

投资者高频情绪对股票日内 收益率的预测作用

尹海员、吴兴颖
2019 中国工业经济

王健
2020-06-17

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
- 数据和变量
- 实证结果
- 文章结论

一. 简介

背景和动机

- 近年来关于投资者情绪与股票市场互动关系的研究，重点不再仅局限于判断情绪是否能预测收益水平，而是如何衡量投资者情绪并量化其更广泛的影响效应。
- 传统的统计指标有其局限性，依托海量在线信息来构建的情绪指标具有显著优势。
- 在股票市场收益的预测方面，学者们已经基于月度、季度等低频数据展开了大量的研究，数据频率越低，则过滤掉的市场信息越多，最终导致所构建的度量指标的有效性减弱。

一. 简介

研究问题：

1、高频投资者情绪是否对股票盘中收益率有预测作用？

区分以下样本环境：

- 全样本
- 牛、熊市和极端市场
- 隔夜和半日预测

2、如何解释高频情绪指标的预测作用？

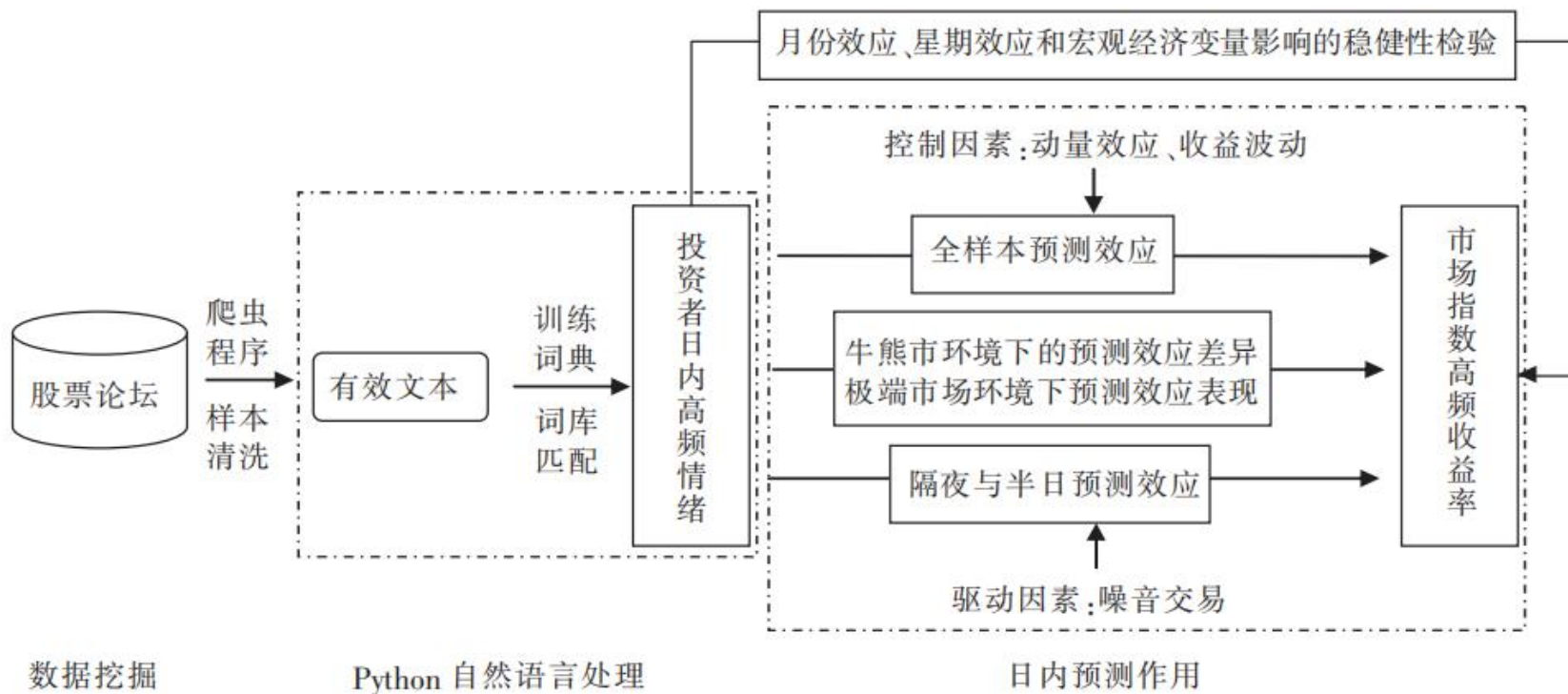
噪音交易者

一. 简介

研究内容

- 利用爬虫和文本分析工具，爬取论坛文章并构造高频情绪指标。
- 探究高频投资者情绪对日内盘中收益率的预测作用，以及在不同市场环境下，预测结果是否具有显著性。
- 探究隔夜投资者情绪的预测作用和投资者情绪的“半日效应”。并探究解释高频投资者情绪预测作用的可能解释原因。
- 稳健性检验

一. 简介



一. 简介

相关文献（投资者情绪度量）

- 直接指标：王美今和孙建军（2004）利用央视看盘指数构建了投资者情绪指数。间接指标：Derrien（2005）利用IPO发行量及首日收益衡量投资者情绪。复合指标：Baker and Wurgler（2006）将多个单一指标进行主成分分析，在控制宏观经济变量等理性因素后得到度量投资者情绪的复合指数。
- 林振兴（2011）、孟雪井等（2016）、段江娇等（2017）、石善冲等（2018）分别从东方财富股吧、百度搜索指数、上证A股公司股吧和股市类微信公众号提取投资者情绪指标。

一. 简介

相关文献（股票收益率的日内效应）

- Gao et al. (2018) 研究了标准普尔 500 指数的日内动量效应现象，发现市场首个半小时的回报可以预测市场上最后一个半小时的回报。
- Zhang et al. (2018) 利用半小时交易数据研究了中国股市的日内动量效应，发现第一个和第七个半小时的收益可以显著预测最后半小时的收益，并且早上的市场收益也能显著预测下午的市场收益。

一. 简介

文章贡献

- 直接对获取到的文本信息进行情感分析，构建了能更准确反映投资者情绪的日内高频指标。
