

市场情绪、投资因子与资产定价 ——基于新闻与社交媒体的对比分析

周颖刚^{1,2}, 唐诚蔚³, 林哲晖⁴

(1. 厦门大学经济学院, 厦门 361005; 2. 厦门大学王亚南经济研究院, 厦门 361005; 3. 厦门大学邹至庄经济研究院, 厦门 361005; 4. 厦门国际银行总行公司金融部, 厦门 361005)

摘要 本文利用汤森路透市场心理指数 (Thomson Reuters MarketPsych Indices) 中个股层面的情绪数据和美国股票市场 2010 至 2019 年期间的交易数据, 比较分析了新闻情绪和社交媒体情绪在日度和月度两个不同时间维度下对股票定价能力的差异。实证结果表明, 社交媒体情绪在日度层面的表现要优于新闻情绪, 而新闻情绪在月度层面对股票收益率的解释能力要强于社交媒体情绪。具体来说, 在日度层面, 本文构建了新闻情绪因子和社交媒体情绪因子, 发现在 Fama-French 五因子模型下, 社交媒体情绪因子有显著的超额收益, 而新闻情绪因子不存在超额收益, 并且社交媒体情绪因子能够解释大部分日度层面的市场异象, 而新闻情绪因子无法解释日度层面的异象。格兰杰因果检验的结果表明社交媒体情绪因子的反应速度比新闻情绪因子快 3 至 4 个交易日, 说明了社交媒体情绪因子领先于新闻情绪因子。在月度层面, 本文发现新闻情绪因子对异象的解释能力有所改善, 而社交媒体情绪因子对异象的解释能力大幅下降。此外, 对于波动率异象和特质波动率异象来说, 月度新闻情绪因子有较为显著的解释能力, 而月度社交媒体情绪因子的解释能力不显著。

关键词 社交媒体; 新闻; 市场情绪; 资产定价

收稿日期: 2024-02-01

基金项目: 国家自然科学基金 (71988101); 国家社会科学基金重大项目 (19ZDA060)

Supported by National Natural Science Foundation of China (71988101); Major Program of National Social Science Foundation of China (19ZDA060)

作者简介: 周颖刚, 博士, 厦门大学经济学院、王亚南经济研究院教授, 研究方向: 实证资产定价、国际金融与贸易、金融与房地产市场, E-mail: yinggang.zhou@gmail.com; 通信作者: 唐诚蔚, 博士研究生, 厦门大学邹至庄经济研究院, 研究方向: 金融风险管理、空间计量经济学、实证资产定价, E-mail: tangchw_20@126.com; 林哲晖, 硕士, 厦门国际银行总行公司金融部, 研究方向: 实证资产定价, E-mail: 729862481@qq.com.

Market Sentiment, Investment Factor and Asset Pricing: Comparative Analysis Based on News and Social Media

ZHOU Yinggang^{1,2}, TANG Chengwei³, LIN Zhehui⁴

(1. School of Economics, Xiamen University, Xiamen 361005, China; 2. Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen 361005, China; 3. Paula and Gregory Chow Institute for Studies in Economics, Xiamen University, Xiamen 361005, China; 4. Corporate Finance Department, Xiamen International Bank Head Office, Xiamen 361005, China)

Abstract This paper compares and analyzes the differences in stock pricing between news sentiment and social media sentiment in two different time dimensions, daily and monthly, using individual sentiment data from the Thomson Reuters MarketPsych Indices and trading data from the US stock market from 2010 to 2019. The empirical results indicate that social media sentiment performs better at the daily level than news sentiment, and news sentiment has a stronger explanatory power on stock returns at the monthly level than social media sentiment. Specifically, at the daily level, this paper constructs news sentiment factor and social media sentiment factor, and finds that social media sentiment factor still exhibits significant excess returns under the Fama-French five-factor model, while news sentiment factor no longer exhibits excess returns. In addition, social media sentiment factor can explain most market anomalies at the daily level, while news sentiment factor cannot. In order to investigate the reasons, this paper conducts a Granger causality test, indicating that the response speed of social media sentiment factor is 3 to 4 trading days faster than that of news sentiment factor. At the monthly level, this paper finds that news sentiment improves its ability to explain anomalies, while the explanatory power of social media decreases significantly. In addition, for volatility anomalies and idiosyncratic volatility anomalies, the monthly news sentiment factor has a significant explanatory power, while the explanatory power of the monthly social media sentiment factor is not significant.

