测步长( $h$ )、加入全信心指数集(Xall)后、使用不同模型和参数的场景下，对照模型之于基准模型的相对预测误差。结果小于1说明对照模型改进了基准模型的预测效果，数值越小，对照模型表现越好。使用DM-test 检验预测性能的改进是否统计显著，\*\*\*、\*\*、\* 分别表示在1%、5%、10%水平下显著。

## 5 结束语

近年来，受到国内外复杂形势的影响，我国经济发展的不确定性升高、短期波动显著。同时，经济主体信心的监测和引导也成为复杂形势下宏观调控工作的一个抓手。经济专家的信心水平对于监测主体信心状态、维护经济运行、预测宏观经济态势具有重大现实意义和政策启示。本文利用TextRank和FinBERT的人工智能自然语言处理方法，借助“首席经济学家论坛”在线文本大数据，构建了首席经济学家信心指数和子类信心指数系列(CECI)。该指数系列可以在月度频率及时更新，清晰地反映经济学家对经济不同侧面的信心变化，具有良好的经济景气指示作用，可以改进对重要宏观变量的预测效果，是政府利用海量大数据实现科学、精准和高效决策的有力工具。

研究发现：第一，首席经济学家信心指数系列具有良好的代表性。首席经济学家信心总指数和

官方的季度经济学家信心指数的基本走势一致，产出信心子指数和官方经济学家信心指数的相关性高于0.7(1%水平显著)、和GDP的相关性则高于0.5(在1%水平显著)。和官方信心指数相比，基于在线文本的首席经济学家信心指数系列更新频率更高、构造成本更低、时效性更强，在识别高频经济变化时优于官方的季度指数；第二，首席经济学家信心指数系列是有效的经济景气指标。时差相关性检验显示，其景气领先性和景气指示性均较好，尤其是总指数和产出子指数，优于采购经理信心指数系列，更是远好于消费者和投资者的信心指数系列。首席经济学家信心指数系列可以作为有效的景气指标，扩充我国的经济景气监测体系；第三，首席经济学家信心指数系列可以改进对宏观变量的预测效果，包括规模以上工业增加值、社会消费品零售总额、房地产开发新增固定资产、进出口金额、美元兑人民币、M2、CPI这7个重要宏观经济变量。全样本上的格兰杰因果检验显示，不同的信心子指数对相应的宏观变量具有增量预测作用，符合经济对应逻辑。基于不同

预测目标、不同预测步长、不同信息集的样本外滚动窗口预测对比实验显示,加入首席经济学家信心指数能显著地改善对这些宏观变量的预测效果,对不同变量的预测改善效果分布在10%~

40%之间.最后,本研究结合了自然语言处理的文本大数据技术和景气指数理论,也对更多基于在线大数据的景气指数构建和经济预测实践具有借鉴意义.

### 参考文献:

- [1]梁方,沈诗涵,黄卓.预测中国宏观经济变量:专家与模型的组合预测[J].金融研究,2021,(7):58-76.  
Liang Fang, Shen Shihan, Huang Zhuo. Forecasts of macroeconomic variables in China: Combination forecasts of surveys and models[J]. Journal of Financial Research, 2021, (7): 58-76. (in Chinese)
- [2]郑京平.中国宏观经济景气监测指数体系研究[M].北京:中国统计出版社,2013:90-134.  
Zheng Jingping. Study of China's Economic Cycles Monitoring Indicators System[M]. Beijing: China Statistics Press, 2013: 90-134. (in Chinese)
- [3]高铁梅,陈磊,王金明,等.经济周期波动分析与预测方法[M].北京:清华大学出版社,2015:197-219.  
Gao Tiemei, Chen Lei, Wang Jinming, et al. Methods of Economic Cycle Fluctuation Analysis and Forecast[M]. Beijing: Tsinghua University Press, 2015: 197-219. (in Chinese)
- [4]Lehmann R. The forecasting power of the ifo business survey[J]. Journal of Business Cycle Research, 2023, 19(1): 43-94.
- [5]Baker S R, Bloom N, Davis S J. Measuring economic policy uncertainty[J]. The Quarterly Journal of Economics, 2016, 131(4): 1593-1636.
- [6]Huang Y, Luk P. Measuring economic policy uncertainty in China[J]. China Economic Review, 2020, (59): 101367.
- [7]Antweiler W, Frank M Z. Is all that talk just noise? The information content of internet stock message boards[J]. The Journal of Finance, 2004, 59(3): 1259-1294.
- [8]部慧,解峥,李佳鸿,等.基于股评的投资者情绪对股票市场的影响[J].管理科学学报,2018,21(4):86-101.  
Bu Hui, Xie Zheng, Li Jiahong, et al. Investor sentiment extracted from internet stock message boards and its effect on Chinese stock market[J]. Journal of Management Sciences in China, 2018, 21(4): 86-101. (in Chinese)
- [9]罗琦,张志达,吴希梅,等.股利情绪、股利迎合与股价崩盘风险——基于百度指数平台搜索量的经验证据[J].管理科学学报,2023,26(2):87-103.  
Luo Qi, Zhang Zhida, Wu Ximei, et al. Dividend sentiment, dividend catering and stock price crash risk: Evidence based on searching volume of Baidu Index Platform[J]. Journal of Management Sciences in China, 2023, 26(2): 87-103. (in Chinese)
- [10]Li F. The information content of forward-looking statements in corporate filings: A Naïve Bayesian machine learning approach[J]. Journal of Accounting Research, 2010, 48(5): 1049-1102.
- [11]Jiang F, Lee J, Martin X, et al. Manager sentiment and stock returns[J]. Journal of Financial Economics, 2019, 132(1): 126-149.
- [12]王克敏,王华杰,李栋栋,等.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,2018,34(12):120-132,194.  
Wang Kemin, Wang Huajie, Li Dongdong, et al. Complexity of annual report and management self-interest: Empirical evidence from Chinese listed firms[J]. Management World, 2018, 34(12): 120-132, 194. (in Chinese)
- [13]胡楠,薛付婧,王昊楠.管理者短视主义影响企业长期投资吗?——基于文本分析和机器学习[J].管理世界,2021,37(5):139-156,11,19-21.  
Hu Nan, Xue Fujing, Wang Haonan. Does managerial myopia affect long-term investment?: Based on text analysis and machine learning[J]. Management World, 2021, 37(5): 139-156, 11, 19-21. (in Chinese)
- [14]成程,杨胜刚,田轩.金融科技赋能下供应链金融对企业价值的影响[J].管理科学学报,2024,27(2):