- 从多个维度验证了高频投资者情绪对中国股票日内收益率的可预测性，并针对中国股市的特殊交易机制，分析了隔夜信息以及午休期间的投资者情绪是否会对后期股价产生冲击。
- 分别从交易量和市场长期收益反转的角度探讨了噪音交易者行为对情绪预测作用的影响。

二. 数据和变量

1、数据来源

- 情绪指标来源：“金融界”股吧论坛帖。
- 选择原因：“金融界”网站主要优点在于股票论坛中的资讯丰富，发帖频率高，同时股吧论坛的专题分类页齐全，保留的发帖相较其他论坛有更长的历史追溯性，为日内数据的挖掘提供了便利。
- 收益指标：上证指数。
- 收益指标来源：Wind
- 时间：2013-07-01 到 2018-08-14


全部帖 | 精华帖 | 准确帖 | 附件帖 | 荐股帖

<< 上一页 | 1 | 2 | 3 | 4 | 5 | 下一页 >> | 发新帖

主题	作者	点击数	回复数	按发表时间
 [原创]价值和成长是投资的两面【江南燕园】	天下第地	2996	4	06-17 11:13
 提醒：野火压不住，春风吹又生！【江南燕园】	王府	2225	9	06-17 10:03
 用力的活着（5）：静与动！【江南燕园】	孙武强	5410	4	06-16 18:55
 不可能依靠高位庄股裹腹御寒【江南燕园】	蓝海孤舟	6071	4	06-16 15:53
 [原创]淘气天尊：市场3连阴后要不…【江南燕园】	淘气天尊	5131	3	06-16 15:46
 [原创]自说自话【江南燕园】	长城脚下	5541	3	06-16 13:01
 行情好看难做【江南燕园】	蓝海孤舟	5815	2	06-16 11:32
 仓又嘉措【江南燕园】	蓝海孤舟	6684	2	06-16 06:58
 [原创]股市投资是否也需要一个“洪…【江南燕园】	证星笑寒	6841	6	06-15 21:21
 用利的活着（4）：回压的底限【江南燕园】	孙武强	8942	4	06-15 18:36
 [原创]周末低吸【江南燕园】	理科魔法师	8097	6	06-15 16:30
 [原创]淘气天尊：市场新低到3547点…【江南燕园】	淘气天尊	8078	4	06-15 15:23
 是什么让我们胆寒？【江南燕园】	蓝海孤舟	9072	4	06-15 12:09
 [原创]淘气天尊：市场节后如期跳水…【江南燕园】	淘气天尊	7761	4	06-15 12:03
 我的心愿【江南燕园】	蓝海孤舟	11964	6	06-14 16:50
 [原创]你要错过这样的暴增机遇吗？【江南燕园】	名花流的剑	19296	16	06-13 07:36
 [原创]跌无需担心，涨必须防守！！…【江南燕园】	孤独de血狼	16358	9	06-12 11:21
 [原创]牛市行情进行中【南平1号】	dong12301230	15566	2	06-12 10:42