Keywords social media; news; market sentiment; asset pricing

1 引言

自进入 21 世纪以来,大量的实证研究文献已经明确证实,市场情绪对股票价格具有显著的影响.特别值得一提的是,2021 年 1 月,美国股票社交论坛爆发了一股逼空热潮,并迅速蔓延至推特等社交媒体平台.在这股热潮的推动下,投资者们被舆论引导,纷纷哄抬 GME (游戏驿站) 这只基本面并不理想的股票的价格.原本被众多机构做空、面临退市风险的 GME 股票,其股价竟从最低的 3 美元一路飙升,最高涨至 483 美元.这股高涨的市场情绪还进一步波及到黑莓、AMC 院线和诺基亚等其他股票,导致它们的股价也多次暴涨暴跌,严重脱离了其基本面,引发了市场的广泛关注.同样地,在 A 股市场,利用消息面引发市场情绪来炒作短线题材股的现象十分普遍.投资者们往往热衷于炒作那些基本面不佳但题材性强的股票,这些股票能够引起人们极大的兴趣.例如,在 2020 年年中火爆一时的抖音概念股广集

团,尽管其基本面相对一般,但作为抖音的海外代理,却因其题材性吸引了大量投资者的关注,股价在炒作初期便一路涨停,最后翻倍上涨,期间更是频频登上各大股票论坛和微信公众号的头条。这种高涨的情绪也带动了其他相关概念股的股价上涨。以上案例表明,媒体在市场情绪的传播中扮演着重要的角色,会对股票价格产生一定的影响。研究市场情绪对股价的影响是十分重要的,运用市场情绪来解释股票收益率已成为当前资产定价领域的一个热门话题。通过深入研究市场情绪,我们可以更加完善地解释股票收益率的变动,为投资者提供更加准确的投资决策依据。

在信息化时代,社交媒体和新闻媒体作为信息传播的重要载体和渠道,对投资者的情绪有着深远影响。然而,值得注意的是,这两类媒体在情绪表达和传播方式上存在着差异,这种差异可能会导致它们对股票市场产生不同的影响。社交媒体以其即时性、互动性和广泛覆盖性等特点,成为投资者获取信息和交流观点的重要平台。在社交媒体上,投资者可以迅速获取市场动态、分享投资经验、讨论股票走势,并通过点赞、转发和评论等方式表达自己的情绪和态度。这些情绪和态度不仅能及时反映投资者的心理预期,还可能在一定程度上影响他们的行为,进而对股票市场产生一定的影响 (Antweiler and Frank, 2004; Das and Chen, 2007)。相比之下,新闻媒体作为传统的信息传播媒介,在报道新闻事件、分析市场趋势等方面具有更高的权威性和专业性。同样地,新闻报道的情绪表达也可能对投资者的心理和行为产生影响 (Tetlock et al., 2007, 2008, 2010, 2011),例如,乐观的新闻报道可能激发投资者的乐观情绪,推动股价上涨;而悲观的新闻报道则可能引发投资者的恐慌情绪,导致股价下跌。和社交媒体不同的是,新闻媒体的报道往往需要经过严格的筛选和审核,这虽然能够提供更全面、更深入的市场信息,但也放缓了信息的传播速度,使得新闻媒体对情绪的影响相对于社交媒体而言更为滞后。为了研究不同媒体情绪对股票市场的影响,本文将投资者情绪分为新闻情绪与社交媒体情绪,通过投资组合和因子模型的方法,对两种情绪的日度定价能力和月度定价能力进行实证分析。

本文使用了汤森路透市场心理指数中个股层面的新闻情绪和社交媒体情绪数据,其样本期为 2010–2019 年,研究了不同类型的媒体情绪对美国股票收益率的解释能力。本文首先根据日度层面的新闻情绪和社交媒体情绪的大小,将股票平均分成十个投资组合,持有期为 1 个交易日,每个投资组合的收益率为组合内所有股票按其市值加权的平均收益率。本文发现随着新闻和社交媒体的情绪由低到高,其对应投资组合的收益率和夏普比率都会逐渐增加。双变量分组分析也得到了类似的结果,但社交媒体情绪对组合收益率的影响展现出了比新闻情绪更好的单调性。

