

# 通胀预期形成与信息黏性特征： 基于媒体新闻视角

郑挺国 范馨月 靳 炜 方匡南\*

**内容提要** 本文使用主流中文财经报刊中的近百万篇新闻报道构成语料库,利用潜在狄利克雷分配模型将新闻文本转换为结构化的高维新闻主题关注度序列,旨在探讨媒体新闻对通胀预期的驱动作用。为此,本文运用 LASSO、LASSO-UMIDAS 模型识别对通胀预期具有重要影响的新闻主题,并基于噪声信息模型估计时变信息黏性,通过将其与媒体报道持续期和信噪比进行联合建模分析,进一步考察了媒体新闻对通胀预期的影响机制。研究表明,媒体信息会从数值驱动和速度驱动两个方面对通胀预期产生重要影响。在数值方面,媒体信息中的经济增长、物价波动、国家要闻与金融信贷等多个主题在通胀预期的形成中发挥着重要作用;在速度方面,媒体报道的持续期和信噪比会对信息黏性产生显著的正向影响,进而影响通胀预期的调整速度。

**关键词** 通胀预期 信息黏性 数值驱动 速度驱动 媒体报道

## 一 引言

通货膨胀预期是经济中的个体基于自身知识水平,通过分析当前已有信息做出

\* 郑挺国(通讯作者):厦门大学宏观经济研究中心 经济学院统计与数据科学系 王亚南经济研究院福建省厦门市思明区思明南路422号 361005;范馨月:厦门大学王亚南经济研究院;靳炜:鹏华基金管理有限公司;方匡南:厦门大学经济学院统计与数据科学系。电子信箱:zhengt@ gmail.com(郑挺国);xinyue@stu.xmu.edu.cn(范馨月);jinwei@phfund.com.cn(靳炜);xmufkn@xmu.edu.cn(方匡南)。

作者感谢教育部人文社会科学重点研究基地重大项目(22JJD790050)、国家自然科学基金面上项目(71973110)以及全国统计科学研究重点项目(2022LZ37)的资助。感谢匿名审稿专家的建设性意见,当然文责自负。

的对未来价格水平的综合判断。近年来,央行不断强调要通过合理有效地引导公众预期来实现货币政策目标。正确理解通胀预期的特征、形成机制及驱动因素,对于稳定现实通胀率、制定科学的前瞻性货币政策具有重要的现实意义。

通胀预期具有两个现实特征:一是预期与真实经济紧密相关,根据预期理论,通胀预期具有自我实现的特征,即人们预期未来价格上涨往往会导致实际价格上涨(Carlson and Parkin, 1975),因此预期的不确定性不利于经济的持续稳定发展,还可能会导致实体经济衰退和资产价格波动(张健华和常黎, 2011);二是通胀预期与货币政策相互影响,预期的形成会受到宏观经济状况和当前货币政策的影响,而后者反过来也会对居民的心理及行为产生影响,因此通胀预期是各国央行制定货币政策的重要参考指标之一。

国内外学者对通胀预期的影响因素开展了大量研究。例如, Hori and Shimizutani (2005)使用日本的通胀预期调查数据与现实通胀率进行建模分析,发现当前的价格变动会对通胀预期产生显著影响。Leduc *et al.* (2007)利用美国二战后的数据,研究石油价格、财政政策和货币政策对通胀预期的影响。Ueda (2010)采用向量自回归模型分析宏观经济变量对美国和日本居民通胀预期的影响,发现能源价格、食品价格和货币政策均是影响通胀预期的因素。赵留彦(2005)基于理性预期假设,使用月度通胀率和利率序列估算了中国通胀预期,并证实通胀预期对货币持有量有显著负向影响。张健华和常黎(2011)基于新凯恩斯菲利普斯曲线,研究各类宏观经济变量对居民通胀预期的影响,发现真实通胀对预期的影响最大,而名义汇率、国际油价等对通胀预期的影响较小。洪智武和牛霖琳(2020)构建了混频无套利 Nelson-Siegel 利率期限结构拓展模型,提取中国通胀预期的期限结构,发现通胀预期受到真实通胀、货币增长率和全球食品价格的显著影响。然而,这些研究普遍存在三个方面的潜在问题:一是均选取少量的宏观经济变量与通胀预期数据进行建模,受传统计量模型的限制,所能容纳的变量维度较低。这种做法遗漏了许多其他变量的影响,忽视了其他信息在通胀预期形成中的重要作用。二是忽略了新闻媒体在预期形成过程中的潜在影响。由通胀预期的定义可知,信息在其形成过程中发挥着十分重要的作用,而新闻媒体不仅是信息传播的主要媒介,同时媒体也会在信息传播过程中对信息进行筛选、整合和包装,从而影响通胀预期的形成(Nimark and Pitschner, 2019)。三是部分研究依赖理性预期假设,认为公众在形成预期时能够获取到当前所有的有效信息,并据此进行理性判断,忽视了信息黏性的存在(Coibion and Gorodnichenko, 2015)。

为改进此前研究的不足,国内外学者从两个方面开展了探索。部分学者修正了

理性预期的假设,将信息黏性纳入模型进行分析。例如,Dopke *et al.*(2008)基于黏性信息菲利普斯曲线模型,估计法国、德国、英国和意大利的通胀预期更新频率。Reid (2015)运用流行病学模型探讨通胀预期的形成过程,并估计了信息黏性。Hung and Kwan(2021)基于香港的通胀及预期数据,对比黏性价格模型和黏性信息模型的有效性,发现黏性信息模型具有更好的预测表现,而且黏性信息模型更强调预期对真实通胀的影响。何运信等(2014)发现专家吸收信息的广度和速度强于普通居民,即专家预期更接近理性预期的假设。汪莉和王先爽(2015)在信息黏性模型的基础上估算通胀预期的波动,证实央行预期管理政策对于降低通胀预期有显著作用。张成思和党超(2015)基于流行病学模型研究中国的异质性通胀预期,发现居民预期具有较高的信息黏性,而专家预期不具有黏性。这些研究均表明,信息黏性是通胀预期形成机制中的一个重要概念,对信息黏性进行有效的估计,有助于理解信息对通胀预期的影响和驱动作用。然而,这些研究大多把信息黏性视作一个存在性的假设,然后基于菲利普斯曲线或流行病学模型分析通胀及通胀预期,很少估算信息黏性的数值。虽然有少部分文献采用通胀预期对其滞后项的回归系数作为信息黏性的估计值,但仅能得到一个长期恒定不变的信息黏性,无法适应不断变化的经济和信息环境。

国内外学者已经开始从媒体信息的角度研究通胀预期。例如,Lamla and Maag (2012)基于德国的数据构建贝叶斯学习模型,发现公众主要通过媒体报道获取信息并形成自己的预期。Pfajfar and Santoro(2013)研究新闻报道与通胀预期之间的关系,认为新闻与通胀预期之间的关系取决于公众对新闻的接受程度以及处理新信息的能力。Lamla and Lein(2014)在贝叶斯学习框架下研究媒体报道对通胀预期的影响,证实了密集的新闻报道能够帮助消费者获取更多的信息,从而提高他们通胀预期的准确性。卞志村和宗旭姣(2014)发现当媒体披露更多的信息时,公众通胀预期的准确性会增加,且在短期内媒体信息披露对通胀的影响程度大于央行信息。张成思和芦哲(2014)证实媒体舆论对通胀预期具有显著的驱动作用,且二者间存在双向互动关系。此外,还有较多经典研究也证实了媒体报道会对公众预期产生显著影响(Doms and Morin, 2004; Dräger and Lamla, 2017; Ehrmann *et al.*, 2017; Mazumder, 2021; 张成思等, 2014)。

新闻报道是信息传播中不可忽视的一个重要环节,基于新闻信息开展通胀预期的相关研究有现实需要。然而,由于受到文本分析方法和传统计量模型的限制,这些研究主要关注与通货膨胀相关的新闻报道,未考虑大量其他新闻话题包含的丰富信息。同时,这些研究大多从预测准确性或回归显著性角度分析新闻报道对通胀预期

的作用,未深入探究媒体信息与信息黏性的关系,因此无法全面分析媒体信息对通胀预期的驱动作用。因此,本文从另一个视角出发,不再局限于以通胀为主题的新闻报道,而是采用无监督的主题模型来提取高维新闻信息,分析其对通胀预期和信息黏性的影响。本文使用的主题模型能够将复杂、高维的语料库以无监督的方式转换为结构化时间序列,并通过多样化的主题尽可能地反映出语料库中的丰富信息,相比于存在一定主观性的有监督方法,主题模型能更客观、更充分地利用文本信息。