 『普通』主题：用力的活着（5）：静与动！

阅读 5412 | 回复 4 | 举报 | 收藏本页



孙武强

查看TA的论坛文集 | 查看TA的荐股记录

头衔：股市猛将 | 积分：150519 | 关注 | 加为好友

发表日期：2021-06-16 18:55:47

引用 | 评分 | 只看该作者

 嗨爆新春！开户得大奖，人人有份儿

 炒股训练营：如何短线炒强势股？

今天上证指数收盘3518点，向上的趋势延续，短线注意3440点得失。

盘面观察，由于指数跌破强势底限，应首先回补5月25日缺口3498点。

对于市场而言，慢牛前行指数的宽幅震荡是常态。

那么市场到位我们买什么样的品种，我日志中提到的如安宁股份这样的品种仍可继续操作，方法就是在趋势的拐点重仓买入持有，余下资金盘中跟随趋势做波段，不断降低持仓成本，以应对市场的宽幅震荡。

操作，持股观望，（资产缩水5%，可我的持仓成本再次降低）

关注：60329X(等三浪筑底结束)

一切都是个人所想，可读，勿信！

“金融界” 股吧论坛 帖

二. 数据和变量

2、情绪指数构建

- 原始数据抓取：
 - 在Java环境的支持下调用R中的Rselenium和Rwebdriver等程序包实现动态网页爬虫，并在详情页中抓取每条发帖的标题、作者、阅读量、详细时间等字段。
 - 对原始文本内容进行数据清洗，去除无效符号和冗余字段、删除重复发帖、剔除空缺行等，导入jieba，对内容进行分词，统计各单词出现的次数，剔除停用词语得到排名前1000的词。
 - 人工筛选出与股市无关的关键词，以此为依据查找并剔除无效发帖80281条，得到有效帖406134条，论坛的日均发贴数量大约为160条。
 - 以半小时为间隔，按照不同时间区间对所有帖子分类汇总，将发帖日期与股市行情数据的日期进行匹配，得到交易时间段内的发帖120665条，非交易时间段的发帖285469条。

二. 数据和变量

2、情绪指数构建

- 投资者情绪的量化
 - 词典语料：SnowNLP中自带的pos(neg)语料库、中国知网(HowNet)情感词典和台湾大学简体中文情感极性词典(NTSUSD)作为基础情感词典语料库。BosonNLP情感词典作为领域情感词典语料库
 - 调用 SnowNLP 中的 sentiment 模块并编写循环语句，基于词库匹配法对所有样本数据逐条进行量化处理，返回情感为“正面”的概率，取值区间为[0,1]。针对中性类别（0.5），导入NLTK 资料库，对这些内容进行分词后重新抓取关键词，并重新进行文本化分。
 - 将所有交易日的帖子按照半小时划分，分别在每个半小时中抽取 50 条，人工对这些帖子进行情感类别划分并标注（83.2%）。

二. 数据和变量

2、情绪指数构建

- 变量指标设计

(1) 投资者高频情绪。本文以半小时为间隔将中国股市每日的交易时间分为 8 个 30 分钟时间段，分别汇总各半小时内的乐观/悲观帖子数，以此构建高频投资者情绪指数。