然后,本文参考 Fama and French (1993, 2015) 的因子构造方法,构建了新闻情绪因子和社交媒体情绪因子,发现日度的社交媒体情绪策略要比新闻情绪策略表现得更好,并且在 Fama-French 五因子模型下,社交媒体情绪因子仍有显著的超额收益,而新闻情绪因子不存在超额收益。进一步地,本文对新闻情绪因子和社交媒体情绪因子的定价能力进行了比较,发现相较于五因子模型来说, Fama-French 五因子 + 新闻情绪因子的六因子模型无法很好地解释市值-社交媒体情绪的 3×3 组合 (即根据市值和社交媒体情绪构造的 9 个投资组合) 的超额收益率,而 Fama-French 五因子 + 社交情绪因子的六因子模型能够进一步提高对市值-新闻情绪的 3×3 组合的超额收益率的解释能力。从情绪因子的回归系数 β 来看,社交媒体

情绪因子系数的显著个数和大小均远高于新闻情绪因子,这说明了日度社交媒体情绪因子能够解释与新闻情绪相关的投资组合的收益率,反之则不行。在对异象的解释能力方面,本文发现日度社交媒体情绪因子能够解释盈利异象、动量异象、短期反转异象和长期异象,而日度新闻情绪因子则无法解释这些异象。以上结果均表明,在日度层面,社交媒体情绪因子对股票收益率的解释能力远远强过新闻情绪因子。本文认为社交媒体情绪因子对股票收益率具有更强解释能力的原因可能是,社交媒体对信息反应具有即时性,而新闻媒体对信息反应存在滞后性。本文通过格兰杰因果检验发现社交媒体情绪因子领先新闻情绪因子大约3至4个交易日。

从长期来看,日度的信息会被市场较为充分地消化,新闻媒体的可靠性和专业性可能会使得新闻媒体情绪对股票收益率具有更强的解释能力。为验证该猜想,本文在进一步的研究中将每家公司的日度情绪数据进行算数平均,得到月度层面的新闻情绪变量和社交媒体情绪变量,并在此基础上构建了月度的新闻情绪因子和社交媒体情绪因子。同样地,本文检验了月度情绪因子对账面市值比异象、投资异象、盈利异象、动量异象、短期反转异象和长期反转异象的解释能力,发现从日度到月度,新闻情绪对异象的解释能力有所改善,而社交媒体情绪因子对异象的解释能力大幅下降。此外,对于波动率异象和特质波动率异象来说,月度新闻情绪因子有较为显著的解释能力,而月度社交媒体情绪因子的解释能力不再显著。

本文的边际贡献有三点:首先,已有文献大多是关注于市场层面的整体情绪 (Baker and Wurgler, 2006; Huang et al., 2015; Jiang et al., 2019; Chen et al., 2023) 或者根据交易数据构建个股层面的情绪指标 (Aboody et al., 2018),而本文则是基于新闻媒体和社交媒体的非结构化文本数据,从日度和月度的角度,研究了个股层面的情绪对股票价格的影响;其次,现有文献只是单独地去研究社交媒体情绪或者新闻情绪,鲜有文献将这两者进行对比研究,而本文无论是在月度还是日度层面,都将社交媒体情绪和新闻情绪进行了对比分析;第三,从实证结论上来看,本文发现社交媒体情绪在日度层面的表现要优于新闻情绪,而新闻情绪在月度层面对股票收益率的解释能力要强于社交媒体情绪,丰富了市场情绪相关的文献。

本文的结构安排如下:第二章是文献综述,介绍了情绪与资产定价、市场情绪构建方法、以及新闻情绪和社交媒体情绪相关的文献。第三章是数据部分,详细介绍了本文所使用的汤森路透的市场心理指数数据库。第四章是实证部分,研究了日度社交媒体情绪和日度新闻情绪对股票收益率的解释能力。第五章是进一步研究,从月度的层面,比较了新闻情绪和社交媒体情绪对股票收益率的解释能力。第六章是结论与展望。