本文利用财经媒体大数据探究媒体新闻、通胀预期与信息黏性之间的关系。结合大数据背景及现实需要,本文综合利用媒体新闻信息对通胀预期进行驱动因素识别,通过噪声信息模型估算时变的信息黏性,并进一步分析媒体对信息黏性的影响,从数值驱动和速度驱动两个层面剖析了媒体在通胀预期形成中的作用。本文的贡献主要体现在以下4个方面:(1)使用无监督文本分析方法提取了高维新闻信息,目前国内基于有监督文本分析方法的研究已较为丰富,但尚无文献利用无监督狄利克雷分配模型(latent Dirichlet allocation, LDA)研究通胀预期;(2)构建适用于高维、混频数据的计量模型,全面考察了不同主题的媒体信息对通胀预期的驱动作用,并由此筛选出国家要闻、物价波动、经济增长、证券交易等具有重要影响的新闻主题;(3)尊重非理性预期的经验事实,构建噪声信息模型,由此估计出不同期限内的时变信息黏性,并进一步分析媒体报道与信息黏性之间的关系,从而更全面揭示媒体在通胀预期形成中的作用;(4)对比Larsen *et al.* (2021)的研究,本研究关注中文文本分析技术<sup>①</sup>,同时从驱动作用的视角详细分析媒体报道对通胀预期的影响机制,在此基础上,本文还使用混频数据模型,能够更加及时、有效地利用高频数据信息(刘汉和刘金全,2011;郑挺国和尚玉皇,2013)。

值得注意的是,本文从数值和调整速度两个方面研究媒体信息对通胀预期的影响。其中,数值驱动衡量的是媒体信息对通胀预期具体数值的作用,其内在逻辑体现为:信息→新闻媒体→通胀预期,即经济个体通过新闻媒体获取信息,并基于此形成对未来价格水平的预期。而速度驱动指的是媒体信息会影响公众对通胀预期的调整速度,具体体现为媒体对信息黏性的影响。这是由于媒体报道的差异会影响公众获取信息的效率和处理信息的难度,从而影响公众调整预期的速度。因此,本文通过提取高维主题时间序列作为媒体报道的量化指标,研究媒体信息对通胀预期及信息黏

<sup>①</sup> 中英文文本之间存在着较大的差异,国外文献使用的文本处理技术很难直接用于中文文本分析。本文在中文文本清洗、分词方面做了大量工作,详细步骤见本刊网站中的文章附录(下文简称在线附录)。

性的影响,以便从数值、速度两个方面探究通胀预期的形成机制。

本文剩余部分结构安排如下:第二部分为数据处理与描述;第三部分重点考察媒体信息对通胀预期的数值驱动作用,并由此筛选出具有重要影响的新闻主题;第四部分基于噪声信息模型估计时变的信息黏性;第五部分从高维的媒体信息中提取出媒体报道的整体持续期和信噪比,并分析二者对信息黏性的影响,从而考察媒体对通胀预期的速度驱动作用;最后是结论与政策建议。

## 二 数据处理与描述

### (一)通胀预期计算

本文使用的通胀预期数据为季度频率,样本区间为2002年第1季度至2021年第二季度。原始数据来源于中国人民银行《城镇储户问卷调查报告》中的“未来物价预期指数”,该报告以每季度向全国50个城市2万城镇储户开展的问卷调查为基础,其中公布的三个重要原始数据可用于计算通胀预期,即预期下个季度物价“上升”“下降”和“基本不变”的居民占比(2009年第二季度起新增“看不准”选项),将上述居民占比数据结合真实的通胀数据,采用Carlson and Parkin(1975)提出的概率法(C-P概率法)可计算得到居民通胀预期。

首先,令 $\psi_{1t}$ 、 $\psi_{2t}$ 、 $\psi_{3t}$ 、 $\theta_t$ 分别表示在 $t$ 时刻预期未来物价上升、下降、基本不变和看不准的居民占比。假设居民预期存在一个敏感区间,临界值为 $a$ ,且未来的真实通胀服从正态分布, $\pi_{t+1} \sim N(E_t \pi_{t+1}, \sigma_t^2)$ ,当 $\pi_{t+1} \in (-\infty, -a)$ 时,居民预期未来物价下降;当 $\pi_{t+1} \in (-a, a)$ 时,居民预期未来物价基本不变;当 $\pi_{t+1} \in (a, \infty)$ 时,居民预期未来物价上升。上述假设可表示为如下形式:

$$\psi_{1t} = (1 - \theta_t)P(\pi_{t+1} > a), \quad \psi_{2t} = (1 - \theta_t)P(\pi_{t+1} \leq -a) \quad (1)$$

经过标准化转换,上式等价于:

$$\begin{aligned} \frac{\psi_{1t}}{1 - \theta_t} &= P\left(\frac{\pi_{t+1} - E_t \pi_{t+1}}{\sigma_t} > \frac{a - E_t \pi_{t+1}}{\sigma_t}\right) = 1 - \Phi\left(\frac{a - E_t \pi_{t+1}}{\sigma_t}\right) = 1 - \Phi(N_{1t}) \\ \frac{\psi_{2t}}{1 - \theta_t} &= P\left(\frac{\pi_{t+1} - E_t \pi_{t+1}}{\sigma_t} \leq \frac{-a - E_t \pi_{t+1}}{\sigma_t}\right) = \Phi\left(\frac{-a - E_t \pi_{t+1}}{\sigma_t}\right) = \Phi(N_{2t}) \end{aligned} \quad (2)$$

其中, $\Phi(\cdot)$ 表示标准正态分布的累积概率密度函数,根据上式可以解得:

$$E_t \pi_{t+1} = -\frac{N_{1t} + N_{2t}}{N_{1t} - N_{2t}} a, \quad \sigma_t = \frac{2a}{N_{1t} - N_{2t}} \quad (3)$$

假设通胀预期具有长期无偏性,即在长期中通胀预期均值等于真实通胀的均值,可以得到:

$$\sum_{t=1}^T \pi_t = \sum_{t=1}^T \frac{N_{1t} + N_{2t}}{N_{2t} - N_{1t}} a \Rightarrow a = \frac{\sum_{t=1}^T \pi_t}{\sum_{t=1}^T \frac{N_{1t} + N_{2t}}{N_{2t} - N_{1t}}} \quad (4)$$

最后,根据 $\psi_{1t}$ 、 $\psi_{2t}$ 、 $\theta_t$ 可计算得出 $N_{1t}$ 和 $N_{2t}$ ;再结合真实通胀可计算得出预期的敏感区间临界值 $a$ ;由 $N_{1t}$ 、 $N_{2t}$ 和 $a$ 可计算得到通胀预期 $E_t \pi_{t+1}$ 。本文选择CPI的同比增长率作为真实通胀的衡量指标,并使用线性插值法处理《城镇储户问卷调查报告》中的缺失数据<sup>①</sup>。图1为真实通胀和本文测算得出的居民通胀预期序列。

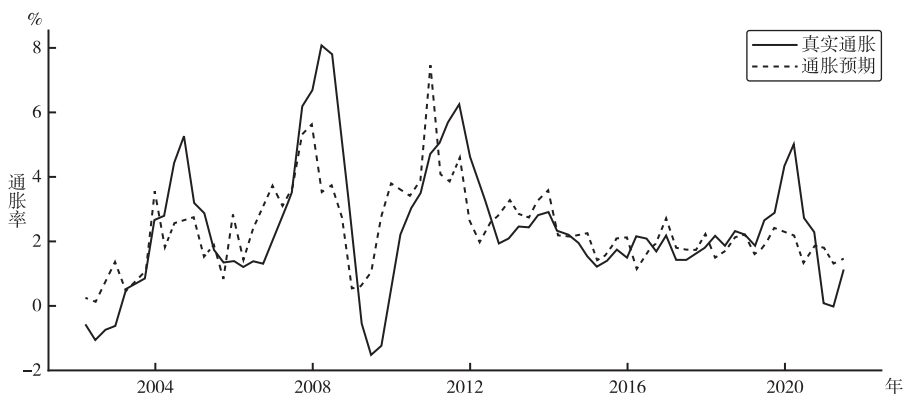


图1 通胀预期与真实通胀

## (二)媒体信息的提取

本文使用的新闻媒体语料库由每天发布的《经济日报》《经济观察报》和《21世纪经济报道》的近百万篇报道构成<sup>②</sup>,样本区间为2002年1月1日至2021年6月30日。这些报刊对中国的经济状况、政治事件进行实时报道,为居民提供了较为权威的新闻来源。为了更好地使用这些媒体信息进行建模分析,需要先将其转化为结构化数据,然后再将其转化为月度和季度数据。本文采用潜在狄利克雷分配模型(latent Dirichlet allocation, LDA)来处理文本数据,假设每天的新闻报道包含不同的主题,每