$$sentiment_{i,t} = \ln \left[(1 + pos_{i,t}) / (1 + neg_{i,t}) \right]$$

(2) 股票收益指标。为了情绪指标的时间相匹配，将半小时收盘价取对数差分后乘以 100，得到上证指数收益率。

$$R_{i,t} = 100 \times [\ln(P_{i,t}) - \ln(P_{i,t-1})]$$

在计算各交易日第一个半小时的收益率时，使用前一个交易日的收盘价作为起始价格。第五个半小时的回报时间段为[11:30-13:30]。

三. 实证结果

1、描述性统计

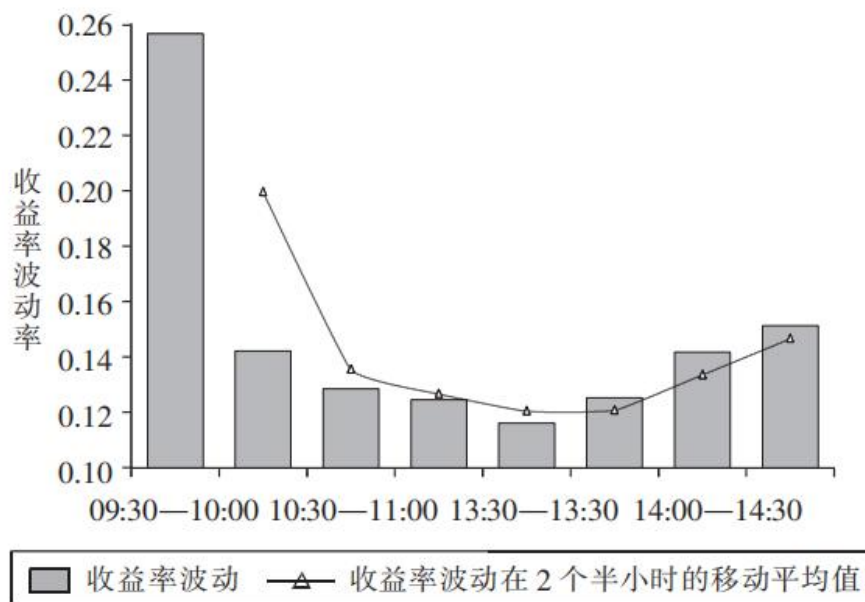


图 2 市场(上证指数)收益率日内波动率
市场收益率通常在交易日开盘后第 1 个半小时波动的最为剧烈, 随后逐渐降低, 在时间分布上呈现较为明显的“U”型特征。

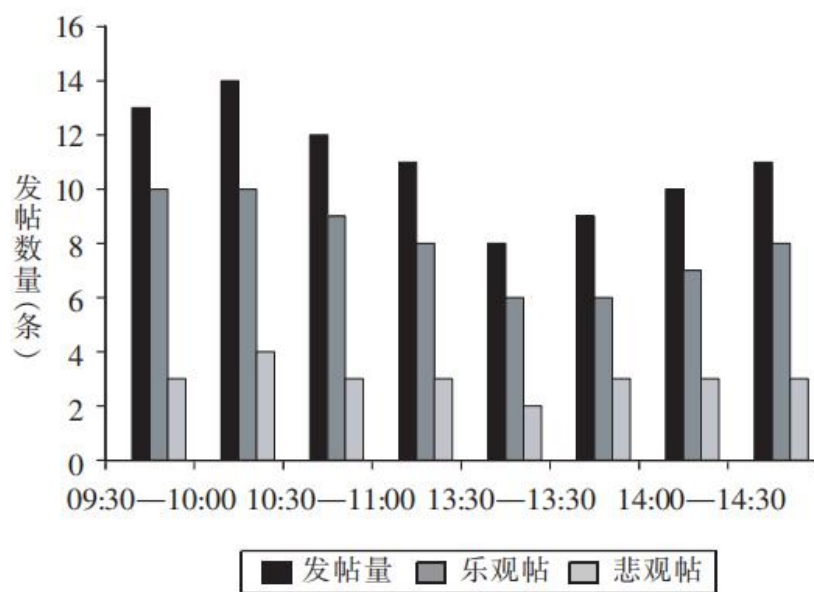


图 3 论坛日各时间段发帖数量均值
上午发帖量总体上高于下午发帖量, 且开盘后的一小时是投资者发帖的高峰期, 随后逐渐减少, 但会在下午开盘后恢复上涨趋势、

三. 实证结果

2、日内投资者情绪对市场收益率的预测作用

- 由于本文主要关注情绪的日内效应，所以预测回归从各交易日的第二个半小时开始。此外对每个回归变量进行了NW调整。

$$R_{i,t} = \beta_0 + \beta_1 sentiment_{i,t-1} + \delta_{i,t} \quad t = 2, 3 \dots 8$$

