

# 央行沟通有助于改善宏观经济预测吗?<sup>\*</sup>

## ——基于文本数据的高维稀疏建模

林建浩 陈良源 罗子豪 张一帆

**内容提要:** 宏观经济预测是宏观调控精准施策的重要前提,一直以来是方法论研究的前沿议题。随着央行沟通在预期管理中的频繁使用,其传达的信息受到普遍关注。本文致力于利用央行沟通文本进行宏观经济预测。首先生成符合央行沟通表达习惯的专用词典用于构建完整语料库,继而利用栅栏分布式多项回归模型从高维和稀疏的语料库中提取有效信息,得到央行沟通测度。基于152个指标构建基准动态因子模型,进一步引入央行沟通测度作为新的预测因子,结果显示央行沟通测度有助于提升模型样本内拟合效果。考察样本外预测效果,在不包括预测变量历史信息时,央行沟通测度能够使得不同期限的预测精度提高6.80%—16.65%;包含预测变量历史信息时则出现分化,在期限较短时,央行沟通未能提升预测精度,这是因为主要沟通信息与预测变量历史信息重叠;当期限较长时,预测精度有所提升,表明沟通中少量的前瞻性指引具有持续的预测能力。本文研究从预测角度验证了中国央行沟通在预期管理中的作用,并为进一步利用非结构化的文本大数据提升中国宏观经济实时预测能力提供了新思路。

**关键词:** 央行沟通 文本分析 宏观经济预测 栅栏分布式多项回归模型

## 一、引言

宏观经济预测不仅是重要的经济学方法论问题,更是政策实践的重要挑战。对于中国而言,无论是长期性的结构改革,还是周期性的宏观调控,其精准施策都建立在科学的宏观经济预测之上。然而,GDP、CPI等核心宏观经济指标官方数据存在时效性和更新频率的局限。虽然大数据为经济学研究带来了新变革,但是大数据往往具有专属性甚至垄断性,这阻碍了其在宏观经济预测上的理论探索和实践应用。在各类大数据中,文本数据的可获得性较高,成为经济学中发展最快的非结构化数据。诺贝尔奖得主 Shiller(2017, 2019) 提倡的叙事经济学(narrative economics)认为,叙事能够影响人们的决策乃至改变整个经济和社会的走向,因此这些信息可以用于预测金融危机、萧条和其他经济事件。信息驱动经济周期理论(news-driven business cycle)认为公众预期是影响经济波动的关键因素,而信息则是形成预期的独立驱动力量(Beaudry & Portier, 2014)。文本数据作为叙事和信息的载体,已经广泛应用于刻画经济政策不确定性、行业动态分类、关注度、情绪、隐含波动率和意见不一致性等,详细综述可参考 Gentzkow et al. (2019) 和沈艳等(2019)。那么,文本数据是否可用于宏观经济预测呢?

在众多文本数据中,央行沟通文本具有明确而独特的价值。央行沟通指央行向市场传递货币政策目标与规则、经济形势判断以及前瞻性指引等信息,央行特殊的地位决定了其释放信息的重要

<sup>\*</sup> 林建浩、陈良源(通讯作者)、罗子豪、张一帆,中山大学岭南学院,邮政编码:510275,电子邮箱:linjh3@mail.sysu.edu.cn; chenly58@mail2.sysu.edu.cn, luozh35@mail2.sysu.edu.cn, zhangyf29@mail2.sysu.edu.cn。本研究得到国家自然科学基金项目(71773147; 72073148; 71991474)、中山大学中央高校基本科研业务费项目(19wkzd09)的支持。作者感谢匿名审稿专家的宝贵意见,文责自负。

性(Blinder et al. 2008)。随着货币政策透明化趋势和预期管理思潮的形成,央行沟通成为重要的新型货币政策工具(Bernanke 2020)。McMahon et al. (2018)指出,尽管中国央行沟通受到的限制更多,但是中国央行仍负责起草并执行货币政策,通过公开市场操作或常备及中期贷款便利来设定政策利率,具有操作独立性。目前中国的央行沟通包括书面沟通和口头沟通两种渠道。书面沟通主要包括每季度定期发布的《货币政策执行报告》和不定期的会议纪要,其语言组织规范、信息传达全面;口头沟通主要包括央行官员或发言人以讲话、采访、新闻发布会、时报文档等形式开展的沟通,其沟通频率更高、时效性更强,对市场的影响力也更强(吴国培和潘再见,2014;林建浩等,2019)。

央行沟通文本之所以具备预测宏观经济的潜力,是因为其在预期管理中的核心作用。<sup>①</sup> 市场根据央行释放的信息对未来货币政策和宏观经济形成判断和预期,进而改变消费、投资决策(Morris & Shin 2008; Hansen & McMahon 2016)。Eusepi & Preston(2010, 2011)研究发现央行公布其货币政策规则有助于改进经济主体对于经济运行的理解,稳定通胀预期;缺乏信息披露可能导致经济在繁荣和萧条中反复波动,充分的央行信息披露有助于维持经济稳定。

具体而言,央行沟通可以通过创造信息、减少噪声和信息效应的期限结构三个信息渠道影响经济主体的预期和行为。第一,央行具有独特的研究能力和信息优势,<sup>②</sup>可以直接“创造新信息”(creating news)影响市场预期(Larsen & Thorsrud 2019)。Peek et al. (2003)强调美联储的预测能力正是来自于长期的监管信息优势,尤其是非公开的监管信息可以用于预测通胀和失业率。中国央行同样对重要经济议题投入了大量的研究资源(陈雨露,2019),而且针对城镇储户、企业家、银行家等进行定期大规模问卷调查,直接获取经济活动的一手数据。第二,央行沟通具有“减少噪声”(reducing noise)功能,能够提高未来货币政策的可预测性(陈良源等,2021),引导公众对物价水平和利率的预期趋同(Lamla and Vinogradov 2019),有利于降低市场波动,最终影响宏观经济的运行。第三,央行沟通的信息效应具有期限结构,能够对不同期限的利率产生差异化影响(Hanson & Stein 2015; Hansen et al. 2019)。利率期限结构反映了长短期利率水平之间的关系,是市场对当前经济情况判断以及未来经济走势预期的均衡结果,央行沟通尤其是前瞻性指引会促使投资者调整其长期投资组合,进而影响长期信贷成本。自2008年金融危机以来,美联储、欧央行、日本央行等主要央行都出现了零利率下限问题,以前瞻性指引为代表的央行沟通成为为数不多的可操作工具。尽管中国尚未面临零利率下限问题,但是中国央行日渐丰富的沟通实践对金融市场和宏观经济调控均有重要影响(Su et al. 2019; Sun 2020)。

央行沟通渠道多样,文本卷帙浩繁,如何从中提取有效信息面临着三大难题。第一,为保证全面性,必须解决文本的高维建模问题。早期的主观赋值法、关键词计数法以及基于通用词典的情感指数法简单快捷,<sup>③</sup>但对于词语个数通常数以万计的央行沟通文本语料库而言,这些方法所考虑的词语范围均十分有限,较低的词语利用率可能遗漏许多影响宏观变量的重要信息。为充分利用文本信息,必须处理词语的高维问题。一种思路是利用与计算机语言相结合的自然语言处理方法,其核心思想是通过聚类降维,应用较多的是LDA方法提取主题(Thorsrud 2020; Hansen et al. 2019);

① 除了央行沟通影响公众预期的形成,还需预期的自我实现环节,但后者不是本文研究重点,因此未展开详细综述与讨论。

② Trichet(2007)总结了欧洲央行的四个研究动机:追踪最新的经济理论、弥补学界和政策制定者之间分歧、保证政策制定依赖工具的有效性、掌握尚未研究议题的相关知识,并因为这四个动机而在学术研究中投入资源。

③ Romer & Romer(1989)将货币政策决策沟通内容所传达的政策倾向进行离散赋值;Jansen & De Haan(2007)通过计算央行沟通中“vigilance”这一单词出现的频率来量化价格波动风险;通用词典的测度方法通常借助哈佛社会心理学词典或者由Loughran & McDonald(2011)构造的LM金融词典,分别计算文本数据中出现正面词语及负面词语的频率,从而提取情感指数,类似的方法也为Hansen et al. (2018)等所应用。