- 95 - 119.
- Cheng Cheng, Yang Shenggang, Tian Xuan. The impact of supply chain finance on firm value in the context of fintech empowerment[J]. *Journal of Management Sciences in China*, 2024, 27(2): 95 - 119. (in Chinese)
- [15] Moore G H. *Statistical Indicators of Cyclical Revivals and Recessions*[M]. Princeton: Princeton University Press, 1961.
- [16] Stock J H, Watson M W. New indexes of coincident and leading economic indicators[J]. *NBER Macroeconomics Annual*, 1989, (4): 351 - 394.
- [17] Aruoba S B, Diebold F X, Scotti C. Real-time measurement of business conditions[J]. *Journal of Business & Economic Statistics*, 2009, 27(4): 417 - 427.
- [18] Bok B, Caratelli D, Giannone D, et al. Macroeconomic nowcasting and forecasting with big data[J]. *Annual Review of Economics*, 2018, 10(1): 615 - 643.
- [19] Lewis D J, Mertens K, Stock J H, et al. Measuring real activity using a weekly economic index[J]. *Journal of Applied Econometrics*, 2022, 37(4): 667 - 687.
- [20] Wegmüller P, Glocker C, Guggia V. Weekly economic activity: Measurement and informational content[J]. *International Journal of Forecasting*, 2023, 39(1): 228 - 243.
- [21] 林 娟, 陈海强, 林 青. 基于模型平均的中国产出增长和通货膨胀密度预测[J]. *管理科学学报*, 2024, 27(2): 82 - 94.
- Lin Juan, Chen Haiqiang, Lin Qing. Forecasting China's output growth and inflation: A model averaging approach[J]. *Journal of Management Sciences in China*, 2024, 27(2): 82 - 94. (in Chinese)
- [22] 陈 磊, 高铁梅. 利用 Stock-Watson 型景气指数对宏观经济形势的分析和预测[J]. *数量经济技术经济研究*, 1994, (5): 53 - 59.
- Chen Lei, Gao Tiemei. Analysis and forecasting of macroeconomic conditions using Stock-Watson Type Business Cycle Index[J]. *Journal of Quantitative & Technical Economics*, 1994, (5): 53 - 59. (in Chinese)
- [23] 刘 汉, 刘金全. 中国宏观经济总量的实时预报与短期预测——基于混频数据预测模型的实证研究[J]. *经济研究*, 2011, 46(3): 4 - 17.
- Liu Han, Liu Jinquan. Nowcasting and short-term forecasting of Chinese macroeconomic aggregates: Based on the empirical study of MIDAS model[J]. *Economic Research Journal*, 2011, 46(3): 4 - 17. (in Chinese)
- [24] 郑挺国, 王 霞. 中国经济周期的混频数据测度及实时分析[J]. *经济研究*, 2013, 48(6): 58 - 70.
- Zheng Tingguo, Wang Xia. Measuring China's business cycle with mixed-frequency data and its real time analysis[J]. *Economic Research Journal*, 2013, 48(6): 58 - 70. (in Chinese)
- [25] 费兆奇, 刘 康. 中国宏观经济波动的高频监测研究——基于混频模型对日度经济先行指数的构建和分析[J]. *管理世界*, 2019, 35(6): 27 - 38.
- Fei Zhaoqi, Liu Kang. Research on high frequency monitoring of macroeconomic fluctuation in China: Daily Economic Leading Index construction and analysis based on mixed frequency model[J]. *Management World*, 2019, 35(6): 27 - 38. (in Chinese)
- [26] 洪永淼, 汪寿阳. 大数据如何改变经济学研究范式? [J]. *管理世界*, 2021, 37(10): 40 - 55, 72, 56.
- Hong Yongmiao, Wang Shouyang. How is big data changing economic research paradigms? [J]. *Management World*, 2021, 37(10): 40 - 55, 72, 56. (in Chinese)
- [27] Lin W, Wei Y. Economic forecasting with big data: A literature review[J]. *Journal of Management Science and Engineering*, 2024, 9(2): 254 - 270.
- [28] Xin B, Jiang K. Economic uncertainty, central bank digital currency, and negative interest rate policy[J]. *Journal of Management Science and Engineering*, 2023, 8(4): 430 - 452.
- [29] 欧阳资生, 陈世丽, 杨希特, 等. 经济政策不确定性、网络舆情与金融机构系统性风险[J]. *管理科学学报*, 2023, 26(4): 62 - 86.
- Ouyang Zisheng, Chen Shili, Yang Xite, et al. Economic policy uncertainty, network public opinion and systemic risk of financial institutions[J]. *Journal of Management Sciences in China*, 2023, 26(4): 62 - 86. (in Chinese)
- [30] 谢海华, 陈雪飞, 都仪敏, 等. 结合统计特征和图模型的半监督式中文关键词抽取方法[J]. *中文信息学报*,