表 1

投资者情绪对日内收益的预测作用

	第二个 半小时	第三个 半小时	第四个 半小时	第五个 半小时	第六个 半小时	第七个 半小时	第八个 半小时
滞后半小时投资者情绪	0.0266* (1.6726)	0.0228 (1.2790)	0.0042 (0.2097)	0.0127** (2.2363)	0.0250*** (2.5956)	0.0272*** (2.6208)	0.0497** (2.2840)
常数项	0.0366** (2.2346)	-0.0235 (-1.3017)	0.0046 (0.2182)	0.0361** (1.9720)	0.0250 (1.5258)	-0.0284 (-1.5306)	0.0076 (0.3334)
N	1171	1173	1175	1169	1121	1131	1144
Adj-R ²	0.0615	0.1050	0.0617	0.1004	0.0516	0.0706	0.0937
F	2.7977	1.6359	0.0440	3.5134	2.7743	3.7318	5.2167

三. 实证结果

- 考虑到收益率自身的盘中动量效应，本文参照Sun et al.(2016)的做法，进一步在预测模型中控制收益率滞后项。同时，考虑到市场波动对收益率的影响，加入市场收益在滞后半小时中的波动水平。

$$R_{i,t} = \beta_0 + \beta_1 sentiment_{i,t-1} + \beta_2 R_{i,1} + \beta_3 R_{i,t-1} + \beta_4 \sigma_{i,t-1} + \delta_{i,t}$$

	第二个 半小时	第三个 半小时	第四个 半小时	第五个 半小时	第六个 半小时	第七个 半小时	第八个 半小时
滞后半小时投资者情绪	0.0249* (1.6845)	0.0244 (1.5037)	0.0050 (0.2091)	0.0097 (0.5734)	0.0200** (2.2654)	0.0219** (2.0844)	0.0397* (1.6737)
第一个半小时收益率	0.0346** (2.3454)	0.0241 (0.6079)	-0.0228 (-1.2601)	0.0399** (2.4647)	0.0421*** (2.7539)	0.0539*** (3.3567)	0.0651*** (3.3735)
滞后半小时收益率		0.0969*** (3.2423)	-0.0301 (-0.8579)	0.1292*** (5.1058)	0.0270 (0.9468)	0.1201*** (3.7212)	0.2269*** (6.7125)
滞后半小时波动率	0.1087** (2.3479)	0.0207 (0.2253)	0.0135 (0.1019)	0.1724 (0.6087)	0.3492*** (3.9907)	-0.0782 (-0.3490)	-0.2546** (-2.2061)
常数项	0.0098 (0.1275)	-0.0243 (-1.2321)	0.0004 (0.0147)	0.0137 (0.6113)	-0.0167 (-0.8478)	-0.0136 (-0.5348)	0.0588** (2.1371)

三. 实证结果

3、日内情绪的预测作用在不同市场环境下的表现

- 借鉴Pagan and Sossounov(2003)的方法，将市场环境划分为牛熊市两种状态：
 - 1) 对市场指数设定一个单向运行周期，将样本区间细分为多个子区间；
 - 2) 将各个运行周期中第*i*日的市场指数与相邻交易日进行对比，运行周期中会交替出现极高点 and 极低点，选取这些连续极值点中的最高者（或连续波谷中的最低者）作为该周期内牛熊市状态转换的转折点。
- 本文将上证指数的单向运行周期设定为5个月，可以将样本区间大致划分为牛市和熊市各五个阶段
- 牛市环境：投资者情绪主要在第二、三、六、七、八这五个半小时的收益率先产生显著的正向影响，显著的时间区间甚至多于全样本；
- 熊市环境：情绪的预测作用远不及牛市，但投资者情绪至少在收盘前的一小时内始终保持显著；滞后收益率等控制变量在熊市中的显著性总体上高于牛市

