

央行货币政策报告文本信息、 宏观经济与股票市场

姜富伟 胡逸驰 黄 楠

2021 金融研究

王健

2021-12-24

目录

- 简介
 - 背景和动机
 - 研究问题
 - 研究内容
 - 相关文献
 - 文章贡献
- 理论模型
- 数据和变量
- 实证结果
- 文章结论

一. 简介

背景和动机

- 中央银行沟通是指中央银行向公众披露货币政策目标、货币政策策略及未来货币政策走向等信息的过程。
- 在人民银行的政策沟通工具中，《货币政策执行报告》历史最为悠久形式最为成熟，受到金融市场的广泛关注。
- 本文利用金融情感词典和文本分析技术手段，对我国央行货币政策报告进行文本分析，并探究货币政策报告的文本信息与宏观经济和股票市场的关系。

一. 简介

研究问题：

1、报告文本信息与宏观经济和股票市场的有怎样的关系？

报告的文本信息对与宏观经济之间相关性显著，且对股票市场的收益和波动都会造成影响。

2、报告文本情绪对股票市场影响的传导路径有哪些？

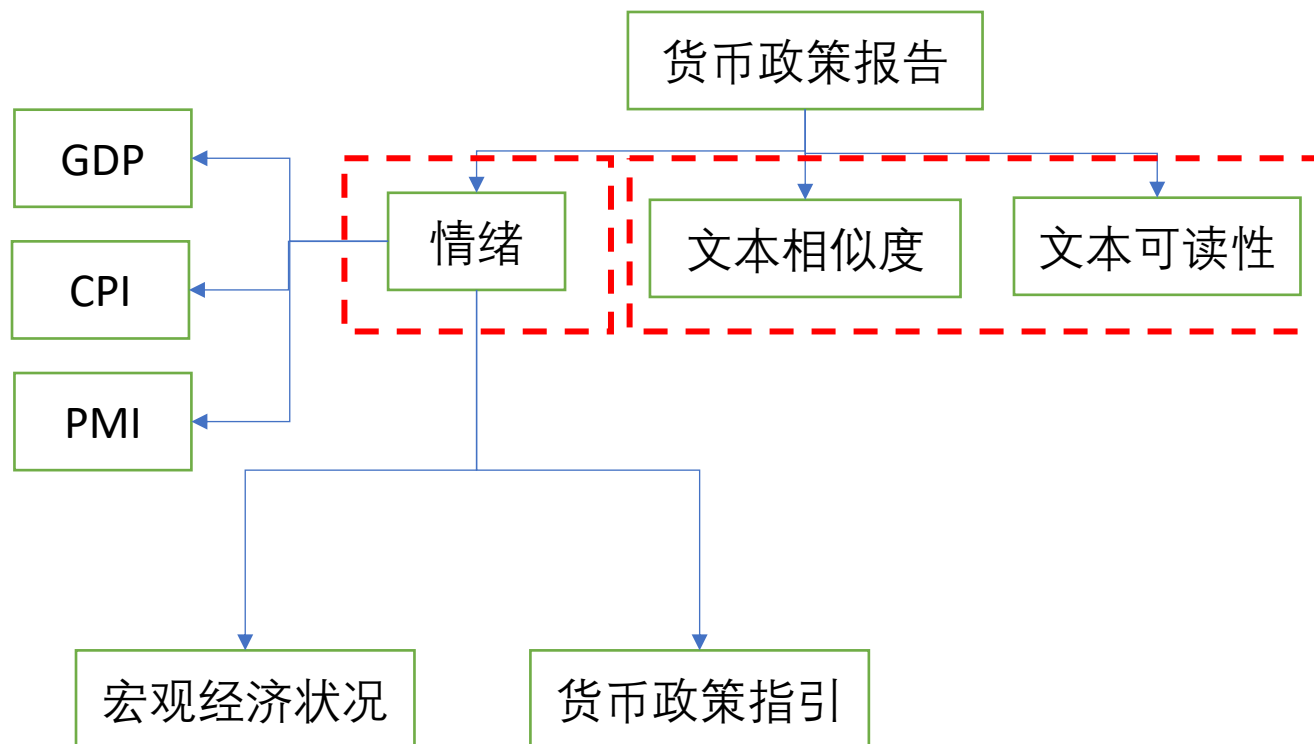
将报告的文本情绪分为反映宏观经济状况和反映央行货币政策指引的部分，而市场有显著影响的主要是反映央行货币政策指引的部分。