另一种思路是用分布式多项回归( distributed multinomial regression ,DMR) 这一微观计量模型模拟文档生成过程中的高维选词决策( Taddy 2013 2015) 。第二,为保证精确性,必须解决文本的稀疏性问题。稀疏性往往与高维性相伴而生,即在某些较短的文档中所出现的所有词语,通常仅占高维整体语料库的很小一部分。例如在 LDA 模型中,文本的稀疏性将使得缺乏足够信息以学习词与词之间的关系,有限的上下文也难以识别简短文档中的歧义词( Rashid et al. 2019) 。尤其是央行的口头沟通文本较短,稀疏性将使得 LDA 的主题学习不够稳定和精确。又如,DMR 虽然可以直接对高维的原始词汇进行分析,但 Kelly et al. ( 2019) 指出 DMR 所使用的泊松分布仅对正词频部分有很好的近似,而对包含众多零词频的文本数据并不适用。而对于某次央行沟通而言,由于总是使用整体语料库中的小部分词汇,因此大多数词汇的词频为 0,这种文本稀疏性将导致 DMR 所拟合出的词频概率分布不够精确,进而使得预测精度大大降低。第三,为保证专业性,必须解决自建词典问题。央行沟通具有很强的专业性,直接套用现有软件的中文分词功能,往往不能体现完整的信息,例如“货币供应量”将被分为“货币”和“供应量”两个词,不利于建立起央行沟通文本与预测变量之间的机器学习模式。为此,需要自建囊括专业术语和固定短语搭配的专用词典,才能准确分词完成语料库构建。

本文致力于解决央行沟通文本分析的难题,并将其运用于宏观经济预测。首先,在文本分词过程中,本文考虑中国央行沟通的特殊表达习惯和专有名词用法,自建央行沟通用户词典。自建词典与通用词典的结合实现了精准而有效的分词,从而构建了央行沟通的完整语料库。其次,本文通过 Kelly et al. ( 2019) 提出的栅栏分布式多项回归模型( HDMR) 解决了文本数据的高维性和稀疏性问题,成功提取出语料库中与预测核心宏观经济变量紧密相关的央行沟通测度指标。最后,动态因子模型是宏观预测领域的基准模型,本文在其基础上引入央行沟通测度作为新的预测因子,讨论央行沟通文本数据对中国核心宏观变量的预测能力。本文研究与两支文献紧密相关并做出边际贡献。第一支文献是中国的央行沟通与预期管理。央行信息披露对居民通胀预期具有引导作用( 卞志村和张义 2012) ,同时能够减少通胀及其预期的波动性( 汪莉和王先爽,2015) 。郭豫媚和周璇( 2018) 发现央行沟通能够加快市场主体向理性预期均衡收敛的适应性学习速度,提高货币政策有效性。不同于以上研究关注央行沟通对预期变量的影响,本文从央行沟通是否有助于提高宏观经济预测这一角度验证了预期管理的有效性,在研究视角上有所创新。第二支文献是大数据与中国宏观经济预测。随着信息技术和统计手段的发展,大数据分析对研究产生了显著的影响,一个重要趋势是高维数据和实时数据在宏观经济预测中的作用越来越大。刘汉和刘金全( 2011) 利用 MIDAS 模型实现了中国宏观经济总量的实时预报和短期预测,并指出忽略高频数据信息将影响预测的及时性和精确度;郑挺国和王霞( 2013) 则是基于动态因子模型实现了中国经济周期指数的测度;刘涛雄和徐晓飞( 2015) 探究如何利用互联网搜索大数据对宏观经济进行预测。沿着大数据时代下宏观经济预测文献的发展脉络,本文利用高维稀疏建模方法提取央行沟通这一重要文本的关键信息并用于改进宏观经济预测,为利用非结构化的文本大数据进行实时预测提供了新思路。

后文安排如下:第二部分是央行沟通文本的高维稀疏建模与数据处理;第三部分是宏观经济数据的收集和预测模型设定;第四部分是实证分析;最后是主要结论和启示。

## 二、央行沟通文本的高维稀疏建模

### (一) 文本数据的高维稀疏建模

Kelly et al. ( 2019) 提出栅栏分布式多项回归模型( hurdle distributed multinomial regression , HDMR) 同时刻画词语选择和词语计数,尤其是通过词语选择方程很好地解决文本稀疏性所导致的

近似效果差的问题。因此,本文基于 HDMR 模型进行文本数据的高维稀疏建模,从中国央行沟通文本中提取关键信息。模型具体如下:

$$c_{ij} = (1 + c_{ij}^*) h_{ij} \quad (1)$$

$$c_{ij}^* = \lambda(\mu_i + \alpha_j + v_i' \varphi_j) + \varepsilon_{ij} \quad (2)$$

$$h_{ij} = 1(h_{ij}^* > 0) \quad (3)$$

$$h_{ij}^* = \gamma_i + \kappa_j + w_i' \delta_j + v_{ij} \quad (4)$$

其中  $i=1, 2, \dots, n$  表示文档数  $j=1, 2, \dots, d$  表示词语数,文本数据中  $d$  的维度一般非常高。

(1) 式和(2) 式描述了词语的计数模型  $c_{ij}$  是从语料库中统计而得的词频,表示在第  $i$  篇文档中词语  $j$  出现的次数。模型将词频划分为是否选取  $h_{ij}$  和使用频率  $(1 + c_{ij}^*)$  两部分  $h_{ij}$  表示词语  $j$  在文档  $i$  中是否出现  $c_{ij}^*$  表示词语  $j$  在文档  $i$  中词频超过 1 的部分,采用刻画计数数据的泊松分布设定  $\lambda(\cdot)$ 。在(2) 式表示的泊松回归模型中  $\mu_i$  为影响词语频率的协变量集合  $\varphi_j$  表示  $v_i$  对词语  $j$  的计数的影响程度  $\mu_i$  和  $\alpha_j$  分别表示文档  $i$  和词语  $j$  中与协变量  $v_i$  无关的固有特征。 $c_{ij}^*$  与  $c_{ij}$  的区分使得泊松分布设定仅针对词语  $h_{ij}$  频率不少于 1 的部分,确保词频取值为 0 的概率只由  $h_{ij}$  决定。实际数据中,只要词语  $j$  出现则词频  $c_{ij}$  大于等于 1,但是(2) 式的泊松分布使得  $c_{ij}^*$  从 0 开始计数,因此需要  $(1 + c_{ij}^*)$  的设定。

(3) 式和(4) 式描述了词语的选择模型,即决定在第  $i$  篇文档中是否加入词语  $j$ 。 $h_{ij}$  由二值选择模型对应的潜变量  $h_{ij}^*$  决定; $w_i$  为影响词语选择的协变量集合; $\delta_j$  表示  $w_i$  对词语  $j$  的选择的影响程度; $\gamma_i$  和  $\kappa_j$  分别表示文档  $i$  和词语  $j$  中与协变量  $w_i$  无关的固有特征。模型允许(2) 式与(4) 式中的协变量  $w_i$  和  $v_i$  相同或者存在差异,本文设定协变量  $w_i$  和  $v_i$  相同,包括关注的 GDP、CPI 等核心宏观经济变量以及从高维指标中提取的动态因子信息。上述模型设定中,(2) 式与(4) 式中词语  $j$  的系数不受其他词语系数的影响,使得每个词语的系数求解可以单独进行,并且在具体的极大似然估计中对参数设置了惩罚项,以此解决了文本数据的高维问题;(1) 式中  $h_{ij}$  和  $(1 + c_{ij}^*)$  将词语  $j$  是否被选以及使用次数的建模分离开,允许词语  $j$  出现 0 的概率较高,解决了文本数据中大量单词词频为 0 的稀疏性问题。

样本外预测是文本分析或机器学习的重要应用,在 HDMR 模型中体现为如何利用文本数据构造预测因子。Kelly et al. (2019) 认为 HDMR 可以通过逆回归进行基于文本数据的有监督学习,对于关注的预测变量  $y$  而言,将其作为监督变量  $w_i$  和  $v_i$ ,利用极大似然估计得到参数估计值  $\hat{\delta}_y$  和  $\hat{\varphi}_y$ ,进而构造  $z_{iy}^0 = \hat{\delta}_y' h_i$  和  $z_{iy}^+ = \hat{\varphi}_y' (c_i - h_i)$  作为预测因子,包含了与预测变量  $y$  相关的所有文本信息。例如,当协变量  $v_i$  设定为利率时,则  $z_{iy}^0, z_{iy}^+$  表示了沟通文本中与利率紧密相关的信息,可以认为其反映了宽松还是紧缩的信号;当选取 GDP 增长率为协变量时,此时  $z_{iy}^0, z_{iy}^+$  代表了经济向好还是衰退的信息。

## (二) 央行沟通文本数据的收集和处理

### 1. 央行沟通文本的收集

在针对中国央行沟通的研究中,文本信息主要涉及书面沟通和口头沟通两种形式(卞志村和张义 2012; 林建浩等 2019)。其中,《货币政策执行报告》是最规范和系统的书面沟通文本,本文从中国人民银行官方网站下载了从 2003 年 1 月到 2019 年 6 月期间发布的所有季度《货币政策执行报告》,共计 66 篇。口头沟通是指央行的官员或发言人以讲话、采访、新闻发布会、时报文档等形式开展的沟通。参考卞志村和张义(2012)、林建浩等(2019),本文利用百度搜索引擎收集整理 2003 年 1 月到 2019 年 6 月所有的口头沟通文本数据,不仅涵盖最为权威的央行行长沟通,还涉及国务院分管经济工作的副总理、中国人民银行分管货币政策相关事务的副行长等其他官员