<sup>①</sup> 本文还参照张成思和党超(2015、2016)的做法处理缺失数据,并与线性插值法所得的结果进行对比,两种方法计算出的通胀预期高度重合,详见本刊网站中的在线附录。

<sup>②</sup> 本文在选择新闻媒体报刊时综合考虑了报刊的发行时间跨度、发行流通范围、数据可得性以及在经济金融领域的影响力。根据“2020胡润中国最具影响力财经媒体榜”,本文所选的报刊均入选中国十大最具影响力的财经报刊。其中《经济日报》是中国经济领域发行量最大、权威性和公信力最强的媒体。

个主题由不同的关键词构成,主题和关键词的先验分布均为狄利克雷分布。通过LDA模型,可以将复杂的新闻语料库转化为多个主题时间序列,用于建模分析。

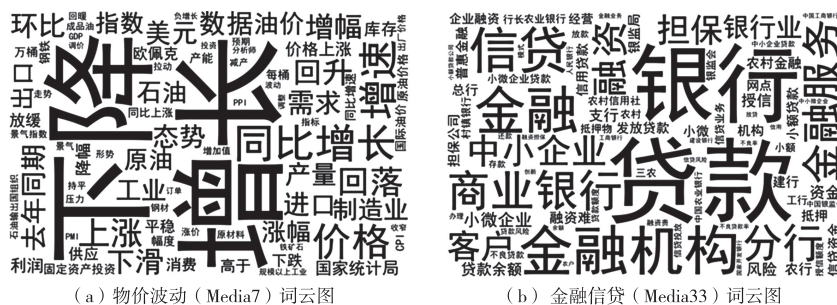
表1 新闻报道主题分类

类别	LDA 输出	主题命名	类别	LDA 输出	主题命名	类别	LDA 输出	主题命名
消费	Media01	网络消费	宏观经济	Media07	物价波动	通信交通	Media08	民航运输
	Media20	服装消费		Media09	经济增长		Media14	车辆运输
	Media53	旅游消费		Media10	外资外贸		Media21	货运物流
	Media54	餐饮消费		Media15	汇率变动		Media28	通信工程
生产投资	Media05	资本投资	国际合作	Media51	国际货币	科教文化	Media02	文旅发展
	Media11	家电制造		Media43	国际会议		Media12	文化艺术
	Media16	农业生产		Media55	对外政策		Media18	社会文化
	Media19	跨境投资		Media56	国际贸易		Media25	科技创新
	Media39	投资贷款		Media58	国际外交		Media30	影视行业
	Media42	工业产业		Media65	合作交流		Media31	产权保护
	Media60	房产市场		Media66	贸易政策		Media36	乡村教育
	Media61	汽车制造		Media04	全国两会		Media45	发明专利
	Media69	科技产品		Media13	人事任免		Media52	高等教育
	金融	Media06		港股市场	国家政治		Media35	党政建设
Media29		上市交易	Media37	国企改革		Media23	医疗健康	
Media33		金融信贷	Media49	财政税收		Media24	奥运赛事	
Media41		国际金融	Media59	国家要闻		Media26	质检保障	
Media46		股市行情	Media68	基层建设		Media32	意外事故	
Media48		证券交易	Media22	法律监管		Media34	博彩公益	
Media63		人寿保险	Media47	脱贫攻坚		Media40	债务纠纷	
Media70		金融理财	Media57	边疆建设		Media50	财务审计	
环境资源	Media17	环境保护	发展政策	Media62	福利政策			
	Media27	自然资源		Media64	乡村建设			
	Media38	绿色发展		Media67	招商引资			
	Media44	新型能源						

本文将主题数范围设置为10至120,并以10为增量分别估计了12个模型,参照Lau *et al.* (2014)的方法计算出每个模型对应的一致性得分,最终以最大化一致性得分为依据将主题数确定为70<sup>①</sup>。使用LDA模型提取出70个不同主题的媒体时间序

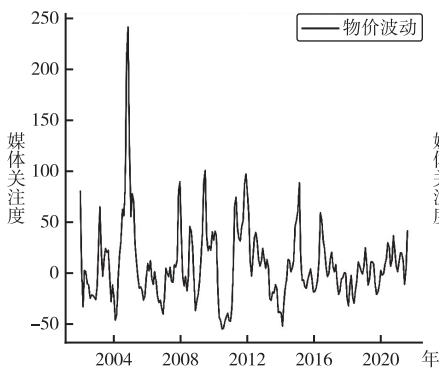
① 各模型的一致性得分见在线附录。此外,在线附录还针对主题数进行了稳健性检验,结论与正文一致。

列,然后按照以下步骤对这些数据进行预处理:首先,采用60天移动平均对数据进行平滑处理;其次,对于部分缺乏报道文本的日期,将当天的数据定义为缺失值并采用随机森林算法进行插值填补;再次,利用4分位范围准则(interquartile range)进行异常值处理;最后,对日度的主题时间序列分别按月、季取均值。如表1所示,根据主题中的关键词分布可以将70个主题分为宏观经济、生产投资、消费、金融等11个类别。

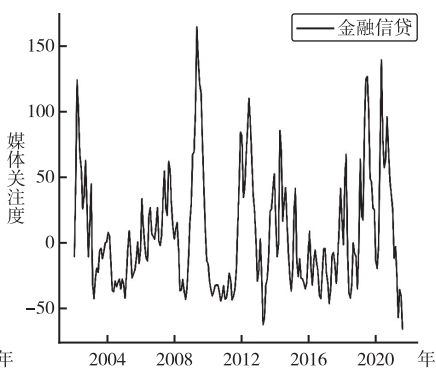


(a) 物价波动 (Media7) 词云图

(b) 金融信贷 (Media33) 词云图



(c) 物价波动 (Media7) 主题序列



(d) 金融信贷 (Media33) 主题序列

图2 关键词分布与主题时间序列

说明:子图(a)和(b)中字体越大表示该词语在描述特定主题时具有更大权重;子图(c)和(d)为月度媒体主题关注度序列,本文对LDA输出的主题时间序列进行了插值、去异常值等处理,并计算其同比增长率。

图2为利用LDA模型提取并处理后得出的第7个和第33个月度主题时间序列,按照关键词的分布,将这两个主题命名为“物价波动”和“金融信贷”<sup>①</sup>。

① 文本数据清洗、LDA原理详见在线附录。



### (三)经济数据处理

参照 McCracken and Ng(2016)构建的 FRED-MD 数据库,本文综合利用国家统计局、Choice 和 CEIC 数据库构建中国月度宏观经济数据库。该数据库覆盖了劳动力市场、房地产、消费等 9 个类别,包含 239 个月度宏观经济指标,样本区间与媒体信息一致,为 2002 年 1 月至 2021 年 6 月<sup>①</sup>。

所有变量均转换为平稳序列。首先将原始数据进行编号并排序,采用随机森林算法对多维数据缺失值进行联合填补,其次按照数据特征将数据划分为当月值、同比值、累计值和指数四个类别,然后分别进行如下处理:(a)对于当月值,将数值为 0 的观测点更改为缺失值,再使用样条法进行插值,然后计算其同比增长率;(b)对于同比值,直接使用原始的数据序列(c)对于累计值,保留每年 1 月的数据,并对 2 月-12 月的数据进行差分,计算得出当月值,然后按照当月值数据的处理方法进行计算;(d)指数类数据一般为价格指数(上年同月=100),如 CPI、PPI 等,用原始数据序列减去 100,即可得到其同比增长率。在进行上述处理后,选出 2002 年 1 月至 2021 年 6 月的数据构成经济数据的完整样本。

## 三 媒体信息对通胀预期的数值驱动

媒体报道中含有丰富的信息,运用 LDA 主题模型可提取出关于不同新闻主题的时间序列。为研究高维的媒体信息对通胀预期和真实通胀的影响差异,并进一步识别出对通胀预期具有重要驱动作用的新闻主题,本文参考 Tibshirani(1996)提出的套索回归方法(least absolute shrinkage and selection operator, LASSO),把高维媒体数据集映射到通胀预期,将解释力度较弱的变量系数压缩为 0,降低模型的复杂度,筛选出具有重要影响的新闻主题。该模型可以表示为如下形式:

$$E_t \pi_{t+h} = a + \beta NT + \varepsilon_t = a + \sum_{n=1}^N \sum_{j=0}^p \beta_{n,j} NT_{n,t-j} + \varepsilon_t \quad (5)$$

$$(\hat{a}, \hat{\beta}) = \arg \min_{a, \beta} \left[ \sum_t (E_t \pi_{t+h} - \beta NT) + \lambda \sum_{n=1}^N \sum_{j=0}^p |\beta_{n,j}| \right]$$