表 3 牛市环境下投资者情绪、滞后收益和收益波动对日内收益的预测作用

	第二个 半小时	第三个 半小时	第四个 半小时	第五个 半小时	第六个 半小时	第七个 半小时	第八个 半小时
滞后半小时投资者情绪	0.0388** (2.1881)	0.0537*** (2.6962)	0.0048 (0.1391)	0.0005 (0.0305)	0.0176** (2.4491)	0.0267*** (2.6382)	0.0287** (2.3198)
第一个半小时收益率	0.0003 (0.0103)	0.0738** (2.4122)	-0.0488 (-1.0296)	0.0338* (1.6756)	0.0227 (0.8159)	0.0234 (1.0486)	0.0615** (2.2695)
滞后半小时收益率		-0.1625*** (-3.1001)	0.0191 (0.2944)	-0.0384 (-0.5093)	0.0177 (0.2239)	0.2058*** (5.3281)	0.1951*** (4.3446)
滞后半小时波动率	0.1448 (0.9262)	0.1910 (0.9053)	-0.2030 (-0.8135)	-0.1721 (-0.5086)	0.3180** (2.5334)	0.1195 (0.7524)	-0.1085 (-0.6810)

表 4 熊市环境下投资者情绪、滞后收益和收益波动对日内收益的预测作用

	第二个 半小时	第三个 半小时	第四个 半小时	第五个 半小时	第六个 半小时	第七个 半小时	第八个 半小时
滞后半小时投资者情绪	0.0265 (0.9829)	0.0231 (0.7947)	0.0078 (0.2261)	0.0130 (0.4095)	0.0263 (1.0611)	0.0336** (2.4325)	0.0570* (1.7679)
第一个半小时收益率	0.0645*** (2.8374)	0.0299 (0.4649)	0.0045 (0.1702)	0.0424* (1.8048)	0.0562** (2.5190)	0.0732* (1.8468)	0.0586** (2.0265)
滞后半小时收益率		0.0065*** (2.7780)	0.0694 (1.3617)	0.3563*** (8.1526)	0.0306*** (2.7199)	0.0107 (0.1106)	0.2606*** (4.9668)
滞后半小时波动率	0.1230** (2.0355)	-0.2256 (-1.2861)	0.3374 (1.5811)	0.4849*** (2.9429)	0.3801*** (2.9401)	-0.1889 (-0.6150)	0.3975** (2.2948)

三. 实证结果

4、日内情绪的预测效应在极端市场环境中的表现

- 我们以2015年股灾作为一个较号的极端市场案例，化分时间区间为暴涨期间（2015-05-07至2015-06-12）；暴跌区间（2015-05-07至2015-07-08）

附表 4： 2015 年股灾之前上涨阶段投资者情绪对日内收益的预测作用 ↵

↵	第二半小 时 ↵	第三半小 时 ↵	第四半小 时 ↵	第五半小 时 ↵	第六半小 时 ↵	第七半小 时 ↵	第八半小 时 ↵
滞后半小时 投资者	0.1733* ↵ (1.7352) ↵	0.0924* ↵ (1.7499) ↵	0.2817* ↵ (1.7520) ↵	0.2151 ↵ (1.1757) ↵	0.0433* ↵ (1.7316) ↵	0.1389* ↵ (1.7483) ↵	0.5250** ↵ (2.0804) ↵

附表 5： 2015 年股灾期间下跌阶段投资者情绪对日内收益的预测作用 ↵

↵	第二半小 时 ↵	第三半小 时 ↵	第四半小 时 ↵	第五半小 时 ↵	第六半小 时 ↵	第七半小 时 ↵	第八半小 时 ↵
滞后半小时 投资者情绪 ↵	0.0690 ↵ (0.1834) ↵	0.5957 ↵ (1.4718) ↵	0.0389* ↵ (1.7239) ↵	0.0141 ↵ (0.0213) ↵	0.7605** ↵ (2.4187) ↵	0.4133 ↵ (0.8257) ↵	0.3897** ↵ (2.1850) ↵

三. 实证结果

5、隔夜投资者情绪的预测作用

$$R_{i,t} = \beta_0 + \beta_1 sentiment_{i,h0} + \beta_2 R_{i,1} + \beta_3 R_{i,t-1} + \beta_4 \sigma_{i,t-1} + \delta_{i,t}$$

	第一个半小时 *		第二个半小时		第三个半小时		第四个半小时	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
隔夜投资者情绪	0.0543 (0.9469)		0.0271* (1.7347)	0.0257 (0.8976)	0.0268*** (2.5921)	0.0311** (2.1409)	0.0656* (1.8186)	0.0662* (1.8337)
第一个半小时收益率				0.0377** (2.3532)		0.0200 (1.2689)		-0.0263 (-1.3893)
滞后半小时收益率						0.1193*** (-3.8500)		-0.0262 (-0.7365)
滞后半小时波动率				0.1072** (2.3712)		0.0042 (0.0405)		0.0320 (0.2402)