对中国货币政策走向和宏观经济形势的沟通。对于每个口头沟通事件, 本文重点记录了其对应的沟通时间、沟通渠道和沟通内容三个主要维度的信息。最终, 本文得到央行行长的有效沟通事件共计 345 次, 其他官员的有效沟通事件共计 72 次。图 1 展示了将合计的 417 次沟通事件按四类沟通渠道划分后的季度频次统计。根据宏观经济预测的月度频率, 本文将每月的全部沟通文本定义为一篇文档进行测度, 最终得到  $n = 198$  篇文档。

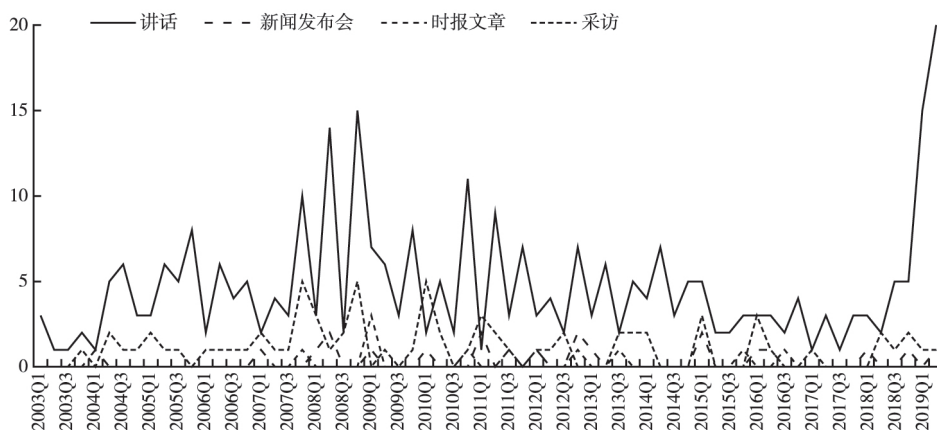


图1 央行口头沟通事件季度频次

## 2. 央行沟通文本数据的处理

对原始文本去除标点符号、停用词等一系列标准化预处理之后, 本文基于 Python 软件的“jieba”分词模块进行初步分词。考虑到央行沟通的专业性, 简单分词往往不能体现完整的信息, 需要采用  $n$  词短语( $n$ -gram)更为合适, 例如完整保留“货币供应量”而不是分为“货币”和“供应量”两个词。具体而言, 本文设定短语包含词数  $n$  的上限为 5, 由此构成一个自建的用户词典, 称之为“央行沟通词典”。将该词典输入 Python 的“jieba”模块进行二次分词。最终得到词语集合共有  $d = 63328$  个词语。

文本数据的高维性体现为词汇较多, 需要进一步筛选出更具信息含量的词语集合, 从而减少噪音和计算成本。本文采用逆文档频率方法(term frequency-inverse document frequency, TF-IDF)对词语筛选, 一方面, 词语出现的频率越高(TF 越大)则越重要。另一方面, 词语并非所有文档都出现, 而仅在一些特定文档中出现(IDF 越大), 则该词语具有良好的区分能力(Gentzkow et al., 2019; Hansen et al., 2018)。计算公式如下:

$$TF - IDF_j = TF_j \cdot IDF_j = \frac{f_j}{\sum_{k=1}^d f_k} \cdot \log \frac{n}{n_j} \quad (5)$$

其中,  $TF$  (term frequency) 即词频, 表示一个给定词语  $j$  在所有文档中出现的频率;  $IDF$  (inverse document frequency) 即逆文档词频, 用于衡量词语  $j$  的区分能力, 即其在所有文档中的常见程度;  $f_j$  是词语  $j$  在所有文档中的出现次数,  $\sum_{k=1}^d f_k$  是所有文档中所有词语的出现次数总和;  $n$  是语料库中所有文档总数, 分母  $n_j$  是包含词语  $j$  的文档个数。兼顾信息含量和计算成本, 本文仅保留逆文档频率最高的前  $d^*$  个词语。通过交叉验证, 本文选取了逆文档频率最高的  $d^* = 15000$  个词语所构成的词频矩阵。

由于央行沟通文本按照月度进行划分, 因此第  $i$  篇文档亦指第  $t$  月份文档, 为了方便用于时间序列预测, 后文将经过 HDMR 模型得到的央行沟通测度下标改为  $t$ , 即得到  $z_{ty}^0 = \hat{\delta}_y h_t$  (缩写为  $z_{-zero}$ ) 和  $z_{ty}^+ = \hat{\varphi}_y (c_t - h_t)$  (缩写为  $z_{-pos}$ ) 并基于这两个预测因子对中国宏观经济变量进行预测。

### 三、变量选取与模型构建

#### (一) 宏观经济变量数据的收集和选取

本文选取消费者物价指数(CPI)、生产价格指数(PPI)、GDP同比增长率(GDP)、工业增加值同比增长率(IVA)、固定资产投资同比增长率(FAI)、广义货币供应量同比增长率(M2)等6个宏观经济变量作为目标预测变量。GDP是反映经济运行的重要指标,为了匹配月度序列,本文参照 Higgins et al. (2016) 和 Chen et al. (2016) 选取的由宏观金融研究中心<sup>①</sup>所公布的月度实际GDP同比增长率。动态因子模型涉及众多宏观经济变量的选择和处理,<sup>②</sup>主要包括以下14个大类指标:实际产出类、物流类、实际投资类、房地产类、实际消费类、进口类、期货市场、股市类、外汇类、财政收支类、货币和信贷类、价格类、利率类、债市类共139个月度指标。此外,考虑到中国开放经济的事实,添加美国经济指标、国际大宗商品指标共13个月度指标体现中国与世界经济的联系。数据来源于国家统计局和中经网统计数据库,数据时间范围为2003年1月到2019年6月。

#### (二) 宏观经济预测基准模型设定和比较<sup>③</sup>

基于对常用预测模型的梳理,本文主要比较以下五种宏观经济预测模型的预测精度,并以此确定基准模型。

1. 小型VAR模型。借鉴央行公布的工作论文(马俊等 2014 2015),本文针对上述6个目标预测变量构建VAR模型,根据BIC准则确定最优滞后阶数为2阶。

2. 中型VAR模型。在上述小规模VAR模型基础上纳入更多的经济变量:社会商品零售总额同比增速、出口额同比增速、进口额同比增速、房地产开发投资同比增速、7天同业拆借利率、财政支出同比增速、财政收入同比增速、商品房销售面积同比增速、美国工业生产指数、美国联邦基金利率(马俊等 2014 2015),这种方法直接扩展了小型VAR模型的维度,根据BIC准则设定滞后阶数为1。

事实上,当变量个数较多时,VAR模型存在“维度诅咒”问题,即模型参数过多导致其难以估计。然而,小型数据集包含的信息有限,导致VAR模型所包含的信息少于央行与经济个体做决策时所利用的信息,难以全面捕捉经济变量的影响关系。

3. 动态因子模型(dynamic factor model, DFM)。动态因子模型可以解决传统VAR模型的部分缺陷。由于不同的经济变量之间普遍存在着共同波动现象,通过提取共同因子可以实现对高维数据的降维处理。动态因子模型基于这一基本事实,从高维经济变量中提取共同因子,用不可观测因子刻画多个变量间的相关关系,解决“维度诅咒”问题,是当前文献处理大数据的主要工具之一(Bok et al. 2018)。该方法在中国宏观经济预测、经济周期指标构建和核心通货膨胀率估计等方面也得到广泛应用(白强和白仲林 2017)。具体模型可表示为:

$$X_t = A f_t + e_t \quad (6)$$

$$f_t = A_1 f_{t-1} + A_2 f_{t-2} + \dots + A_p f_{t-p} + u_t \quad (7)$$

其中  $X_t$  是高维的经济变量,  $f_t$  是  $q \times 1$  维的动态因子,  $A$  是因子载荷矩阵,  $e_t$  是特定误差项,假定动态因子和特定误差项之间不相关。模型允许不同宏观变量的特定误差项  $e_t$  存在同期相关,因此也被称为近似因子模型。假定  $f_t$  服从自回归过程,  $A_1, A_2, \dots, A_p$  是因子自回归系数矩阵,  $u_t$  是  $q \times 1$

① 宏观金融研究中心由上海交通大学上海高级金融学院成立,由张春教授主持,并由诺贝尔经济学奖得主 Christopher Sims 教授和多位海外知名宏观经济学者任学术顾问。

② 指标选取具体参考 Bernanke et al. (2005)、Hansen & McMahon (2016) 以及黄卓等 (2018); 指标名称和处理细节备案。