其中,  $E_t \pi_{t+h}$  表示在第  $t$  期形成的对第  $t+h$  期的通胀预期,  $NT = [NT_1, NT_2, \dots, NT_N]'$ ,  $NT_n = [NT_{n,t-p}, NT_{n,t-p+1}, \dots, NT_{n,t}]$ ,  $NT_{n,t-j}$  表示第  $n$  个新闻主题在

<sup>①</sup> 详细分类和变量描述省略备索。

$t-j$ 期的关注度,  $|\beta_{n,j}|$ 为  $L_1$  正则化, 能够对变量系数施加惩罚, 从而实现数据降维和变量选择。此外, 为减小估计偏误, 本文在后续估计中引入真实通胀的滞后项作为控制变量。现有的通胀预期、真实通胀和媒体信息分别为季度、月度和日度数据, 如果使用 LASSO 进行建模估计, 则需要将高频数据进行大量的加总处理, 这将会损失高频信息, 大大降低数据的利用效率和预测的时效性。因此, 本文参考 Uematsu and Shinya (2017) 的研究, 构建正则化无约束混频数据抽样模型 (LASSO-UMIDAS), 在实现数据降维和变量选择的同时, 兼顾混频数据的处理和运用。该模型可表示为如下形式:

$$E_t \pi_{t+h}^{(Q)} = a + \beta NT^{(M)} + \varepsilon_t = a + \sum_{n=1}^N \sum_{j=0}^{l_n} \beta_{n,j} NT_{n,t-j/m}^{(M)} + \varepsilon_t \quad (6)$$

$$(\hat{a}, \hat{\beta}) = \arg \min_{a, \beta} \left[ \sum_t (E_t \pi_{t+h}^{(Q)} - \beta NT^{(M)}) + \lambda \sum_{n=1}^N \sum_{j=0}^{l_n} |\beta_{n,j}| \right]$$

其中,  $E_t \pi_{t+h}^{(Q)}$  表示低频的季度通胀预期,  $NT^{(M)}$  表示高频的月度媒体主题时间序列,  $l_n$  为第  $n$  个主题时间序列的最大滞后阶数,  $NT_{n,t-j/m}^{(M)}$  为第  $n$  个月度主题时间序列在  $t$  期经频率对齐转换后得到的低频变量,  $m$  为高频解释变量与低频被解释变量之间的频率倍差,  $h$  为向前预测的步长。本文在研究信息对通胀预期的驱动作用时, 取  $h = 1$ 。在使用媒体信息进行预测时, 取  $h = 1$ 、 $h = 2$ 、 $h = 3$  及  $h = 4$ , 并采用 5 倍交叉验证选择其余参数。

为重点考察媒体信息对通胀预期的数值驱动作用, 本文构建了“媒体”和“媒体+经济”两个信息集分别加入 LASSO-UMIDAS 模型, 其中“媒体”信息集包括 70 个媒体主题时间序列和 1 个控制变量, 即 CPI 的滞后项。利用稀疏主成分分析 (sparse PCA) 从高维经济数据中提取出 5 个主成分 (记为 PC1-PC5) 作为经济变量的共同因子, 并将其与 CPI 滞后项以及 70 个媒体主题时间序列一起构成“媒体+经济”信息集, 以进一步考察在经济信息共存的情况下, 媒体信息对通胀预期的影响。

表 2 给出了通胀预期、真实通胀对媒体信息的回归结果。首先, 第 I 列的结果显示 70 个媒体主题时间序列中有 28 个被 LASSO-UMIDAS 模型选择, 其中有 11 个在多元线性回归中显著。显著性水平较高的主题包括: 经济增长、物价波动、国家要闻和金融信贷等。此外, 生产投资、科教文化与国际合作等类别中也有主题被选择, 可见新闻信息的多个主题均能对通胀预期发挥驱动作用, 这是因为人们会广泛关注各个方面的信息, 以此形成对未来价格走势的判断。其次, 对比第 I 列和第 II 列, 可以发现大多数第 I 列中被选择的主题, 在加入了经济信息后依然会被继续选择, 这表明能够影响通胀预期的媒体信息, 在有经济变量共存的情况下也依然能发挥作用。

值得注意的是,LDA是一种无监督的文本分析方法,通过该方法计算出的主题时间序列大多没有明确的符号或情绪导向,不能直观地根据系数符号来判断特定主题对通胀预期或真实通胀的影响方向,需要结合各个主题的词云图和序列特征来进行详细分析<sup>①</sup>。尽管如此,LDA仍然在处理复杂、高维语料库时具备明显的优势,同时能相对客观、无损地利用丰富的文本信息。

表2 媒体信息对通胀预期与真实通胀的影响

类别	主题	通胀预期		真实通胀	
		I 媒体	II 媒体+经济	III 媒体	IV 媒体+经济
宏观经济	物价波动	-0.8592**	-0.4366		
	经济增长	0.6806**	0.3622**		
	汇率变动		0.1438		
生产投资	资本投资	0.0872		0.3702	0.3560
	家电制造	-0.2111	-0.1284		
	跨境投资	0.3257		0.0356	0.0323
	投资贷款	0.5578	0.5951		
	工业产业	0.9872	0.5137	0.2656	0.2236
消 费	旅游消费	0.5388*			
金 融	上市交易			0.0061	0.0214
	金融信贷	-0.3766**	-0.2824	-0.2916***	-0.2997**
	国际金融			0.0003	0.0026
	股市行情			-0.0851***	-0.1124***
	证券交易	1.5289**	1.7555**		
	人寿保险	-0.5782	-0.2221	-0.0596***	-0.0295
	金融理财	1.2050**	0.5620**	0.1778***	0.1248**
科教文化	文化艺术	-0.5571	-0.3767		
	社会文化	-0.0196	-0.0092		
	科技创新	1.0146	0.8354		
	高等教育	0.0505			
通信交通	民航运输	0.1399	0.3030		
	通信工程			0.0076	0.0167
环境资源	环境保护			0.1449	0.1551**
	国际会议			0.0029	0.0025
国际合作	对外政策	0.3387			
	国际外交			-0.2636***	-0.2449**
	合作交流	0.3169			

① 本文结合主题时间序列的词云图和序列特征讨论了表2中的回归系数符号,见在线附录。

(续表)

类别	主题	通胀预期		真实通胀	
		I	II	III	IV
		媒体	媒体+经济	媒体	媒体+经济
国家政治	全国两会	-0.3328*	-0.1390		
	党政建设	-0.3362	-0.3027		
	财政税收	-0.8417**	-0.7712**		
	国家要闻	1.5840***	1.7790***		
	基层建设	-0.0086	-0.0302		
其他	自然灾害	1.0421	0.8554*		
	质检保障	0.5837	0.6965		
	博彩公益	-0.4162**	-0.3595		
	财务审计	-0.3164*	-0.4162		
概要	控制变量	3.6665***	2.9445***	9.6269***	9.5861***
	筛选数量	28/70	23/70	14/70	14/70
	主要模型	LASSO-UMIDAS	LASSO	LASSO-UMIDAS	LASSO

说明:(1)表中数值为正则化回归模型的估计系数;(2)\*、\*\*和\*\*\*分别表示变量在10%、5%和1%的置信水平下显著;(3)列I和列III的控制变量为通胀滞后项,列II和列IV在此基础上加入了经济变量;(4)估计时所用的数据包括宏观经济数据、媒体主题关注度、真实通胀及通胀预期,除通胀预期为季度频率外,其他数据均为月度频率。