这一结果说明隔夜情绪的释放具有时滞性，它并非在开盘后短时间内就能冲击市场，而是集中作用于第三个半小时的收益率，并在后续时间里持续减弱。

三. 实证结果

6、投资者情绪的“半日效应”

$$R_{i,aft} = \beta_0 + \beta_1 sentiment_{i,mor} + \beta_2 sentiment_{i,noon} + \beta_3 R_{i,mor} + \delta_i$$

	模型(7-1)	模型(7-2)	模型(7-3)
上午投资者情绪	0.0323* (1.8314)	0.0268 (0.4435)	0.0154 (0.2979)
上午股票收益率		0.1071*** (3.0769)	0.1049*** (3.0427)
午休投资者情绪			0.0388** (2.3245)

说明在中国股市收益率的预测中，确实存在“半日效应”，这种效应不仅取决于上午的总体收益率，还来源于上午和午休期间的投资者情绪，相比之下，投资者在午休期间的情绪状态发挥着更明显的作用

三. 实证结果

7、噪音交易对日内情绪预测作用的影响（1）

- 前述实证证明了高频投资者情绪对上证指数的日内收益有显著的预测作用。一个可能的解释是，这种预测效应主要由噪音交易者的交易行为驱动。
- Barber and Odean（2008）指出，噪音交易活动的主要表现是市场上的成交量显著提高，且成交量可作为噪音交易的代理变量。
- 本文认为，若噪音交易假设成立，则投资者情绪的预测作用理应会在成交量高的时间区间内更显著。

$$R_{i,t} = \beta_0 + \beta_1 sentiment_{i,t-1} + \beta_2 HighVol_{i,t} + \beta_3 sentiment_{i,t-1} HighVol_{i,t} + \beta_4 R_{i,1} + \beta_5 R_{i,t-1} + \beta_6 \sigma_{i,t-1} + \delta_{i,t}$$

- 其中， $HighVol_{i,t}$ 为交易量在*i*日第*t*个半小时的虚拟变量，如果第*t*个半小时的交易量高于所有样本在该时段的均值，则取值为1，否则取值为0。

表 7

交易量和投资者情绪对日内收益的预测作用

	第二个 半小时	第三个 半小时	第四个 半小时	第五个 半小时	第六个 半小时	第七个 半小时	第八个 半小时
滞后半小时投资者情绪	0.0009* (1.8388)	0.0087 (0.5189)	0.0211 (1.4008)	0.0081 (0.5463)	0.0115** (2.5324)	0.0157** (2.2512)	0.0274* (1.8756)
第一个半小时收益率	0.0346 (1.1279)	0.0246 (0.5849)	-0.0196 (-0.5160)	0.0378 (1.3956)	0.0424 (1.3437)	0.0543** (2.0813)	0.0660** (2.0467)
滞后半小时收益率		0.0974 (1.4997)	-0.0339 (-0.9685)	0.1278 (1.6332)	0.0259 (0.4565)	0.1190** (2.4141)	0.2257*** (3.5740)
滞后半小时波动率	0.1109 (1.1071)	0.0209 (0.1075)	0.0678 (0.2737)	0.1518 (0.5271)	0.3437 (1.0573)	-0.0777 (-0.3852)	-0.2418 (-0.9187)
交易量	0.0211 (0.6325)	-0.0300 (-0.8672)	-0.1031* (-1.8103)	0.0265 (0.6585)	0.0220 (0.6219)	-0.0231 (-0.6326)	-0.0562 (-1.1296)
交易量×滞后半小时投资者情绪	0.0473* (1.6667)	0.0357 (1.1508)	0.0613 (1.1422)	0.0027 (0.0694)	0.0212** (2.2928)	0.0298*** (2.8954)	0.0269** (2.5123)
常数项	-0.0022 (-0.0829)	-0.0114 (-0.3487)	0.0387 (1.1280)	0.0036 (0.1111)	-0.0266 (-0.0840)	-0.0026 (-0.1074)	0.0842** (2.0056)