③ 特别感谢审稿专家对于宏观经济预测基准模型比较的重要建议。

维的扰动向量。

Stock & Watson( 2016) 指出虽然主成分估计是常用方法 ,但是随着参数数量的增加 ,贝叶斯估计有助于提高估计效率和模型预测效果。Koopman & Mesters( 2017) 通过模拟证实了贝叶斯估计方法的优势在于可以通过收缩估计来减少预测函数参数的均方误差。因此 ,本文使用贝叶斯方法进行估计 ,因子个数设定为已有文献认为预测效果最佳的  $q = 5$ ( Stock & Watson 2012; Forni et al. , 2000; Kelly et al. 2019) ,BIC 信息准则确定滞后阶数  $p = 2$ 。参考 Stock & Watson( 2012) 和 Kelly et al. ( 2019) ,本文构建以下实时预报和短期预测的方程 ,后文简称为“DFM 模型”:

$$Y_{t+\tau} = \beta_0 + f_t'\beta + \varepsilon_{t+\tau} \quad (8)$$

其中  $Y_{t+\tau}$  表示需要预测的目标宏观变量在未来  $\tau$  期时点上的数值大小( 对于本文所讨论的月度序列而言  $\tau$  的取值一般为 1 ,3 ,12 等) 。

4. DFM + 历史信息模型。参考 Forni et al. ( 2000) 和 Stock & Watson( 2012) ,本文考虑宏观经济变量的动态相依特征 ,添加其历史信息来增强预测能力 ,简称为“DFM + 历史信息模型” ,即:

$$Y_{t+\tau} = \beta_0 + f_t'\beta + \beta_1 Y_t + \beta_2 Y_{t-1} + \dots + \beta_p Y_{t-p+1} + \varepsilon_{t+\tau} \quad (9)$$

其中  $p$  表示历史信息的滞后阶数。

5. FAVAR 模型( factor-augment vector autoregressive) 。Bernanke et al. ( 2005) 提出的 FAVAR 模型同时考虑了动态因子结构和少数重要变量的动态相依性。本文将 6 个目标预测变量作为可观测因子 ,使用因子信息准则得到不可观测的因子数为 5 ,根据 BIC 准则设定 1 阶滞后。

预测( forecasting) 与实时预报( nowcasting) 均有助于市场对未来的经济形势形成合理判断 ,但是两者又有所区别。从分析目标来看 ,预测侧重于对未来经济进行预测 ,即对  $Y_{t+\tau}$  进行建模; 实时预报强调对当前或最近经济情况的分析 ,是对存在发布时滞的  $Y_t$  进行建模; 从数据基础来看 ,预测依赖于当期完整的经济信息 ,而实时预报利用除目标变量以外的部分当期指标 ,更新对当期目标变量的实时预测( Giannone et al. 2008; Bok et al. 2018; Kapetanios & Papailias 2018) 。全样本范围为 2003 年 1 月至 2019 年 6 月 ,本文将 2009 年 1 月至 2019 年 6 月作为预测效果的测试集 ,计算滚动样本外的预测误差 RMSE ,结果如表 1 所示。整体而言 ,“DFM + 历史信息模型”在绝大多数情况下达到或者接近最优 ,其次是不含历史信息的“DFM 模型” ,其在长期预测上表现更好。因此 ,本文采用“DFM + 历史信息模型”作为基准模型 ,在此基础上考察央行沟通对预测效果的提升程度。

表 1 基于 RMSE 指标的宏观经济预测模型比较

目标变量	预测方法	预测阶数 $\tau$	小型 VAR	中型 VAR	DFM	DFM + 历史信息	FAVAR
GDP	实时预报		1. 1217	1. 0995	1. 5081	0. 7577	1. 1027
	预测	1	1. 1124	1. 045	1. 5081	0. 7577	1. 0909
		3	1. 4065	1. 4883	1. 6027	0. 8154	1. 7023
		12	1. 9619	2. 5162	1. 4580	1. 2466	2. 3452
	累计 RMSE		5. 6025	6. 1490	6. 0769	3. 5774	6. 2411
IVA	实时预报		2. 1113	2. 4685	1. 8492	1. 8711	2. 3070
	预测	1	2. 0636	2. 1020	1. 8492	1. 8711	1. 9927
		3	2. 7452	2. 4518	2. 1938	2. 2187	2. 6106
		12	4. 0019	4. 7406	3. 0280	3. 0669	4. 6259
	累计 RMSE		10. 9221	11. 7629	8. 9202	9. 0278	11. 5034

续表 1

目标变量	预测方法	预测阶数 $\tau$	小型 VAR	中型 VAR	DFM	DFM + 历史信息	FAVAR
FAI	实时预报		8.4213	6.1481	6.9511	5.3979	6.0546
	预测	1	8.2426	6.4319	6.9511	5.3979	6.1007
		3	10.9991	11.2127	8.7342	8.7030	10.4752
		12	15.2368	17.2176	8.1026	8.3966	16.4388
累计 RMSE			42.8997	41.0103	30.7390	27.8954	39.0693
M2	实时预报		1.4458	0.9707	1.0871	0.8973	1.0360
	预测	1	1.4461	0.9639	1.0871	0.8973	1.0770
		3	2.5883	1.8314	1.5298	1.4660	1.9145
		12	6.4959	6.5168	2.9213	3.0768	7.3847
累计 RMSE			11.9761	10.2828	6.6253	6.3374	11.4122
CPI	实时预报		0.7417	0.5954	0.5374	0.4862	0.6116
	预测	1	0.7410	0.6140	0.5374	0.4862	0.6056
		3	1.2370	1.1280	0.6654	0.6761	1.0115
		12	2.3190	3.8007	1.7314	1.4887	2.7826
累计 RMSE			5.0387	6.1381	3.4716	3.1372	5.0113
PPI	实时预报		1.7585	0.8343	0.8471	0.5460	0.7734
	预测	1	1.7687	0.8707	0.8471	0.5460	0.7210
		3	3.3655	2.5829	1.9550	1.8268	2.4030
		12	6.3559	8.2507	3.8684	3.7639	8.4589
累计 RMSE			13.2487	12.5385	7.5176	6.6828	12.3564

注: 本表是基于 2009 年 1 月至 2019 年 6 月测试集, 计算不同宏观预测模型滚动样本外的预测误差 RMSE; 累计 RMSE 为实时预报加上预测的 RMSE 之和。

#### 四、央行沟通文本对宏观经济变量的预测效果

##### (一) 预测和实时预报模型

在预测中, 本文主要考察央行沟通测度能否在动态因子模型的基础上, 提高对  $\tau$  期后目标预测变量的预测效果, 即在模型(9)的基础上, 加入了本文所构造的央行沟通测度  $z^+$  和  $z^0$  作为额外的预测因子:

$$Y_{t+\tau} = \beta_0 + [df_t^1, \dots, df_t^5, z_{tY}^+, z_{tY}^0, m_t, d_t] \beta + \beta_1 Y_t + \beta_2 Y_{t-1} + \dots + \beta_p Y_{t-p+1} + \varepsilon_{t+\tau} \quad (10)$$

其中  $t$  表示当前时点,  $Y_{t+\tau}$  表示未来  $\tau$  期的目标经济变量,  $df_t^1, \dots, df_t^5$  为五个动态因子,  $p$  是历史信息的滞后阶数,  $\varepsilon_t$  为扰动项。此外, 为了尽可能消除不同文档长短差异的影响, 本文将每期文档的总词数  $m_t$  和文档中出现的词语种类数  $d_t$  也引入预测模型中。

实时预报关注的是当期预测, 即考察滞后一期 ( $t-1$  时期) 的宏观经济动态因子和当期 ( $t$  时期) 的央行沟通信息, 能否对当期 ( $t$  时期) 的目标变量有较好的预测效果。现实中, 宏观经济变量的统计和公布往往是滞后的, 而央行沟通实效性较强、传播速度较快, 利用实时更新的央行沟通文本, 可以在当期宏观经济变量公布之前对其进行实时预报。模型设定如下:

$$Y_t = \beta_0 + [df_{t-1}^1, \dots, df_{t-1}^5, z_{tY}^+, z_{tY}^0, m_t, d_t] \beta + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \varepsilon_t \quad (11)$$



其中  $Y_t$  表示当期（即  $t$  时期）的目标预测变量，央行沟通测度  $z_{tY}^+$ 、 $z_{tY}^0$ 、 $m_t$ 、 $d_t$  均为当期指标，其他指标诸如动态因子  $df_{t-1}^1, \dots, df_{t-1}^5$  和  $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$  都是取滞后期数据。

（二）样本内模型拟合度

表 2 展示了所有目标变量基于动态因子模型的回归结果。当期的自回归项  $Y_t$  和五个动态因子大多在 1% 的水平上显著影响了未来一期的目标宏观经济变量，说明本文利用中国 152 个宏观变量序列所提取出的动态因子能较好地解释核心宏观经济变量未来一段时期的波动。从调整  $R^2$  看，六个模型均在 0.825 以上，平均为 0.904，表明基准模型对下一期的目标宏观经济变量具有较强的样本内拟合效果。