一. 简介

研究内容

- 提取报告的文本情绪、文本相似度、文本可读性等指标
- 研究报告文本情绪和宏观经济状况的关系
- 利用EGARCH模型，探究报告文本信息与股票市场的收益和波动的影响。
- 将报告文本情绪分解为反映宏观经济状况的部分和反映央行货币政策指引的部分，探究两者是否都能对股票市场走势有显著影响。

一. 简介



一. 简介

相关文献

- 央行沟通与股票市场反应：
 - 外国央行与外国股票市场, Lucca and Moench(2015) Cieslak et al. (2019) Guthrie and Wright(2000),
 - 中国央行与中国股票市场 冀志斌和宋清华 (2012) ; 张强和胡荣尚 (2013) ; 吴国培和潘再见 (2014)
- 经济金融文本大数据分析：
 - 第一类文献研究公司公告的文本情绪对公司股票价格的影响, Loughran and McDonald(2011) 与 Jegadeesh and Wu(2013) ;
 - 第二类文献研究媒体报道文本情绪对股票市场的影响, Tetlock(2007), Garcia(2013), Huang et al. (2014) ;
 - 第三类文献研究中央银行沟通的文本情绪与宏观经济和金融市场的关系, Amaya and Filbien(2015), Ehrmann and Talmi(2017)

一. 简介

文章贡献

- 本文使用前沿的文本分析技术对货币政策报告文本进行全方面多维度分析，在内容和技术手段上与以往文献都有明显差异。
- 本文对中文金融经济文本情绪分析中关于央行沟通情绪分析的研究形成了有益补充。
- 本文对央行沟通文本情绪影响股票市场的传导机制开展了深入分析。
- 本文的研究对于我国进行货币政策调整和投资者投资管理也有十分重要的意义。

二. 理论模型

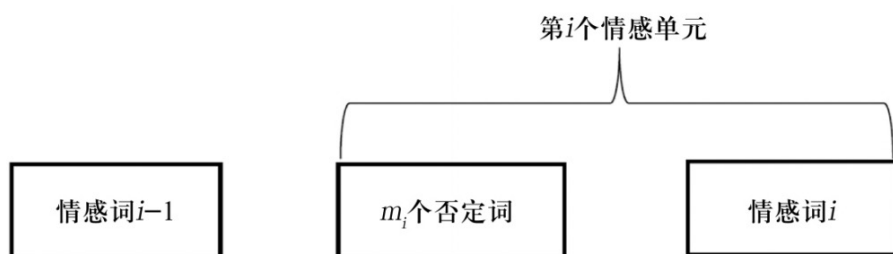
1、分词

文本分析的基础是对文本进行分词，分词的基础是分词工具自带的词库。

- 中文文本中较为常用的分词方法为 jieba 分词并加入搜狗词库，文章挑选了三个下载量较大的经济、金融、财会词库

2、文本情绪

本文先采用姜富伟等（2020）的情感词典计算货币政策报告的文本情绪，在稳健性检验中再用知网情感词典计算出的文本情绪检验我们的结果。



$$tone_i = sig(m_i) \times weight(word_i)$$

$$sig(m_i) = (-1)^{m_i}$$

$$weight(word_i) = \begin{cases} 1, & \text{若 } word_i \text{ 是积极情感词} \\ -1, & \text{若 } word_i \text{ 是消极情感词} \end{cases}$$

二. 理论模型

3、文本相似度

本文采用基于TF-IDF加权的余弦相似度法计算报告的文本相似度。

- 去除报告文本中只有语气停顿、转折等作用的词语。
- 将每个文本看成一个词汇向量，向量的某个维度表示对应词语在这个文本中出现的次数。
- 再次，计算所有词语的idf系数。
- 最后，在每个文本向量中表示词语j频率的维度上乘以词语j的idf系数

$$idf_j = \ln\left(\frac{N}{n_j}\right)$$

4、文本可读性

具体而言，我们将句号、感叹号和问号看作一个句子的结尾对文本句子总数进行统计，进而以文本总字数除以文本句子总数，得到平均句子长度。

三. 数据和变量

1、文本指标

- 本文计算文本情绪的对象是人民银行每季度公布一次的《货币政策执行报告》，第一期报告于2001年5月公布于人民银行官方网站上，此后每隔3个月发布新的报告。
- 本文使用的报告从2001年第一季度起，至2018年第三季度止，共71篇文本。
- 在对报告进行文本分析前，我们对报告进行了预处理，删除了所有标点符号，去掉了所有标题和非正文的文本内容。