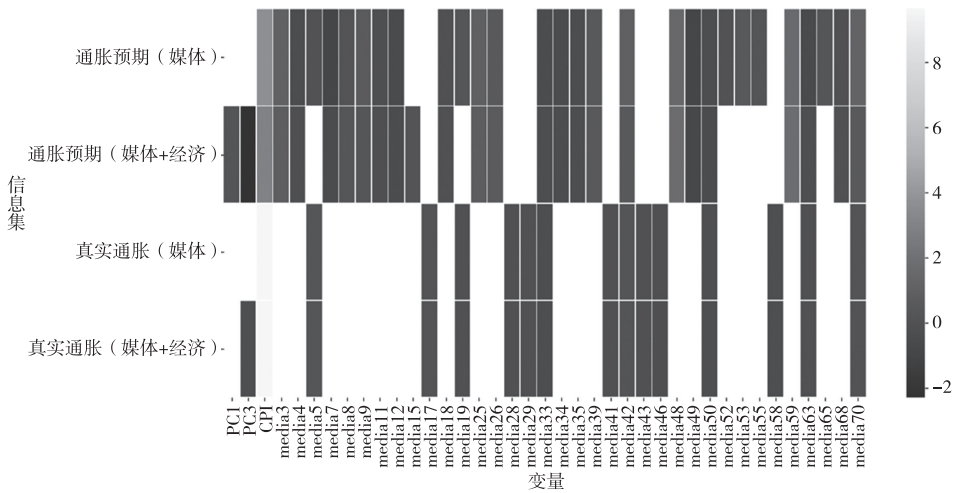


图3 通胀预期、真实通胀系数的矩阵热力图

说明:(1)横轴表示在正则化回归中显著的变量,其中PC1和PC3分别代表第1个和第3个宏观经济因子;(2)左侧纵轴表示在对通胀预期和真实通胀进行预测时使用的信息集,例如“通胀预期(媒体)”表示仅使用媒体数据对通胀预期进行预测。

图3是在不同信息集下预测通胀预期和真实通胀得到的系数矩阵热力图,可以更加直观地观察到媒体信息的影响差异。结合回归结果可知,影响差异体现在两个方面。其一,从被选择的主题数量来看,媒体信息对通胀预期具有更强的影响力。对比表2的第I列和第III列,当使用真实通胀与新闻信息一起构建LASSO时,仅有13个主题被选择,相较于第I列的28个主题而言数量显著减少,这在一定程度上表明媒体信息对通胀预期的影响要比对真实通胀的影响更为普遍。张成思和芦哲(2014)也得出了类似的结论,他们研究媒体舆论、公众预期与现实通胀率三者间的关系,发现媒体舆论与通胀预期之间存在着显著的双向互动关系,而媒体信息与现实通胀率之间的互动关系并不明显。这印证了本文的结论,即公众基于信息形成的通胀预期更容易受到媒体报道的影响。其二,从被选择主题的类别分布来看,通胀预期选出的主题涉及宏观经济、生产投资、科教文化、国家政治等各个方面,而真实通胀选择的主题大多集中于生产投资和金融这些与现实经济紧密相关的类别<sup>①</sup>。进一步地,可以从通胀预期与真实通胀各自的特征出发,解释这一现象产生的原因。通胀预期是公众基于既有信息形成的主观判断,受到经济状况、信息传播、信息处理能力等多方面因素的影响,因为不同个体所处环境不同,可能会产生不同的预期判断,因此通胀预期的影响因素较为广泛。而现实通胀则与之相反,它是对经济现状的直观反映,虽然会间接地受到信息及公众行为的影响,但其数值通常由真实经济直接决定。因此,相较现实通胀而言,媒体对通胀预期的影响更为直接和广泛。

为验证结论的稳健性,本文也考察了媒体信息对居民通胀预期、专家通胀预期、月度真实通胀和季度真实通胀的影响差异,所得结论再一次验证了媒体信息对通胀预期存在显著影响<sup>②</sup>。

## 四 信息黏性的估计

Calvo(1983)提出了黏性定价理论,Mankiw and Reis(2002)在其基础上引入了信息黏性的概念。现实中信息的更新具有一定的成本,人们在判断未来经济形势时,不会立即获取并信任最新的信息,而是综合参考过去和当前的信息,他们对于旧信息的

<sup>①</sup> 与真实经济相关的主题包括“宏观经济”“生产投资”和“金融”,但由于其报道内容具有不同的导向性和时效性,因此对真实通胀也具有不同的前瞻性预测能力。详细讨论见在线附录。

<sup>②</sup> 详细结果见在线附录。

依赖程度被定义为“信息黏性”，可以用于衡量人们对于通胀预期的调整速度。信息黏性越大，人们越依赖过去的旧信息，预期的调整速度越低。

现有文献对信息黏性的估计主要有三种方法：第一种方法以 Mankiw and Reis (2002) 提出的黏性信息菲利普斯曲线为基础，利用工资、失业率、利率等宏观经济数据估计信息黏性；第二种方法是 Carroll (2003) 提出的流行病学模型，其基本思想是假设普通公众主要参考专家预期来形成各自的预期判断，然而当专家进行预期更新时，仅有部分公众能及时获取到信息并更新自身预期，其余公众将继续保持之前的预期，因此关于通胀的新信息在公众中以类似流行病传染的模式传播，未能及时更新预期的居民占比可用于衡量信息黏性；第三种方法是基于 Coibion and Corodnichenko (2015) 构建的噪声信息模型 (noise information model)，该模型从媒体信息角度出发，认为媒体报道是公众信息的主要来源，且报道中的信息和噪声决定了信息黏性。这三种方法都能够估计出信息黏性的具体数值，但前两种方法并未考虑媒体在通胀预期形成中的作用，而且忽略了信息黏性可能存在的时变特性。因此，本文参考噪声信息模型，并结合国内数据的特点，估计不同期限内的信息黏性。

假设人们主要通过媒体来获取有关经济状况和未来价格的信息，并据此做出自己的预期判断。虽然媒体信息是高维的，但人们能够对这些信息进行整理聚合，可用  $\pi_i^N$  来表示在  $t$  时刻人们所获取的全部媒体信息，第  $i$  个个体在  $t$  时刻所获取的有关未来通胀的信号 ( $S_{it}$ ) 为：

$$S_{it} = \pi_i^N + \omega_{it}, \quad \omega_{it} \sim N(0, \sigma_{\omega_{it}}^2) \quad (7)$$

其中， $\omega_{it}$  为经济个体的异质性噪声，它描述不同个体获取、衡量和处理新闻信息的差异，这些差异的大小与当前的经济状况有关。受异质性噪声的影响，不同个体会形成对未来通胀的不同预期。此外，新闻报道具有一定的持续性，根据其时间序列特性，可以构建一个如下自回归过程：

$$\pi_i^N = \rho_i^N \pi_{i,t-1}^N + v_i^N, \quad v_i^N \sim N(0, \sigma_{v_i}^2) \quad (8)$$

此处， $t = 1, 2, \dots, T$  为时间下标， $\rho_i^N$  和  $\sigma_{v_i}^2$  为时变参数，较大的  $\rho_i^N$  代表媒体在一段时间内持续地报道某些重大经济或政治事件；较大的  $\sigma_{v_i}^2$  意味着这一时期的新闻报道波动较大，难以预测。接下来，假设媒体根据真实的经济状况  $\pi_t$  进行新闻报道，即新闻信息与真实通胀之间存在着如下关系：

$$\pi_i^N = \pi_t + \alpha_t, \quad \alpha_{t+j|t} = 0 \quad (9)$$

其中， $\alpha_t$  代表时间固定效应，能够描述潜在的媒体报道偏差。 $\alpha_{t+j|t} = 0$  意味着，预期未来的媒体报道与真实通胀一致，但实际上经济个体获取的信息与真实通胀之间

存在一定的偏差。由式(7)和式(8)可构成一个基础的状态空间模型,通过卡尔曼滤波进行推导,并联合式(9)进行迭代,可以得到如下表达式<sup>①</sup>:

$$\pi_{t+h} - E_t \pi_{t+h} = -\alpha_{t+h} + \beta_t (E_t \pi_{t+h} - E_{t-j} \pi_{t+h}) + e_t \quad (10)$$

其中,  $\pi_{t+h} - E_t \pi_{t+h}$  表示在  $t$  期对第  $t+h$  期通胀的预期误差,  $E_t \pi_{t+h} - E_{t-j} \pi_{t+h}$  表示对通胀的预期平均修正,  $\beta_t$  表示信息黏性。然而,由于当前可获取的通胀预期数据有限,仅能得到提前1期的通胀预期,即  $E_t \pi_{t+1}$ 。为了估计不同期限内的信息黏性,本文使用通胀的向前  $h$  步预测值  $\pi_{t+h|t}$  作为提前  $h$  期通胀预期  $E_t \pi_{t+h}$  的近似值<sup>②</sup>,由此可用如下时变系数模型估计近似的信息黏性:

$$\begin{aligned} y_t &= c_t + r_t x_t + e_t, \quad e_t \sim N(0, \sigma_e^2) \\ r_t &= r_{t-1} + v_t \end{aligned} \quad (11)$$

其中,  $y_t = \pi_{t+h} - \pi_{t+h|t}$  表示预期误差的近似值,  $x_t = \pi_{t+h|t} - \pi_{t+h|t-j}$  表示预期平均修正的近似值,  $r_t$  为真实的信息黏性  $\beta_t$  的近似值。此处采用 Logit 函数将近似值  $r_t$  映射到  $\beta_t$ , 得到取值介于 0 和 1 之间的信息黏性估计值。本文在状态空间模型的框架下采用极大似然方法对上述式子进行估计,具体步骤见在线附录。