表 2 基于 DFM + 历史信息的回归结果

	(1)	(2) ①	(3)	(4)	(5)	(6)
	GDP <sub>t+1</sub>	IVA <sub>t+1</sub>	FAI <sub>t+1</sub>	M2 <sub>t+1</sub>	CPI <sub>t+1</sub>	PPI <sub>t+1</sub>
$Y_t$	0.133* (0.0726)	0.117 (0.0813)	0.708*** (0.0740)	0.582*** (0.0791)	0.472*** (0.0784)	1.177*** (0.160)
DF1	0.458*** (0.131)	0.363 (0.284)	-3.513*** (0.882)	-1.078*** (0.145)	0.162* (0.0858)	-0.412* (0.211)
DF2	-0.338*** (0.0872)	0.338 (0.225)	1.648*** (0.523)	-0.129 (0.0856)	0.269*** (0.0896)	0.101 (0.0921)
DF3	0.188** (0.0743)	-0.130 (0.161)	0.167 (0.437)	-0.0135 (0.0693)	-0.688*** (0.0944)	0.886*** (0.248)
DF4	-0.268*** (0.0784)	-0.00705 (0.168)	-1.706*** (0.564)	-0.466*** (0.0804)	0.895*** (0.122)	0.403*** (0.0710)
DF5	0.358*** (0.0748)	0.301* (0.158)	2.860*** (0.682)	1.516*** (0.201)	-0.882*** (0.133)	-1.093*** (0.248)
滞后项	4 阶	2 阶	2 阶	3 阶	0 阶	2 阶
常数项	控制	控制	控制	控制	控制	控制
观测值	193	195	195	194	197	195
R <sup>2</sup>	0.883	0.837	0.832	0.973	0.937	0.987
adj. R <sup>2</sup>	0.877	0.830	0.825	0.972	0.935	0.987

注：估计模型为  $Y_{t+1} = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + \dots + \beta_p Y_{t-p+1} + \varepsilon_{t+1}$ ；每个模型的滞后阶数由贝叶斯信息准则确定，最优滞后阶数不同也导致参与估计的观测值数量相应变化；括号内为标准误，\*、\*\*、\*\*\* 分别表示在 10%、5% 和 1% 水平上显著，以下各表同。

表 3 验证央行沟通能否提供短期预测（ $\tau = 1$ ）的额外信息。首先，从调整  $R^2$  来看，加入沟通测度之后，六个目标变量回归的调整  $R^2$  均在 0.843 以上，平均为 0.913，相比表 2 有一定提升，表明沟通测度对未来一期的目标经济变量具有额外解释能力。其次，本文所构造的两个核心沟通测度  $z_{tY}^+$ （ $z_{tY}^{pos}$ ）和  $z_{tY}^0$ （ $z_{tY}^{zero}$ ）对所有六个目标经济变量的回归系数大多为正且显著，符合理论预期。以预测 M2 增速为例，“流动性过多”这一短语对应的逆回归系数  $\varphi_y$  为正，即货币供应量增速越高时央

① 对于 IVA 的预测拟合度较高，而各变量系数的显著性较低的原因可能是共线性造成的。单独引入滞后项或动态因子时，回归系数大部分在 1% 水平上显著。

行沟通采用越多的“流动性过多”表述,由此当文本数据中“流动性过多”的频率较高时可以推测货币供应量增速过快,即  $z_{ty}^+ = \varphi_y (c_t - h_t)$  与 M2 增速呈正相关关系;相反地,当词语  $j$  的词频  $c_{ij}$  与 M2 增速负相关时,例如“流动性下降”对应的系数  $\varphi_y$  为负,当“流动性下降”出现频率越低时  $z_{ty}^+ = \varphi_y (c_t - h_t)$  越大, M2 增速越高,因此表现为 M2 增速与  $z\_pos$  仍呈正相关。

表 3 基于 DFM + 历史信息 + 央行沟通文本的回归结果

	(1)	(2)	(3)	(4)	(5)	(6)
	GDP <sub>t+1</sub>	IVA <sub>t+1</sub>	FAI <sub>t+1</sub>	M2 <sub>t+1</sub>	CPI <sub>t+1</sub>	PPI <sub>t+1</sub>
$Y_t$	0.0380 (0.0793)	-0.105 (0.0963)	0.449 *** (0.0763)	0.453 *** (0.0848)	0.341 *** (0.0938)	0.823 *** (0.197)
$z\_pos$	7.964 * (4.068)	24.110 * (12.700)	201.000 ** (78.260)	7.457 *** (2.813)	-0.450 (0.822)	1.553 (0.956)
$z\_zero$	4.228 (2.679)	47.530 * (25.480)	865.700 *** (202.500)	3.096 ** (1.487)	2.378 *** (0.870)	2.552 (1.975)
$m$	0.000593 (0.000635)	0.00264 ** (0.00121)	-0.000711 (0.00298)	0.00138 ** (0.000602)	0.0000354 (0.000298)	0.000122 (0.000316)
$d$	-0.000860 (0.000580)	-0.00220 ** (0.00103)	0.00290 (0.00271)	-0.000583 (0.000573)	0.000243 (0.000254)	0.0000844 (0.000311)
滞后项	4 阶	2 阶	2 阶	3 阶	0 阶	2 阶
DF1 - 5	控制	控制	控制	控制	控制	控制
常数项	控制	控制	控制	控制	控制	控制
观测值	193	195	195	194	197	195
R <sup>2</sup>	0.889	0.853	0.868	0.975	0.941	0.988
adj. R <sup>2</sup>	0.880	0.843	0.859	0.974	0.937	0.987

注:估计模型为  $Y_{t+1} = \beta_0 + [df_t^1 \dots df_t^5 \tilde{z}_{tY}^+ \tilde{z}_{tY}^0 m_t d_t] \beta + \beta_1 Y_t + \beta_2 Y_{t-1} + \dots + \beta_p Y_{t-p+1} + \varepsilon_{t+1}$ 。

为了更直观地体现出央行沟通测度在宏观经济预测上的效果,本文以 GDP 增长率和 CPI 为例,将基于 DFM + 历史信息(表 2)和基于 DFM + 历史信息 + 央行沟通文本(表 3)的拟合序列与目标变量的真实序列在图 2 和图 3 中直观展示。三条曲线较为接近,表明两个模型对 GDP 增长率和 CPI 两个指标都具有很好的拟合效果,其中尤以 CPI 为甚。

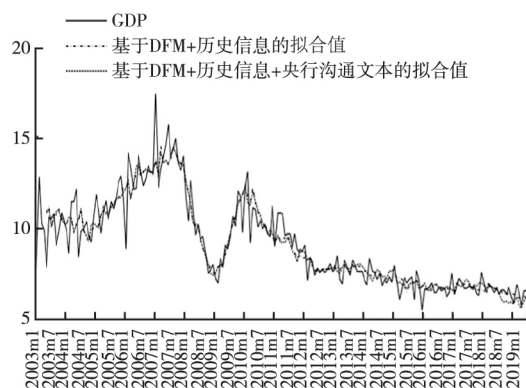


图 2 GDP 增长率预测拟合对比

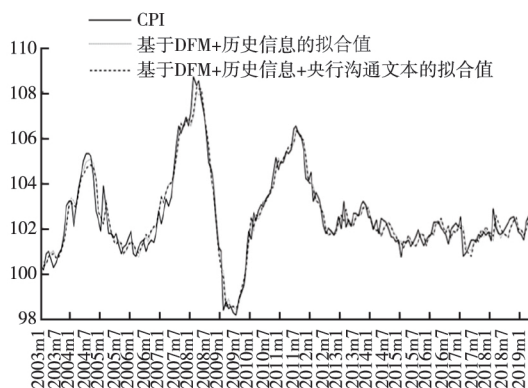


图 3 CPI 预测拟合对比

(三) 样本外模型预测精度

为了考察央行沟通对宏观经济变量的预测能力,本文采用现有文献(Kelly et al. 2019; Stock & Watson 2012)常用的样本外均方根误差(RMSE)对预测能力进行衡量。根据 Kelly et al. (2019)的伪样本外滚动预测(pseudo out-of-sample rolling forward predictions)方法,首先以一定比例的样本作为训练集进行回归并预测下一期目标宏观变量,再将下一期也纳入训练集进行回归并预测再下一期,如此循环迭代直至预测完所有可得样本。数据时间范围是2003年1月—2019年6月,为了保证有充足的样本进行回归拟合(样本数>30),本文选择从2009年1月开始进行伪样本外滚动预测。基准预测模型包括:DFM模型、DFM+1阶历史信息、DFM+多阶历史信息。

基于基准模型统计其样本外预测RMSE,同时在基准模型基础上加入央行沟通测度后统计其样本外预测RMSE,最后分别计算不同预测模型中央行沟通测度带来的预测精度提升比例,<sup>①</sup>结果如表4所示。第(1)列表明,在不包含预测变量历史信息的情况下,央行沟通带来的预测精度提升明显,所有指标预测的平均精度提升比例达到6.80%—16.65%。从实时预报和不同期限预测比较来看,央行沟通对实时预报的精度提升最为明显,表明央行沟通传导时效性较强。综合上述基于DFM模型的讨论,央行沟通能够显著提高对中国核心宏观变量的预测精度,相较于滞后公布宏观变量集的低时效性,央行沟通文本的高时效性能够对中国核心宏观变量的预测带来额外的优势。