三. 实证结果

7、噪音交易对日内情绪预测作用的影响（2）

- 从另一个角度看，如果投资者日内的情绪效应是由噪声交易驱动的话，那么这种预测作用会在长期内发生反转，因为噪音交易活动只会在短期内推动股票价格偏离基本价值，而股价在更长的时间内会发生反转。

$$R_{i,t+h} = \beta_0 + \beta_1 sentiment_{i,t-1} + \beta_2 R_{i,1} + \beta_3 R_{i,t-1} + \beta_4 \sigma_{i,t-1} + \delta_{i,t+h}$$

- 其中，h表示不同的时间区间， $R_{i,t+h}$ 为上证指数从第i天的t半小时起，后续h个半小时时段的收益率，比如h=[0, 5]表示股票在第[t+0, t+5]这一持有期内的收益率。
- 本文以h=[0, 5]衡量短期，以 h=[6, 20]衡量长期，研究情绪的预测作用在长期内是否会发生反转。

表 8

短期内 $h=[0,5]$ 投资者情绪对日内收益的预测作用

	第二个 半小时	第三个 半小时	第四个 半小时	第五个 半小时	第六个 半小时	第七个 半小时	第八个 半小时
滞后半小时	0.0027	0.0379	0.0822	0.0582*	0.0396***	0.0400**	0.0705**
投资者情绪	(0.0798)	(0.7381)	(1.2615)	(1.6534)	(2.8368)	(2.3951)	(2.4465)
第一半小时	0.0650	0.1925***	0.1982**	0.1797*	0.1756***	0.0518	0.0267
收益率	(0.9722)	(2.9253)	(2.0483)	(1.9117)	(3.6158)	(0.5165)	(0.6022)
滞后半小时		0.0940	0.0340	0.0025	0.4469***	0.2777*	0.1786**
收益率		(0.6018)	(0.2115)	(0.0160)	(4.9278)	(1.7879)	(2.3005)
滞后半小时	-0.3028	0.7320	-0.6239	-0.5781	-0.5141*	-1.2923**	-0.5216*
波动率	(-1.5936)	(1.3163)	(-0.9765)	(-0.9593)	(-1.8496)	(-2.0392)	(-1.9685)

表 9

长期内 ($h=[6,20]$) 投资者情绪对日内收益的预测作用

	第二个 半小时	第三个 半小时	第四个 半小时	第五个 半小时	第六个 半小时	第七个 半小时	第八个 半小时
滞后半小时投	-0.0040	-0.1480	-0.1163	-0.1494*	-0.0274	-0.0039**	-0.0428*
资者情绪	(-0.0425)	(-1.6111)	(-1.3684)	(-1.8176)	(-0.3382)	(-2.4465)	(-1.6780)
第一半小时收	0.1524	0.0498	0.0621	0.0058	0.0105	0.0335	0.0269
益率	(0.8637)	(0.2653)	(0.4232)	(0.0705)	(0.1263)	(0.1921)	(0.3303)
滞后半小时收	—	0.3760	0.0981	-0.2525*	-0.3449**	-0.6381***	-0.5625***
益率		(1.3866)	(0.4993)	(-1.8894)	(-2.2259)	(-2.7362)	(-3.9439)
滞后半小时波	0.3954	0.4745	0.6695	-1.5641***	-1.3825***	-1.6621	-0.7843
动率	(0.7032)	(0.4161)	(0.5995)	(-2.8227)	(-2.9130)	(-1.2133)	(-1.6108)

四、文章结论

- 本文抓取上证指数股吧的实时发帖，通过文本语义分析构建了投资者日内高频情绪指标，并研究了其对股市盘中收益的预测效应。研究发现，中国股票市场的日内投资者情绪能正向预测股票市场运行，这种预测作用下午交易时段表现得更显著
- 牛市中投资者情绪对日内收益率的预测作用强于滞后收益率等变量，熊市则相反，但在暴涨或暴跌的极端市场环境中，情绪对日内收益率的影响程度相较于滞后收益率等变量更为显著。
- 隔夜投资者情绪的释放会显著影响次日上午的市场收益率，但存在时滞性；午间休市期间的投资者情绪会与上午收益率一起正向影响下午的市场表现；进一步看，噪音交易是投资者情绪影响股票收益率的重要驱动力。