为估计时变的信息黏性,本文截取 2002 年 1 月至 2015 年 12 月的数据用于样本内拟合,剩余 2016 年 1 月至 2021 年 6 月的数据进行样本外预测,并基于“媒体+经济”信息集,使用 LASSO 方法对月度真实通胀进行向前 1 步至 4 步的递归预测(详见在线附录)。图 4 展示了向前 1 步预测的系数矩阵热力图。

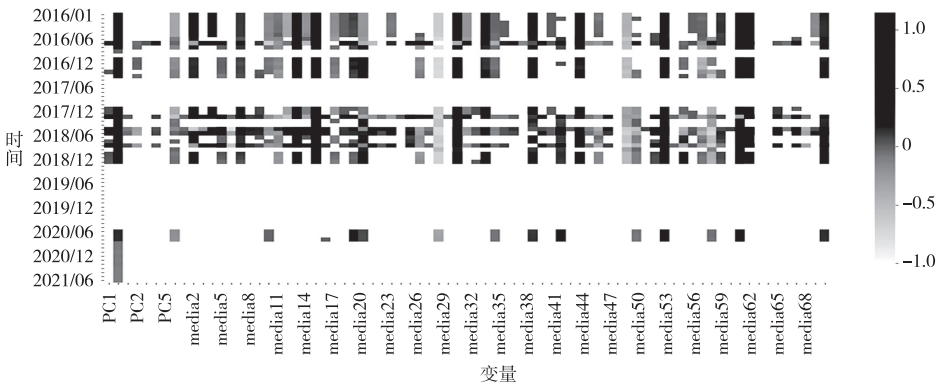


图 4 通胀向前 1 步预测的系数矩阵热力图

① 推导细节见在线附录。

② 本文还参照张成思和芦哲(2014)的思想,采用 TVP-AR 模型估计季度信息黏性,所得结果均具有较强的稳健性,详见在线附录。

图 5a 给出了不同期限下时变系数  $r_t$  的平滑估计值, 图 5b 是对应时变信息黏性  $\beta_t$  的估计值。可以看出随着期限的增加, 信息黏性通常趋于降低, 而最短的期限  $j=1$  对应最大的信息黏性, 这表明期限越短, 信息的更新频率越慢, 公众越难以接受新的信息。进一步分析可以发现, 2020–2021 年信息黏性呈现显著的上升趋势, 该时间点对应 2020 年初爆发的新冠肺炎疫情。究其原因, 在经济不确定性较大的情况下, 信息黏性将会上升, 可以从两个方面来分析这一现象的成因: 一是信息更新成本的增大, 经济体在遇到较重大的负面事件时, 媒体将在一段时间内持续集中地报道与该事件相关的新闻, 很少更新其他新的信息, 因此, 公众很难获取到新的信息来调整自己的通胀预期。二是公众辨别信息的难度增大, 在经济不景气或不确定性较高的时候, 信息中往往有较大的噪声, 公众难以通过混杂的信息进行有效判断和分析, 因此他们对当前新信息的信任减弱, 会更多地参考过去的信息, 导致信息黏性上升。后文分析媒体对信息黏性的驱动作用时亦可以佐证该结论。

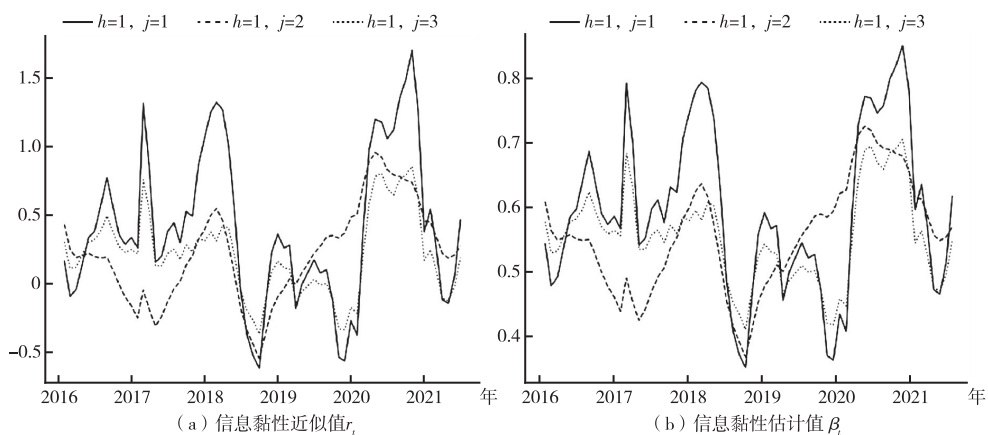


图5 时变的信息黏性

## 五 媒体对信息黏性的驱动作用

新闻媒体在信息的获取、整理和传播过程中有重要作用, 其在不同时间段内报道内容的差异会使得公众获取到不同的信息, 而且会影响公众获取信息的效率和处理信息的难度, 从而影响公众对通胀预期的调整速度。而基于现实通胀及通胀预期所估计出的信息黏性, 可以作为公众通胀预期调整速度的重要衡量指标。此前的结论



已经表明信息黏性是时变的,因此,本部分进一步探究信息黏性的变化是否受到媒体报道的影响。

区别于前文的数值驱动分析,通过研究媒体报道对信息黏性的影响,可进一步分析媒体信息对通胀预期的速度驱动作用。因此,本文首先利用70个媒体主题时间序列提取媒体报道的整体持续期和信噪比,然后构建多元线性回归模型,考察媒体信息对通胀预期的速度驱动作用。为了从高维的媒体信息中提取出媒体报道的持续期和信噪比,本文参考Koop and Korobilis(2013)提出的时变系数向量自回归模型(TVP-VAR),并使用指数加权移动平均(EWMA)刻画时变的方差波动,在状态空间模型框架下估计时变系数和方差,分别用以刻画持续期和信噪比。具体的建模步骤如下:

首先,对每一个媒体主题时间序列 $i = 1, 2, \dots, N$ 构造如下TVP-AR模型:

$$\begin{aligned} NT_{i,t} &= \rho_{i,t} NT_{i,t-1} + v_{i,t}, & v_{i,t} &\sim N(0, \sigma_{i,t}^2) \\ \rho_{i,t} &= \rho_{i,t-1} + u_{i,t}, & u_{i,t} &\sim N(0, \sigma_{\rho_i}^2) \\ \sigma_{i,t}^2 &= k\sigma_{i,t-1}^2 + (1-k)v_{i,t}^2 \end{aligned} \quad (12)$$

其中, $k$ 为衰减因子,一般取值为 $0.9 \leq k < 1$ 。将该模型表示为状态空间的形式,即可通过卡尔曼滤波构造出似然函数,通过极大似然估计得出参数值。由此,可得到每个新闻主题的报道持续期 $\rho_{i,t}$ 和噪声 $\sigma_{i,t}^2$ 。

接下来,对信息集中的新闻主题进行加总,得到关于该信息集的整体持续期 $\rho_t$ 和信噪比 $n_t$ 。在后续分析中,选取的信息集包括所有媒体主题时间序列和通过LASSO-UMIDAS选择出的媒体序列。在特定信息集中,持续期和信噪比的加总方法为:

$$\rho_t = \sum_{i=1}^N \omega_i \rho_{i,t}, \quad n_t = \sum_{i=1}^N \omega_i n_{i,t}, \quad n_{i,t} = \tau_{i,t} / \sigma_{i,t}, \quad \omega_i = R_i^2 / \sum_{i=1}^N R_i^2 \quad (13)$$

其中, $\omega_i$ 表示第 $i$ 个主题的权重,通过通胀预期对新闻主题的简单线性回归中的拟合优度 $R^2$ 计算得出, $\tau_{i,t}$ 为第 $i$ 个主题在LDA中的后验分布标准差。

最后,使用简单线性回归来分析媒体报道对信息黏性的驱动作用。模型表达式如下:

$$\beta_t = c + \gamma_1 \rho_t + \gamma_2 n_t + \varepsilon_t \quad (14)$$

为保证结果的稳健性,本文分别基于两个不同的媒体信息集提取持续期和信噪比。第一个信息集为 $S^w$ ,包含全部70个媒体主题时间序列;第二个信息集为 $S^{MSelect}$ ,包含对通胀预期具有重要影响的媒体主题时间序列,即表2第1列中所筛选出的28个主题。图6为从信息集 $S^w$ 提取出来的媒体报道持续期和信噪比。可见,2019年年末和2020年年初媒体报道持续期和信噪比均处于较高位置。