第(2)列表明,在“DFM+1阶历史信息”模型中,央行沟通仍可以提升模型预测精度,但是提升幅度相比第(1)列有所下降,这种现象在第(3)列“DFM+多阶历史信息”模型中有更明显的体现。第(3)列在控制预测变量的历史信息之后,央行沟通对于预测精度的提升效果出现分化,对于大多数宏观经济变量的实时预报和短期预测( $\tau=1,3$ )的精度提升效果一般,甚至可能造成预测精度下降,只有在预测期限较长时( $\tau=12$ )才呈现出明显的精度提升。以GDP和IVA为例,第(3)列对应预测阶数 $\tau=12$ 的精度提升分别为0.53%和6.16%,明显高于 $\tau=1,3$ 的短期预测精度提升。这一发现可以从以下两个角度进行理解:一方面,中国的央行沟通更多是传达历史和现状信息,经济形势展望和货币政策的前瞻性指引信息相对较少;另一方面,当央行沟通信息与宏观经济变量的历史信息重叠较多时,这种高度共线性将导致模型样本外预测能力下降。对于长期预测而言,预测变量久远的历史信息提供的预测能力有限,此时央行沟通信息反而能够发挥其特殊的作用。

表4 基于DFM模型和不同信息集下央行沟通测度预测能力分析

目标变量	预测方法	预测阶数 $\tau$	DFM 模型	DFM + 1 阶历史信息	DFM + 多阶历史信息
			(1)	(2)	(3)
GDP	实时预报		40.41%	18.47%	-12.99%
	预测	1	30.13%	7.08%	-18.18%
		3	32.37%	6.83%	-3.40%
		12	15.38%	3.74%	0.53%
IVA	实时预报		10.91%	7.82%	6.25%
	预测	1	7.40%	0.63%	-7.13%
		3	1.72%	1.57%	1.17%
		12	2.32%	1.82%	6.16%

<sup>①</sup> 此处和下文提到的预测精度提升比例,计算公式均为:(基准模型的误差-本文模型的误差)/基准模型的误差。由于篇幅限制,仅展示精度提升比例,具体的RMSE结果备索。

续表 4

目标变量	预测方法	预测阶数 $\tau$	DFM 模型	DFM + 1 阶历史信息	DFM + 多阶历史信息
			(1)	(2)	(3)
FAI	实时预报		15.08%	4.28%	3.03%
	预测	1	10.37%	-8.81%	-8.10%
		3	6.62%	5.45%	4.70%
		12	5.21%	4.74%	5.69%
M2	实时预报		9.25%	-0.72%	-1.93%
	预测	1	10.15%	0.95%	-1.29%
		3	0.61%	0.81%	-0.59%
		12	3.20%	3.62%	4.09%
CPI	实时预报		11.22%	4.67%	4.67%
	预测	1	8.76%	1.49%	1.49%
		3	1.69%	2.85%	2.85%
		12	4.10%	-3.28%	-3.28%
PPI	实时预报		13.01%	7.00%	4.75%
	预测	1	4.74%	1.56%	0.34%
		3	0.79%	0.49%	-3.22%
		12	10.56%	10.01%	7.54%
平均	实时预报		16.65%	6.92%	0.63%
	预测	1	11.93%	0.49%	-5.48%
		3	7.30%	3.00%	0.25%
		12	6.80%	3.44%	3.46%

注:本表是基于 2009 年 1 月—2019 年 6 月测试集,利用滚动样本外的预测误差 RMSE 计算的提升比例;DFM 模型为:  $Y_{i+\tau} = \beta_0 + f_i^* \beta + \varepsilon_{i+\tau}$ ,DFM + 1 阶历史信息为:  $Y_{i+\tau} = \beta_0 + \beta_1 Y_i + f_i^* \beta + \varepsilon_{i+\tau}$ ,DFM + 多阶历史信息为:  $Y_{i+\tau} = \beta_0 + \beta_1 Y_i + \dots + \beta_p Y_{i-p+1} + f_i^* \beta + \varepsilon_{i+\tau}$ ,滞后阶数由 BIC 准则确定;对于 CPI 而言,BIC 信息准则确定的 P 为 1 阶,因此其 DFM + 1 阶历史信息和 DFM + 多阶历史信息的实证结果相同;平均数据中的预测精度提升比例为六个目标变量对应处理下提升比例的算术平均值。

#### (四) 稳健性分析

##### 1. 不同央行沟通文本集合的信息含量比较

口头沟通频率更高、时效性更强,季度书面报告篇幅更长,对经济形势的刻画更为细致,信息含量也更高,尤其是每篇季度报告均有“中国宏观经济展望”章节,涉及经济增速、政策目标、市场调查结果等信息,可能对中长期经济判断更具价值。针对上述不同文本的差异,本文分别考虑全部文本、口头沟通文本、口头沟通 + 书面报告政策展望部分等三种文本内容的预测效果。以 GDP 为例,图 4 给出了模型估计结果的 RMSE 比较。<sup>①</sup>

首先,口头沟通在实时预报中具有一定优势,仅用口头沟通对 GDP 进行实时预测的 RMSE 为 0.823,而书面和口头沟通所有文本时为 0.856;但在短期预测中,口头沟通的优势不再,此时仅用口头沟通预测下一期 GDP 时 RMSE 为 0.945,精度低于所有沟通文本的 0.895。这种规律在其他经济变量中也存在。

其次,在包含历史信息的 DFM 中,基于所有沟通文本测度的预测精度在绝大多数情况下都具

<sup>①</sup> 由于篇幅限制,仅展示 GDP 的计算结果,其余宏观变量具体的 RMSE 结果备案。

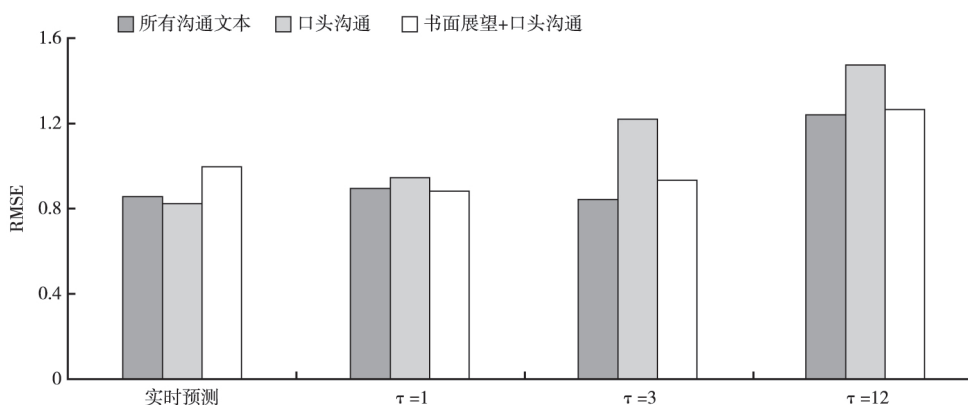


图4 不同沟通文本的预测表现比较(以GDP为例)

注: 本图是基于2009年1月—2019年6月测试集结果; 预测变量为GDP; 预测模型为“DFM+历史信息模型”下引入不同央行沟通文本; 呈现各种情形下的滚动样本外预测误差RMSE。

有最佳表现。这表明, 文本数据同样是未来经济预测的重要信息。经济判读具有延续性和广泛性, 央行书面报告中的历史经济信息或国外经济形势同样是预判未来宏观经济的重要信息。例如, 2018年第一季度的《中国货币政策执行报告》在国际经济展望中明确提出“保护主义、单边主义和反全球化思潮引发全球贸易摩擦升级, 成为影响全球经济增长的重要风险”, 强调了贸易争端可能波及全球较大范围的商品和服务。央行沟通明确指出这些风险因素可能抑制需求、推高通胀、拖累经济增长, 甚至引发更大的全球宏观经济风险, 凸显了这些风险因素对预判未来经济形势的重要性。

## 2. 不同文本测度方法的预测能力比较

在已有文献中, 林建浩等(2019)针对央行行长的口头沟通进行有监督词典法的测度, 本文测度考虑了文本的稀疏性和高维性, 在预测宏观经济变量上是否具有优势? 为此, 本文同样基于有监督词典法测度了货币政策沟通指数  $I_{MP}$  和经济形势沟通指数  $I_{EC}$ , 将其可作为新的预测因子。<sup>①</sup> 与前文一致, 本文同时考虑了每期文档的总词数  $m_t$  和文档中出现的词语种类数  $d_t$ 。具体预测模型表示为:

$$Y_{t+\tau} = \beta_0 + [df_t^1, \dots, df_t^5, I_{MP}, I_{EC}, m_t, d_t] \beta + \beta_1 Y_t + \beta_2 Y_{t-1} + \dots + \beta_p Y_{t-p+1} + \varepsilon_{t+\tau} \quad (12)$$