2、股票收益

- 本文选取的股票指数包括上证综指、深证成指、沪深 300 和中证 500，涵盖大市、大盘股和中小盘股等不同股票类型的收益状况，涵盖的时间范围为2005年10月1日至2018年12月31日。

二. 数据和变量

3、宏观经济与金融数据

宏观经济金融数据	起止时间	频率	数据来源	变频方式
制造业采购经理指数	2005—2018 年	月度	国家统计局	取季度均值
消费者价格指数	2001—2018 年	月度	国家统计局	取季度同比增长率
城镇登记失业率	2002—2018 年	季度	国家统计局	—
1 年期国债到期收益率	2002—2018 年	日度	中债估值中心	取季度均值
美元兑人民币中间价	2001—2018 年	日度	中国人民银行	取季度均值
Wind 全 A 指数静态市盈率	2002—2018 年	日度	Wind 资讯	取季度均值

三. 实证结果

1、描述性统计

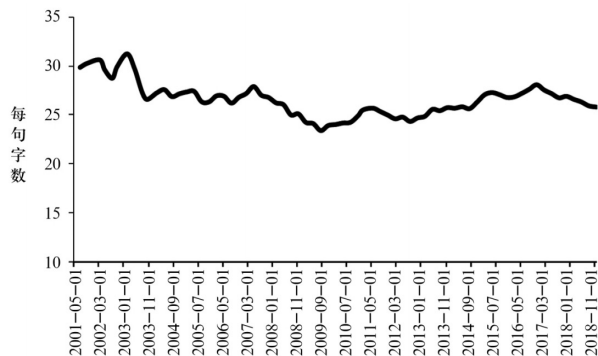
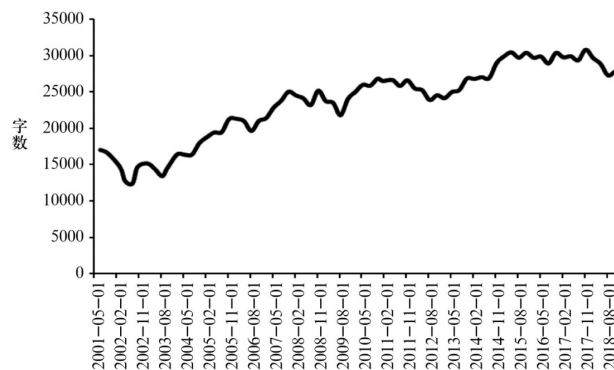


图4 报告文本可读性

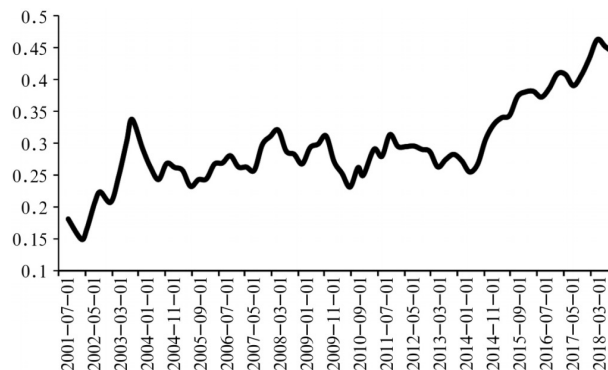
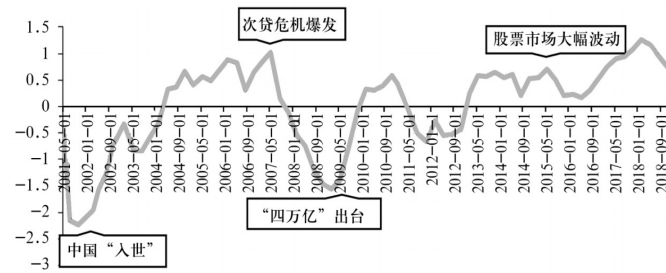