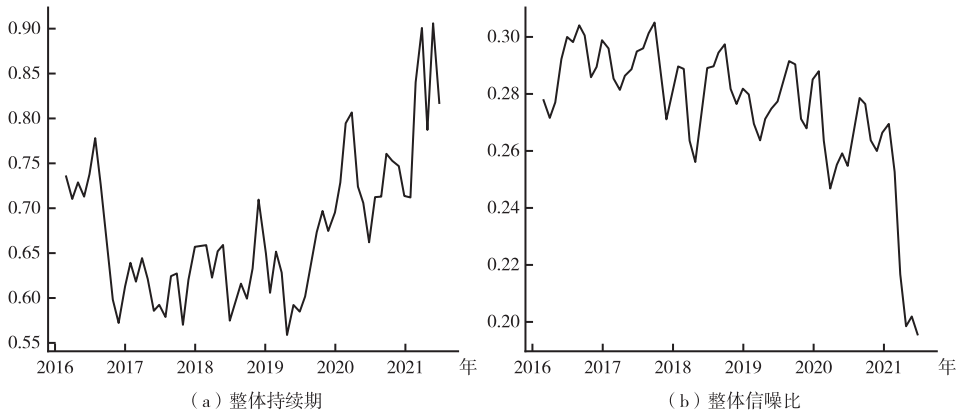


图6 基于全部媒体主题的持续期与信噪比

对两个信息集中提取出的持续期和信噪比进行标准化处理,然后与信息黏性一起构造多元线性回归模型,结果见表3。首先,媒体报道持续性的系数显著为正,说明信息黏性与持续期之间存在正相关关系,即当媒体报道的持续期增加,信息黏性将会增加。这是由于当媒体报道的持续性增加时,媒体会在某一段时间内持续地报道某一些特定主题的新闻,因此公众难以获取来自其他主题的新信息,信息获取难度的增大导致信息黏性的上升。其次,媒体报道信噪比的系数也显著为正,说明信息黏性与信噪比之间存在正相关关系,即当媒体信息中存在较多噪声时,公众处理信息的难度增大,同时对于当前信息的信任程度减弱,因此公众将更多地参考过去的信息形成对

表3 媒体对信息黏性的影响分析

变量类型	变量名称	所有媒体主题, $S^M$		通胀预期所选主题, $S^{MSelect}$	
		I	II	III	IV
媒体信息	媒体报道持续期	0.5025***	0.4217***	0.4577***	0.3714***
	媒体报道信噪比	0.5607***	0.2021***	0.5755***	0.1132
	<i>EPU</i>		0.0086		-0.0226
	<i>CPI</i>		-0.0392*		-0.0077
控制变量	<i>PC1</i>		0.1551		0.0937
	<i>PC2</i>		0.4265***		0.4478***
	<i>PC3</i>		0.1309		0.1868

说明:本表使用第1期信息黏性,其余期限的估计结果见在线附录;考虑到其他因素也可能影响信息黏性,第II列和第IV列加入经济政策不确定性指数*EPU*(Baker *et al.*, 2013)、*CPI*滞后项和宏观因子作为控制变量。

未来的预期。这与前文的结论相符,即经济不景气或不确定性增大的时期,信息黏性往往显著上升。

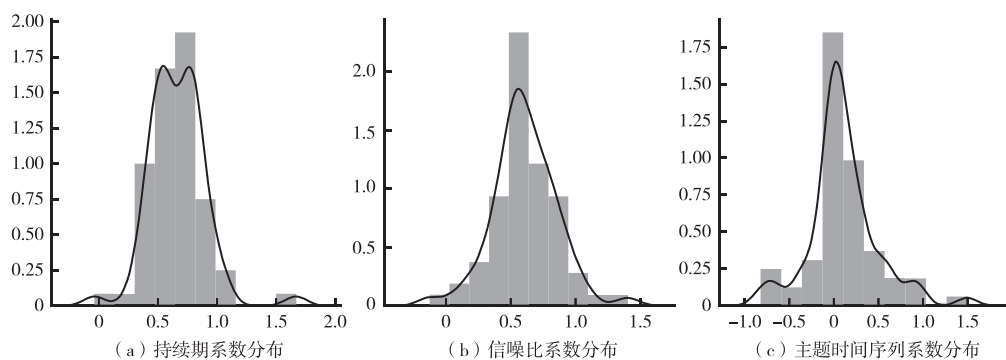


图7 信息黏性与单个主题时间序列

媒体报道影响信息黏性的另一种可能的解释是,报道内容本身会影响信息黏性。为了排除这一可能,并进一步验证媒体报道与信息黏性之间的关系主要来源于报道的持续期和信噪比,本文考察了单个媒体主题持续期、信噪比和其原始主题关注度序列对信息黏性的影响。图7分别展示了70个主题时间序列的影响系数分布。如图7(a)和7(b)所示,大多数新闻主题持续期和信噪比的系数都介于0和1之间,且其分布的峰值在0.5和0.7之间,这再一次验证了前文的结论,即媒体报道的持续期和信噪比对信息黏性具有显著的正向影响。而由图7(c)可知,大多数主题原始时间序列对信息黏性的影响并不显著,即新闻报道本身的内容与信息黏性之间没有明显的关联。

## 六 结论及政策建议

通货膨胀预期是人们基于既有信息对未来物价水平形成的主观判断,它与真实经济相互影响、紧密相关,是各国央行制定货币政策的重要参考指标之一。而新闻媒体作为信息传播的主要媒介,在信息搜集、整理和传播过程中发挥着重要作用,是公众获取信息的主要来源,能够对通胀预期的形成产生不可忽视的影响。因此,从信息的视角出发,厘清媒体报道、通胀预期和信息黏性三者之间的关系,有助于更科学地开展通胀预期的相关研究。本文将国内主要财经报刊的新闻报道转换为高维主题时

间序列,识别对通胀预期具有重要驱动作用的新闻主题,并进一步引入信息黏性的概念,同时提取出媒体报道的持续期和信噪比,揭示了媒体报道对信息黏性的重要影响,从而说明媒体信息能够从数值驱动和速度驱动两个方面影响通胀预期。

本文的主要结论可概括如下:首先,对于通胀预期和真实通胀而言,媒体信息均能发挥重要影响,但其对通胀预期的影响更为普遍。具体而言,影响通胀预期的新闻主题广泛地分布于宏观经济、生产投资、科教文化、国家政治等各个不同类别,而能够对真实通胀产生影响的新闻主题则大多集中于生产投资和金融等这些与现实经济紧密相关的类别。其次,公众在处理信息形成预期判断时存在信息黏性,而且由于经济环境的不断变化会带来不同的信息更新成本和处理难度,因此信息黏性是时变的,当经济不景气或不确定性增大时,信息黏性通常升高。最后,媒体信息能够对通胀预期产生数值驱动和速度驱动。在数值方面,媒体报道中的诸多主题,例如经济增长、物价波动、国家要闻和金融信贷等在公众的通胀预期形成过程中有重要作用;在速度方面,媒体报道的持续期和信噪比均会对信息黏性产生显著正向影响,即更长的持续期和更大的信噪比将降低公众对通胀预期的调整速度。

本文重点关注媒体报道中的新闻信息,并从数值驱动和速度驱动两个方面分析媒体信息在通胀预期形成中的作用,所得结论对现实中的预期管理、政策制定有所帮助,同时对于通胀预期的相关研究具有参考价值。据此,可以得出如下政策建议:

第一,在新闻媒体管理与大数据战略方面,要重视新闻媒体在信息传播、舆论引导和预期形成中的作用,加强媒体管理,推动媒体融合,这也是建立新时代通胀预期管理体系的必然要求。媒体报道会影响公众的信息黏性,从而改变通胀预期的调整速度。当经济不确定性增大或信息中噪声成分增多时,公众对于媒体信息的处理难度增大,同时对新信息的信任程度降低,导致信息黏性增大,使得通胀预期更大程度地偏离理性预期。因此,要科学地管理和引导公众预期,必须进一步增强主流媒体新闻报道的时效性和准确性,使新闻媒体具有更强大的传播力、引导力和公信力,帮助公众及时有效地获取信息,从而形成准确的预期判断。

第二,在预期管理方面,要建立健全国家宏观经济预期管理政策体系,并进一步扩展预期管理主体,加强政府各部门间的沟通协作。公众在判断未来的价格走势时,会广泛地参考不同类别的信息,如宏观经济、金融、国家政治、科教文化等。因此,预期管理的措施不应局限于货币政策,还应关注与宏观经济调控有关的各个领域。相应地,预期管理的主体也不应局限于中央银行,还应包括财政部、证监会、统计局等其他部门,在央行发挥核心作用的同时,其他部门也应当重视并支持预期管理工作。