针对口头沟通文本, 本文比较了用HDMR提取所有文本的信息、HDMR仅提取口头沟通文本的信息、有监督词典法提取口头沟通信息共三种模型的宏观经济变量实时预报和预测能力。<sup>②</sup> 结果表明, 基于HDMR提取所有文本信息能够实现最高的预测精度, 而仅利用口头沟通文本时HDMR和有监督词典法提取的预测精度较为接近。这可能是由于口头沟通文本较短, 监督词典法受到高维性的影响较小。但是, 在研究央行沟通时忽略季度报告将产生严重的信息遗漏, 而书面报告的文本数据存在高维性和稀疏性, 有监督词典法在书面沟通文本中难以推广。因此, 本文仍主要考虑所有文本的HDMR模型。

<sup>①</sup> 由于有监督词典法需要采用人工标注句子倾向, 对于字斟句酌的季度执行报告而言, 这种方法显然不适用(林建浩等, 2019)。一方面, 季度报告篇幅较长, 成千上万的句子进行标注将大大增加误判风险; 另一方面, 季度执行报告的语言风格更加严谨和准确, 涵盖信息量大, 对专业能力有限的公众形成了认知限制。因此, 本文仅利用有监督词典法测度了口头沟通文本。

<sup>②</sup> 由于篇幅限制, 具体的RMSE计算结果备索。

## 五、结论与启示

央行沟通是宏观调控预期管理的核心机制,已发展成为重要的货币政策工具。这些即时、公开的叙事文本影响着市场主体关于宏观经济和货币政策的判断和预期。然而,这些颇有价值的文本信息很少被系统地整理和提取出来,并应用于中国宏观经济的实时预报和预测。本文开发出一整套适应中国央行沟通特有表达方式的文本分析和处理过程,构造了低维的央行沟通测度。与其他央行沟通测度相比,本文具有以下优势:第一,本文收集汇总所有的口头和书面沟通文本,进行n-gram分词、组合进而构成符合央行沟通用语习惯的专用词典,由此构建的语料库更具专业性。第二,本文利用最新的HDMR模型,能够解决央行沟通文本的高维性和稀疏性问题,进而实现对文本信息的全面提取和有效降维。第三,HDMR通过逆回归进行基于文本数据的有监督学习,将预测变量视作监督变量,完整而简洁地概括了对特定宏观经济变量有解释作用的沟通信息,这种机器学习方法能够针对不同预测变量构造不同的央行沟通测度。

与其他常用宏观预测模型相比,基于152个指标构建的动态因子模型对中国主要宏观经济变量具有更优的预测能力。在此基础上,进一步引入央行沟通测度作为新的预测因子,结果显示央行沟通测度有助于提升模型样本内拟合效果。考察样本外预测效果,在不包括预测变量历史信息时,央行沟通测度能够使得不同期限的预测精度提升幅度达6.80%—16.65%;包括预测变量历史信息时则出现分化,在期限较短时,央行沟通未能提升预测精度,这是因为主要沟通信息与预测变量历史信息重叠;当期限较长时,预测精度有所提升,表明沟通中少量的前瞻性指引具有持续的预测能力。

本文聚焦央行沟通文本数据的分析建模,并从宏观经济预测角度验证了中国央行沟通在预期管理中的作用。本文对于理解央行沟通与预期管理,以及推进基于大数据的宏观经济预测均具有重要的参考意义。据此,可以得出以下政策启示:

第一,建立制度化的货币政策沟通机制、有效管理和引导预期,是建设现代中央银行制度的重要举措。首先,央行应该适当增加更明确的未来经济形势研判和前瞻性指引信息的披露。本文研究表明,从宏观经济预测的角度来看,目前中国的央行沟通信息与经济形势和货币政策的历史和现状存在重叠。由于公共信息对私人信息具有挤出效应,如果央行披露信息过多或者有所偏误时,反而可能加剧经济系统的不稳定性。因此,要提供更精确的未来经济形势研判,必须进一步加强央行的研究能力以及数据资源搜集和开发能力,才能在沟通中释放独特信息引导公众预期。其次,2020年7月的中央政治局会议首次提出完善宏观调控跨周期设计与调节,《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议》关于宏观经济治理也明确提出搞好跨周期政策设计,易纲(2020)指出做好跨周期政策设计,以现代化的货币管理促进经济高质量发展,因此本文从央行沟通信息预测效果的期限结构这一角度给出预期管理跨周期政策设计的建议。考虑到央行沟通信息效应的期限结构,央行可以尝试释放期限不同的货币政策信息,既立足于短期经济波动,维持金融市场稳定;同时注重引导长期预期,保证中长期的政策传导效应,实现跨周期调控。

第二,以文本大数据平台建设为推手,加快政府的数据开放与数字公共品提供,促进大数据在宏观经济预测中的应用,服务宏观调控和风险防范的实时需求。习近平总书记2017年12月提出“要充分利用大数据平台,综合分析风险因素,提高对风险因素的感知、预测、防范能力”。<sup>①</sup>无论是2008年的金融危机,近几年的中美贸易摩擦,还是2020年的新冠肺炎疫情,这些冲击变化迅速、影响巨大、传播广泛,使得政府和企业了解宏观经济实时变化的需求变得更迫切。常规的月度和季度统计

<sup>①</sup> [http://www.xinhuanet.com/2017-2/09/c\\_1122084706.htm](http://www.xinhuanet.com/2017-2/09/c_1122084706.htm)。

指标发布难以满足新的数据需求,以周和日度为单位的宏观经济实时预测建模是重要研究方向。本文对此提供了新思路,海量的新闻文本以及社交媒体文本记录了政府、企业和消费者的经济活动和市场情绪,可以提炼出以日甚至以小时为单位的感知大数据,实现宏观经济实时监测。为此,在数字政府建设方面,可以文本大数据平台为推手,合法、有序地推进大数据公共品的有效供给。

## 参考文献

- 白强、白仲林,2017《动态因子模型的广义矩估计(GMM)及其统计性质研究》,《统计研究》第10期。
- 卞志村、张义,2012《央行信息披露、实际干预与通胀预期管理》,《经济研究》第12期。
- 陈良源、林建浩、王少林、詹焯扬,2021《央行沟通对于货币政策实际干预的预测能力研究》,《统计研究》第1期。
- 陈雨露,2019《四十年来中央银行的研究进展及中国的实践》,《金融研究》第2期。
- 郭豫媚、周璇,2018《央行沟通、适应性学习和货币政策有效性》,《经济研究》第4期。
- 黄卓、邱晗、沈艳、童晨,2018《测量中国的金融不确定性——基于大数据的方法》,《金融研究》第11期。
- 林建浩、陈良源、宋登辉,2019《如何测度央行行长的口头沟通信息——一种基于监督学习的文本分析方法》,《统计研究》第8期。
- 刘汉、刘金全,2011《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》,《经济研究》第3期。
- 刘涛雄、徐晓飞,2015《互联网搜索行为能帮助我们预测宏观经济吗》,《经济研究》第12期。
- 马俊、刘斌、贾彦东等,2014《2015年中国宏观经济预测》,中国人民银行工作论文,NO.2014/10。
- 马俊、刘斌、贾彦东等,2015《2016年中国宏观经济预测》,中国人民银行工作论文,NO.2015/15。
- 沈艳、陈赞、黄卓,2019《文本大数据分析在经济学和金融学中的应用:一个文献综述》,《经济学(季刊)》第4期。
- 汪莉、王先爽,2015《央行预期管理、通胀波动与银行风险承担》,《经济研究》第10期。
- 吴国培、潘再见,2014《中央银行沟通对金融资产价格的影响——基于中国的实证研究》,《金融研究》第5期。
- 易纲,2020《建设现代中央银行制度》,载于《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议辅导读本》,人民出版社,280—286页。
- 郑挺国、王霞,2013《中国经济周期的混频数据测度及实时分析》,《经济研究》第6期。
- Beaudry, P., and F. Portier, 2014, "News Driven Business Cycles: Insights and Challenges", *Journal of Economic Literature*, 52(4), 993—1074.
- Bernanke, B. S., 2020, "The New Tools of Monetary Policy", *American Economic Review*, 110(4), 943—83.
- Bernanke, B. S., J. Boivin, and P. Elias, 2005, "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (Favar) Approach", *Quarterly Journal of Economics*, 120(1), 387—422.
- Blinker, A. S., M. Ehrmann, M. Fratzscher, J. De Haan, and D. Jansen, 2008, "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence", *Journal of Economic Literature*, 46(4), 910—45.
- Bok, B., D. Caratelli, D. Giannone, A. M. Sbordone, and A. Tambalotti, 2018, "Macroeconomic Nowcasting and Forecasting with Big Data", *Annual Review of Economics*, 10, 615—43.
- Chen, K., P. Higgins, D. F. Waggoner, and T. Zha, 2016, "Impacts of Monetary Stimulus on Credit Allocation and the Macroeconomy: Evidence from China", NBER Working Paper, No. 22650.
- Eusepi, S., and B. J. Preston, 2010, "Central Bank Communication and Expectations Stabilization", *American Economic Journal: Macroeconomics*, 2(3), 235—271.
- Eusepi, S., and B. J. Preston, 2011, "Expectations, Learning and Business Cycle Fluctuations", *American Economic Review*, 101(6), 2844—2872.
- Forni, M., M. Hallin, M. Lippi, and L. Reichlin, 2000, "The Generalized Dynamic-Factor Model: Identification and Estimation", *Review of Economics and Statistics*, 82(4), 540—54.
- Gentzkow, M., B. Kelly, and M. Taddy, 2019, "Text as Data", *Journal of Economic Literature*, 57(3), 535—74.
- Giannone, D., L. Reichlin, and D. Small, 2008, "Nowcasting: The Real-Time Informational Content of Macroeconomic Data", *Journal of Monetary Economics*, 55(4), 665—76.
- Hansen, S., and M. McMahon, 2016, "Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication", *Journal of International Economics*, 99, S114—S33.
- Hansen, S., M. McMahon, and A. Prat, 2018, "Transparency and Deliberation within the FOMC: A Computational Linguistics

Approach", *Quarterly Journal of Economics*, 133(2), 801—70.