图3 报告文本相似度



三. 实证结果

2、文本情绪与宏观经济金融指标相关性分析

首先就相关性分析所涉及的变量展开ADF检验。除GDP增长率、净出口指标、城镇登记失业率和美元兑人民币中间价一阶差分后平稳外，其他指标均通过检验。

	tone(-1)	D_GDP	D_NX	CPI growth	bond yield	PMI	P/E ratio	D_unemploy	D_forex
tone(-1)	1								
D_GDP	0.42 *** (3.276)	1							
D_NX	-0.05 (-0.373)	-0.14 (-1.043)	1						
CPI growth	0.30 ** (2.125)	-0.29 ** (-2.179)	0.07 (0.524)	1					
bond yield	0.02 (0.150)	-0.30 ** (-2.263)	0.10 (0.736)	0.38 *** (3.035)	1				
PMI	0.29 ** (2.119)	0.46 *** (3.719)	-0.08 (-0.620)	0.39 *** (3.116)	-0.17 (-1.251)	1			
P/E ratio	-0.07 (-0.520)	0.08 (0.571)	-0.17 (-1.231)	0.53 *** (4.583)	-0.10 (-0.761)	0.60 *** (5.525)	1		
D_unemploy- ment	-0.30 ** (-2.225)	-0.14 (-0.995)	0.02 (0.139)	-0.19 (-1.399)	-0.12 (-0.889)	-0.50 *** (-4.233)	-0.23 * (-1.692)	1	
D_forex	0.14 (1.004)	0.14 (1.050)	0.07 (0.544)	-0.46 *** (-3.754)	-0.22 (-1.638)	-0.31 *** (-2.404)	-0.36 *** (-2.798)	0.02 (0.144)	1

三. 实证结果

3、文本情绪与股票市场

- 参考Ehrmann and Talmi (2017) 的模型设定, 我们将TF-IDF加权的文本相似度指标加入 EGARCH 模型的方差方程, 检验货币政策报告的文本相似度是否能够影响股票市场的波动性

$$\begin{aligned} r_t &= \alpha + \beta r_{t-1} + \gamma S_gdp_t + \delta S_policy_t + \theta tone_t \\ \ln(\sigma_t^2) &= a + b \left(\left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - \sqrt{\frac{2}{\pi}} \right) + c \left(\frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right) + d \ln(\sigma_{t-1}^2) \\ &\quad + e D_gdp_t + f D_policy_t + g D_report_t + h Read_t + i Sim_t \end{aligned}$$

- 未预期到的 GDP 增长: 季度GDP-朗润预测数据
- 未预期货币政策冲击:

$$S_policy_t = interbank_rate_{t+1} - interbank_rate_{t-1}$$

变量	上证综指	深证成指	沪深 300	中证 500
均值方程:				
constant	0.0333* (1.688)	0.0224 (0.824)	0.0463** (2.130)	0.0293 (0.971)
r_{t-1}	0.00906 (0.501)	0.0277 (1.501)	0.0223 (1.199)	0.0527*** (2.960)
S_gdp_t	0.347** (2.540)	0.497** (2.576)	0.426*** (2.837)	0.258 (1.394)
S_policy_t	-0.413* (-1.832)	-0.625** (-2.133)	-0.534** (-2.072)	-0.652** (-2.268)
$tone_t$	0.541*** (3.648)	0.726*** (3.126)	0.744*** (4.186)	0.683*** (2.932)
方差方程:				
D_gdp_t	0.0217 (0.304)	0.0980 (1.403)	0.0377 (0.515)	0.0674 (0.918)
D_policy_t	0.0335 (1.013)	0.0770** (2.114)	0.0160 (0.472)	0.149*** (3.862)
变量	上证综指	深证成指	沪深 300	中证 500
D_report_t	0.282 (1.140)	0.450 (1.584)	0.567* (1.918)	0.277 (1.047)
$Read_t$	0.00163 (0.092)	0.00648 (0.332)	0.0128 (0.678)	0.0294 (0.786)
Sim_t	-1.76* (-1.849)	-1.90** (-2.148)	-1.72* (-1.764)	-1.34 (-1.401)