第三,在预期调控和信息披露方面,政府各部门需与公众保持适度沟通,合理地进行信息披露,科学地引导和管理通胀预期。由于信息更新存在时滞,并且不同个体对政策和经济形势的解读存在差异,公众难以形成稳定而理性的预期。权威部门应当及时、准确地公布相关信息,提高公众获取信息的效率,使公众预期向理性预期靠近。此外,公众处理信息的能力有限,政府在进行信息披露的同时还需注意沟通的力度,既要帮助公众排除不实、模糊信息的干扰,又要避免过度沟通导致的反应过度和信息混乱。

综合以上分析,政策制定和宏观经济研究都需要与时俱进,时代的发展对中国未来的预期管理提出了更高的要求。在实践方面,政府部门需要从制度创新、政务公开、部门协作等方面探索构建更为科学的管理体系;在理论方面,学者需要从数据应用、模型理论、分析方法等方面来改进通胀预期的相关研究。本文利用新闻文本及宏观经济大数据对通胀预期进行探索研究,为后续的政策制定和理论发展提供了新的思路和参考依据。

### 参考文献:

- 卞志村、宗旭姣(2014):《公众学习、媒体信息披露与通胀预期形成》,《金融评论》第1期。
- 何运信、沈宏、耿中元(2014):《居民与专家预期通货膨胀预期的差异及两者间关系——流行病学模型在中国的检验》,《金融研究》第5期。
- 洪智武、牛霖琳(2020):《中国通货膨胀预期及其影响因素分析——基于混频无套利 Nelson-Seigel 利率期限结构扩张模型》,《金融研究》第12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。
- 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。
- 张成思、党超(2015):《异质性通胀预期的信息粘性与信息更新频率》,《财贸经济》第10期。
- 张成思、党超(2016):《谁的通胀预期影响了货币政策》,《金融研究》第10期。
- 张成思、芦哲(2014):《媒体舆论、公众预期与通货膨胀》,《金融研究》第1期。
- 张成思、佟冠良、芦哲(2014):《媒体报道、通胀预期与新凯恩斯菲利普斯曲线》,《金融评论》第3期。
- 张健华、常黎(2011):《哪些因素影响了通货膨胀预期——基于中国居民的经验研究》,《金融研究》第12期。
- 赵留彦(2005):《中国通胀预期的卡尔曼滤波估计》,《经济学(季刊)》第4期。
- 郑挺国、尚玉皇(2013):《基于金融指标对中国 GDP 的混频预测分析》,《金融研究》第9期。
- Baker, S.; Bloom, N.; Davis, S.J. and Wang, X. "A Measure of Economic Policy Uncertainty for China." University of Chicago working paper, 2013.
- Calvo, G.A. "Staggered Prices in a Utility-maximizing Framework." *Journal of Monetary Economics*, 1983, 12 (3), pp. 383-398.
- Carlson, J.A. and Parkin, J.M. "Inflation Expectations." *Economica*, 1975, 42(166), pp. 123-138.

- Carroll, C.D. "Macroeconomic Expectations of Households and Professional Forecasters." *The Quarterly Journal of Economics*, 2003, 118(1), pp. 269–298.
- Coibion, O. and Gorodnichenko, Y. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *The American Economic Review*, 2015, 105(8), pp. 2644–2678.
- Doms, M. and Morin, N.J. "Consumer Sentiment, the Economy, and the News Media." Federal Reserve Bank of San Francisco working paper, 2004–09(204).
- Dopke, J.; Dovern, J.; Fritsche, U. and Slacalek, J. "Sticky Information Phillips Curves: European Evidence." *Journal of Money, Credit and Banking*, 2008, 40(7), pp. 1513–1519.
- Dräger, L. and Lamla, M. J. "Imperfect Information and Consumer Inflation Expectations: Evidence from Microdata." *Oxford Bulletin of Economics and Statistics*, 2017, 79(6), pp. 933–968.
- Ehrmann, M.; Pfajfar, D. and Santoro, E. "Consumers' Attitudes and Their Inflation Expectations." *International Journal of Central Bank*, 2017, 13(1), pp. 225–259.
- Hori, M. and Shimizutani, S. "Price Expectations and Consumption under Deflation: Evidence from Japanese Household Survey Data." *International Economics and Economic Policy*, 2005, 2, pp. 127–151.
- Hung, T.H. and Kwan, Y.K. "Hong Kong's New Keynesian Phillips Curve: Sticky Information or Sticky Price?" *Pacific Economic Review*, 2021, 27(1), pp. 1–14.
- Koop, G. and Korobilis, D. "Large Time-varying Parameter VARs." *Journal of Econometrics*, 2013, 177(2), pp. 185–198.
- Lamla, M.J. and Maag, T. "The Role of Media for Inflation Forecast Disagreement of Households and Professional Forecasters." *Journal of Money, Credit and Banking*, 2012, 44(7), pp. 1325–1350.
- Lamla, M.J. and Lein, S.M. "The Role of Media for Consumers' Inflation Expectation Formation." *Journal of Economic Behavior & Organization*, 2014, 106, pp. 62–77.
- Larsen, V. H.; Thorsrud, L. A. and Zhulanova, J. "News-driven Inflation Expectations and Information Rigidities." *Journal of Monetary Economics*, 2021(1), pp. 507–520.
- Lau, J. H.; Newman, D. and Baldwin, T. "Machine Reading Tea Leaves: Automatically Evaluating Topic Coherence and Topic Model Quality." Proceedings of the 14<sup>th</sup> Conference of the European Chapter of the Association for Computational Linguistics, 2014.
- Leduc, S.; Sill, K. and Stark, T. "Self-fulfilling Expectations and the Inflation of the 1970s: Evidence from the Livingston Survey." *Journal of Monetary Economics*, 2007, 54(2), pp. 433–459.
- Mankiw, N.G. and Reis, R. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *The Quarterly Journal of Economics*, 2002, 117(4), pp. 1295–1328.
- Mazumder, S. "The Reaction of Inflation Forecasts to News about the Fed." *Economic Modelling*, 2021, 94, pp. 256–264.
- McCracken, M.W. and Ng, S. "FRED-MD: A Monthly Database for Macroeconomic Research." *Journal of Business and Economic Statistics*, 2016, 34(4), pp. 574–589.
- Nimark, K.P. and Pitschner, S. "News Media and Delegated Information Choice." *Journal of Economic Theory*,

2019, 181, pp. 160–196.

Pfajfar, D. and Santoro, E. “News on Inflation and the Epidemiology of Inflation Expectations.” *Journal of Money, Credit and Banking*, 2013, 45(6), pp. 1045–1067.

Reid, M. “Inflation Expectations of the Inattentive General Public.” *Economic Modelling*, 2015, 46, pp. 157–166.

Tibshirani, R. “Regression Shrinkage and Selection via the LASSO.” *Journal of the Royal Statistical Society*, 1996, 58(1), pp. 267–288.

Ueda, K. “Determinants of Households’ Inflation Expectations in Japan and the United States.” *Journal of the Japanese and International Economies*, 2010, 24(4), pp. 503–518.

Umatsu, Y. and Shinya, T. “High-dimensional Macroeconomic Forecasting and Variable Selection via Penalized Regression.” *The Econometrics Journal*, 2017, 22(1), pp. 34–56.

## **Inflation Expectations and Information Rigidity: An Analysis from the Media News Perspective**

Zheng Tingguo; Fan Xinyue; Jin Wei; Fang Kuangnan

**Abstract:** Using nearly one million news articles from major national financial and economic newspapers, this paper constructs a large corpus of Chinese news and employs the latent Dirichlet allocation (LDA) model to transform news texts into structured high-dimensional news thematic attention time series to study the role of media news in driving inflation expectations. To this end, this paper uses the LASSO and LASSO-UMIDAS models to identify news topics that have a significant impact on inflation expectations, and it estimates the time-varying information rigidity based on the noise information model. By jointly modelling the information rigidity with the duration and signal-to-noise ratio of media coverage, an in-depth discussion is developed about the influence mechanism of media news. According to the results, media information will have both numerical and speed-driving effects on inflation expectations. In terms of numerical driving, some important news information, such as economic growth, price fluctuation, national news and financial credit, play a significant role in shaping inflation expectations. In terms of speed driving, the media coverage duration and signal-to-noise ratio have a significant positive impact on the information rigidity, thereby affecting the adjustment speed of inflation expectations.

**Key words:** inflation expectations, information rigidity, numerical driving, speed driving, media coverage

**JEL codes:** E17, E31

(截稿:2022年11月 责任编辑:曹永福)