Hansen, S., M. McMahon, and M. Tong, 2019, "The Long-Run Information Effect of Central Bank Communication", *Journal of Monetary Economics*, 108, 185—202.

Hanson, S. G., and J. C. Stein, 2015, "Monetary Policy and Long-Term Real Rates", *Journal of Financial Economics*, 115(3), 429—48.

Higgins, P., T. Zha, and W. Zhong, 2016, "Forecasting China's Economic Growth and Inflation", *China Economic Review*, 41, 46—61.

Jansen, D., and J. De Haan, 2007, "The Importance of Being Vigilant: Has Ecb Communication Influenced Euro Area Inflation Expectations?", CESifo Working Paper, No. 2134, Center for Economic Studies and Ifo Institute (CESifo), Munich.

Kapetanios, G., and F. Papailias, 2018, "Big Data & Macroeconomic Nowcasting: Methodological Review", Economic Statistics Centre of Excellence, Discussion Paper.

Kelly, B. T., A. Manela, and A. Moreira, 2019, "Text Selection", NBER Working Paper, No. 26517.

Koopman, S. J., and G. Mesters, 2017, "Empirical Bayes Methods for Dynamic Factor Models", *Review of Economics and Statistics*, 99(3), 486—98.

Lamla, M. J., and D. V. Vinogradov, 2019, "Central Bank Announcements: Big News for Little People?", *Journal of Monetary Economics*, 108, 21—38.

Larsen, V. H., and L. A. Thorsrud, 2019, "The Value of News for Economic Developments", *Journal of Econometrics*, 210(1), 203—18.

Loughran, T., and B. McDonald, 2011, "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10 - Ks", *Journal of Finance*, 66(1), 35—65.

McMahon, M., M. A. Schipke, and X. Li, 2018, "China's Monetary Policy Communication: Frameworks, Impact, and Recommendations", International Monetary Fund.

Morris, S., and H. S. Shin, 2008, "Coordinating Expectations in Monetary Policy", Central Banks as Economic Institutions, 88—104.

Peek, Joe, E. S. Rosengren, and G. M. B. Tootell, 2003, "Does the Federal Reserve Possess an Exploitable Informational Advantage?", *Journal of Monetary Economics*, 50(4), 817—39.

Rashid, J., S. M. A. Shah, and A. Irtaza, 2019, "Fuzzy Topic Modeling Approach for Text Mining over Short Text", *Information Processing & Management*, 56(6).

Romer, C. D., and D. H. Romer, 1989, "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz", *NBER Macroeconomics Annual*, 4, 121—70.

Shiller, R. J., 2017, "Narrative Economics", *American Economic Review*, 107(4), 967—1004.

Shiller, R. J., 2019, *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*, Princeton University Press.

Stock, J. H., and M. W. Watson, 2012, "Generalized Shrinkage Methods for Forecasting Using Many Predictors", *Journal of Business & Economic Statistics*, 30(4), 481—93.

Stock, J. H., and M. W. Watson, 2016, "Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics", *Handbook of Macroeconomics*, 415—525.

Su, S., A. H. Ahmad, and J. Wood, 2019, "How Effective Is Central Bank Communication in Emerging Economies? An Empirical Analysis of the Chinese Money Markets Responses to the People's Bank of China's Policy Communications", *Review of Quantitative Finance and Accounting*, 54, 1195—1219.

Sun, R., 2020, "Monetary Policy Announcements and Market Interest Rates' Response: Evidence from China", *Journal of Banking & Finance*, 113, 105766.

Taddy, M., 2013, "Multinomial Inverse Regression for Text Analysis", *Journal of the American Statistical Association*, 108(503), 755—70.

Taddy, M., 2015, "Distributed Multinomial Regression.", *Annals of Applied Statistics*, 9(3), 1394—414.

Thorsrud, L. A., 2020, "Words Are the New Numbers: A Newsy Coincident Index of the Business Cycle", *Journal of Business & Economic Statistics*, 38(2), 393—409.

Trichet, J., 2007, "The Role of Research in Central Banks and at the ECB", Award of the Germán Bernácer Prize, Madrid, 21.



## Does Central Bank Communication Improve Macroeconomic Forecasting? — High-dimensional Sparse Modeling Based on Text Data

LIN Jianhao , CHEN Liangyuan , LUO Zihao and ZHANG Yifan

(Lingnan College , Sun Yat-sen University)

**Summary:** Macroeconomic forecasting is not only an important issue of economic methodology , but also a main challenge for policy practice. Big data has brought a new revolution to economic research. Due to its high availability , textual data has become the fastest growing and most widely used unstructured data. According to the theory of news-driven business cycle , public expectation is the key factor of economic fluctuation , while news is an independent driving force of expectation. In the international practice , central bank communication usually includes monetary policy objectives and strategies , macroeconomic outlook and monetary policy stance. These communications can reduce the information asymmetry between the central bank and the public , help the public understand the changing economic environment , and guide the public to rational expectations of future economic status and monetary policy. Specifically , central bank communication can influence the expectation and behavior of economic agents through three information channels , namely , creating information , reducing noise and the term structure of information effect. Therefore , among all textual data , central bank communication has a clear and unique value. Based on the importance of macroeconomic forecasting and the uniqueness of central bank communication , this paper focused on whether central bank communication , an important narrative information , can improve the forecasting performance of key macroeconomic variables such as GDP , CPI and so on.

In the era of big data , central bank communication is a typical kind of important unstructured data. The main challenge of textual data analysis is extracting information accurately and effectively from these high-dimensional and sparse texts. Considering there is no literature applying textual data to the forecast of China's macroeconomy yet , this paper attempted to make a scientific text analysis of the textual data of China's central bank communication. The details are as follows.

Firstly , this paper collected all the written and oral communication texts of China's central bank from January 2003 to June 2019 , completed n-gram phrase segmentation and combination to form a field-specific dictionary in line with the unique communication habits of the central bank. Based on these textual data , this paper applied the latest hurdle distributed multinomial regression ( HDMR ) to model the relationship between all phrases and our target variables. HDMR model can simultaneously cope with the issue of high-dimension and sparsity , which allows us to fully extract the embedded information of central bank communication. Secondly , our results show that the dynamic factor model has the best forecasting performance for China's macroeconomy compared with the vector autoregression ( VAR ) model and factor-augmented VAR model. Thirdly , this paper took a dynamic factor model constructed with 152 macroeconomic indicators as benchmark , and investigated whether central bank communication can improve the model performance , the accuracy of nowcasting and forecasting. Our empirical results show that central bank communication helps to improve the in-sample fitness; for the out-of-sample forecasting performance , central bank communication information can significantly improve the prediction accuracy in the absence of the forecasting variable of historical information. After adding historical information , only the long-term forecasting shows a significant accuracy improvement with central bank communication. Finally , this paper verified the robustness by considering different communication samples or measurements.

This paper focuses on the text analysis of central bank communication and its application of macroeconomic forecasting , which is helpful for understanding the communication practice of China's central bank. Particularly , this paper provides some implications for establishing a monetary policy communication system , and implementing effective expectation management. For long-term macroeconomic forecasting , central bank communication can bring a significant improvement. That is , central bank needs to strengthen its research ability and information acquisition for providing more long-term analysis and outlook of future economic situation , guiding the public expectation and achieving its countercyclical adjustment. At the same time , this paper provides a new insight into nowcasting and forecasting with unstructured textual data. More real-time data , news and social media text , or daily forecasting of core macroeconomic variables can be introduced for future research.

**Keywords:** Central Bank Communication; Text Analysis; Macroeconomic Forecasting; Hurdle Distributed Multinomial Model

**JEL Classification:** C35 , C53 , C82 , E58

( 责任编辑: 昱 池 ) ( 校对: 晓 鸥 )