三. 实证结果

3、机制分析

- 通常认为，央行货币政策报告既反映客观经济金融基本面，也反映央行的政策主张和对宏观经济金融的观点。
- 我们将报告分为八个部分：货币信贷概况、金融市场分析、国际经济回顾、国内经济回顾、国际经济展望、国内经济展望、货币政策操作和货币政策思路。
- 前四部分为已有的经济金融信息，可以归类为反映基本面的部分；后四部分为央行对未来国内外局势的预估和货币政策相关分析，可以归类为反映央行指引的部分。

变量	上证综指			深证成指		
	(1)	(2)	(3)	(4)	(5)	(6)
均值方程:						
constant	0.034 [*] (1.745)	0.034 [*] (1.723)	0.035 [*] (1.763)	0.0235 (0.865)	0.0250 (0.923)	0.0249 (0.918)
r_{t-1}	0.00836 (0.461)	0.00819 (0.451)	0.00859 (0.473)	0.0265 (1.431)	0.0272 (1.475)	0.0271 (1.465)
S_gdp_t	0.346 ^{**} (2.514)	0.346 ^{**} (2.525)	0.347 ^{**} (2.539)	0.494 ^{**} (2.548)	0.496 ^{***} (2.572)	0.497 ^{***} (2.574)
S_policy_t	-0.417 [*] (-1.841)	-0.416 [*] (-1.840)	-0.416 [*] (-1.844)	-0.624 ^{**} (-2.114)	-0.625 ^{**} (-2.115)	-0.625 ^{**} (-2.118)
$real_t$	0.143 (0.744)		0.250 (1.367)	0.192 (0.954)		0.102 (0.510)
$guid_t$		0.291 ^{**} (1.988)	0.324 ^{**} (2.131)		0.421 ^{**} (2.063)	0.423 ^{**} (2.039)
方差方程:						
D_gdp_t	0.0153 (0.214)	0.0169 (0.237)	0.0172 (0.241)	0.106 (1.512)	0.103 (1.475)	0.103 (1.475)
D_policy_t	0.0343 (1.033)	0.0333 (1.006)	0.0345 (1.039)	0.0756 ^{**} (2.077)	0.0762 ^{**} (2.086)	0.0761 ^{**} (2.084)
D_report_t	0.312 (1.255)	0.321 (1.300)	0.291 (1.169)	0.460 (1.596)	0.478 [*] (1.661)	0.469 (1.631)
$Read_t$	0.000415 (0.023)	0.00217 (0.122)	0.00168 (0.094)	-0.00628 (-0.316)	-0.00939 (-0.475)	-0.00815 (-0.412)
Sim_t	-1.81 ^{**} (-1.973)	-1.71 ^{**} (-2.001)	-1.75 ^{**} (-2.017)	-2.04 ^{**} (-2.314)	-2.05 ^{**} (-2.437)	-2.03 ^{**} (-2.311)
R squared	0.005	0.006	0.007	0.007	0.009	0.009
N	3220	3220	3220	3220	3220	3220

四、文章结论和启示

- 货币政策报告的文本情绪与诸多宏观经济指标有着显著的相关性。
- 在控制经济增长和货币政策后，报告文本情绪对报告发布后股票市场的收益率有显著的正向影响。报告文本相似度越高，股票市场波动性越小。报告文本可读性高低对股票市场波动性的影响并不显著。
- 进一步研究发现，报告文本情绪中对股票市场有显著影响的是其反映央行政策指引的部分，而反映宏观经济基本面状况的部分对股票市场影响并不显著。
- 一方面，本文的实证结果表明央行沟通可以显著影响股票市场的收益率和波动性，从而证明我国央行沟通具备有效性。
- 另一方面，本文进一步指出，央行报告的文本情绪中影响市场的部分是反映央行对未来经济形势的判断和政策前瞻的部分，从数据层面证实了央行指引对股票市场的显著影响，表明央行有效预期管理的重要性。