- 2022, 36(4): 57–65.
- Xie Haihua, Chen Xuefei, Du Yimin, et al. A semi-supervised method for Chinese key phrase extraction based on statistical features and graph model[J]. *Journal of Chinese Information Processing*, 2022, 36(4): 57–65. (in Chinese)
- [31] Mihalcea R, Tarau P. TextRank: Bringing Order into Text[C]. *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, Spain, Barcelona: Association for Computational Linguistics*, 2004: 404–411.
- [32] 刘涛雄, 汤珂, 姜婷凤, 等. 一种基于在线大数据的高频 CPI 指数的设计及应用[J]. *数量经济技术经济研究*, 2019, 36(9): 81–101.
- Liu Taoxiong, Tang Ke, Jiang Tingfeng, et al. Design and application of novel CPI based on online big data[J]. *Journal of Quantitative & Technical Economics*, 2019, 36(9): 81–101. (in Chinese)
- [33] 梁言顺. 经济景气预测方法的应用[J]. *数量经济技术经济研究*, 1996, (2): 59–62, 69.
- Liang Yanshun. Application of economic climate forecasting methods[J]. *Journal of Quantitative & Technical Economics*, 1996, (2): 59–62, 69. (in Chinese)
- [34] Goldfarb A, Tucker C. Digital economics[J]. *Journal of Economic Literature*, 2019, 57(1): 3–43.
- [35] 姜婷凤, 汤珂, 刘涛雄. 基于在线大数据的通货膨胀“现时”预测[J]. *计量经济学报*, 2022, 2(3): 597–619.
- Jiang Tingfeng, Tang Ke, Liu Taoxiong. Nowcasting the inflation with online big data[J]. *China Journal of Econometrics*, 2022, 2(3): 597–619. (in Chinese)
- [36] 刘涛雄, 徐晓飞. 互联网搜索行为能帮助我们预测宏观经济吗?[J]. *经济研究*, 2015, 50(12): 68–83.
- Liu Taoxiong, Xu Xiaofei. Can internet search behavior help to forecast the macro economy?[J]. *Economic Research Journal*, 2015, 50(12): 68–83. (in Chinese)

## Construction and macroeconomic forecasting of the chief economist confidence indexes based on text analysis

ZHANG Tao<sup>1</sup>, TANG Ke<sup>1</sup>, XIE Hai-hua<sup>2</sup>

1. Institute of Economics School of Social Sciences, Tsinghua University, Beijing 100084, China;
2. Yanqi Lake Beijing Institute of Mathematical Sciences and Applications, Beijing 101408, China

**Abstract:** Economists' comprehensive judgments and confidence levels are increasingly valuable in today's environment of heightened economic fluctuations and uncertainty. However, existing economist confidence indexes are not sufficiently regarded due to their low frequency, high compilation costs, and lack of timeliness. This study constructs a monthly Chief Economist Confidence Index and its sub-indexes (CECI) by leveraging online text data from the "Chief Economists Forum" and employing cutting-edge natural language processing technologies, specifically the TextRank + FinBERT method. The paper finds that the CECI trends consistently with the National Bureau of Statistics' Quarterly Economist Confidence Index, but features stronger timeliness and a higher update frequency. Compared to the confidence indexes of other economic entities, the confidence of chief economists serves as a more effective indicator of the business cycle. In terms of macroeconomic forecasting, the CECI series can significantly enhance the out-of-sample forecasting performance for key macroeconomic variables. This study represents a valuable attempt to construct business cycle indicators and conduct macroeconomic forecasts using artificial intelligence methods. The methodology can be applied to the construction of other high-frequency business cycle indexes.

**Key words:** business cycle indexes; economist confidence; natural language processing; macroeconomic forecasting