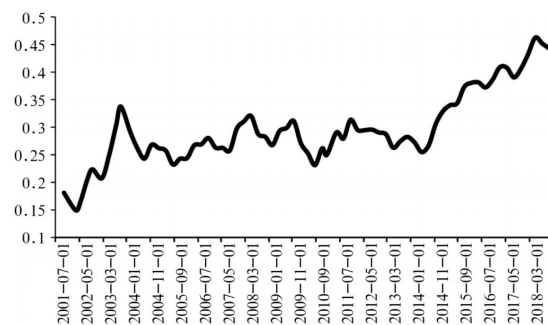
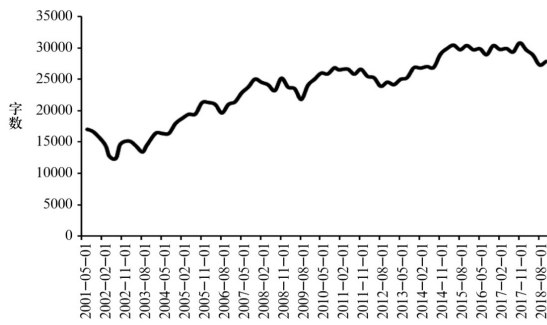
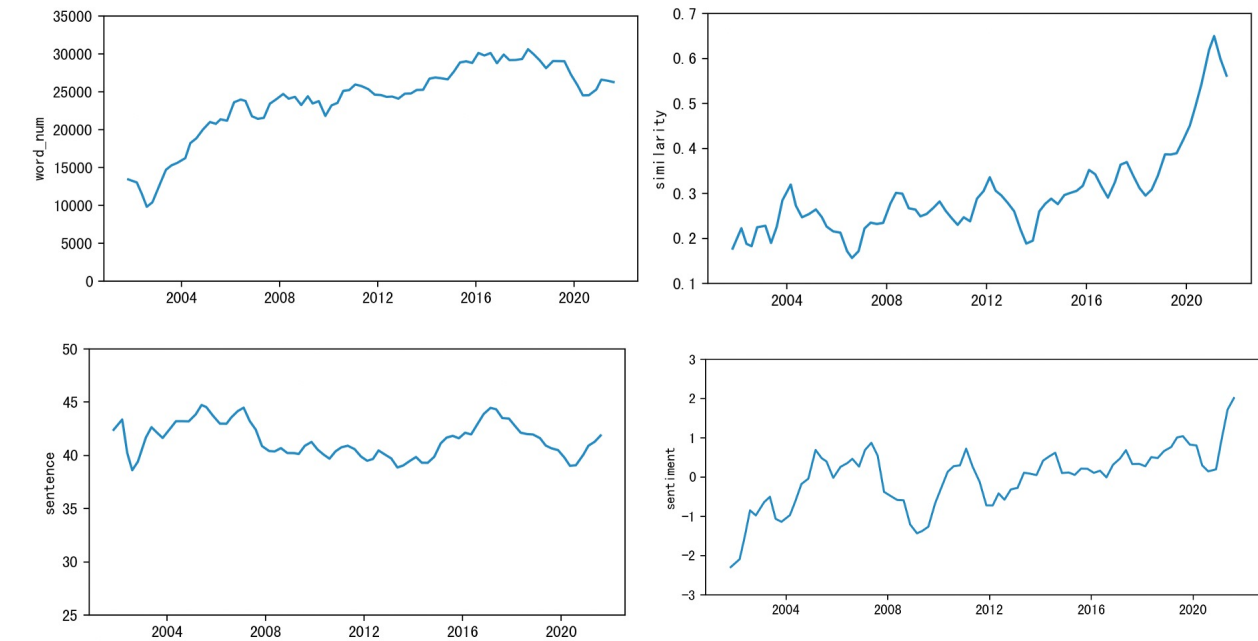


图3 报告文本相似度

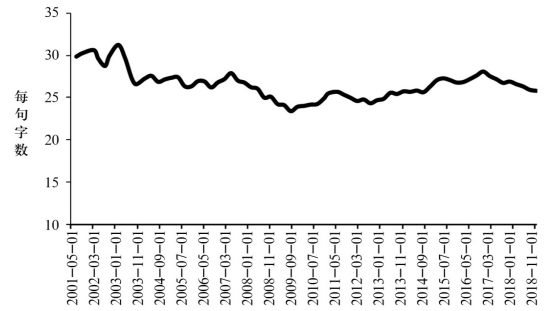


图4 报告文本可读性