2 文献综述

2.1 情绪与资产定价

投资者情绪是影响股票价格的重要因素。一般来说,投资者情绪通过驱动市场中的非理性行为,使得股票被错误定价 (De Long et al., 1990)。投资者情绪主要通过三种渠道对股票价格产生影响:第一个渠道是增强投机交易的倾向和加大套利交易的难度,这主要体现在投资者情绪对投机性股票具有更强的预测能力,具体来说,基本面较差和高波动的股票在投资者情绪较低时会有更高的未来回报,而在投资者情绪较高时会有更差的未来表现 (Baker and Wurgler, 2006; Da et al., 2015);第二个渠道是影响投资者的风险偏好,高风险与高收益之间

的正向关系只在市场情绪低迷时显著,而在市场情绪高涨时不再成立 (Yu and Yuan, 2011; Shen et al., 2017); 第三个是不确定性, Birru and Young (2022) 发现市场的不确定性能增强投资者情绪对股票未来收益率的预测能力,并且在不确定性较高的时候,情绪对被错误定价的股票具有更强的预测能力。此外,投资者情绪能进一步增强市场异象, Stambaugh et al. (2012) 发现投资者情绪会使得股票价格被高估,并且高投资者情绪会进一步加大异象的收益率,而 Liu et al. (2019) 发现加入情绪因子的因子模型能大大提高对市场异象的解释能力。

除了投资者情绪之外,管理层的情绪、雇员情绪对股票收益率也具有一定的预测能力。在情绪高涨的时期,管理者的定性披露往往更加乐观 (Bochkay and Dimitrov, 2014)。Seybert and Yang (2012) 发现,管理层的收入能够较好地预测投资者情绪。Jiang et al. (2019) 构建了一个总体指标来衡量市场上的经理情绪,发现了无论是样本内还是样本外,经理情绪对股票市场的未来总回报都有一个很强的负向预测效果。Chen et al. (2023) 利用公司员工在 Glassdoor 上发布的雇主评论,构建了员工情绪,发现高的员工情绪能够预测随后的低市场回报。

2.2 投资者情绪的测度方法

Baker and Wurgler (2006) 最早通过主成分分析法,将封闭式基金折价率 (Lee et al., 1991; Swaminathan, 1996)、市场换手率 (Scheinkman and Xiong, 2003)、IPO 数量、IPO 首日收益率 (Ljungqvist and William, 2003; 韩立岩和伍燕然, 2007)、新增开户数等情绪代理变量的主成分作为市场层面的投资者情绪。但是, Huang et al. (2015) 认为主成分分析方法不能准确地捕捉投资者情绪,因为这种方法很难过滤情绪代理变量中的偏误。他们提出了运用偏最小二乘的方法来消除情绪代理指标中常见的噪声成分,并构造了一种新的投资者情绪指数,该指数在样本内和样本外均有更强的预测能力。此外,股票层面的交易数据也是测度投资者情绪的良好指标,比如说个股的换手率 (Baker and Stein, 2004; Lee, 2013; Liu et al., 2019)、非主力资金的净流入程度 (Kumar and Lee, 2006; 何诚颖等, 2021) 和隔夜收益率 (Aboody et al., 2018) 等。

除了交易数据之外,文本等非结构化数据为投资者情绪指标的构建提供了新的方向。就目前研究来看,词典法和自然语言处理技术是当前构造情绪指标的主流方法。Henry (2008) 在电信和计算机行业的公司范围内提供了研究经理情绪相关的新闻稿样本。Price et al. (2012) 使用 Henry (2008) 的单词列表构建了能够衡量经理在业绩汇报电话会议期间的情绪。Loughran and McDonald (2016) 创建了一个在商业语境中使用的情感词汇综合列表。唐国豪等 (2016) 整理了词汇分类字典法和文本词汇加权法等热门文本分析方法,分别从横截面和时间序列两个维度研究了资产价格与市场情绪之间的关系。Chen et al. (2023) 将员工情绪定义为积极评论占比与消极评论占比之差。姜富伟等 (2021) 在 Loughran and McDonald (2016) 的基础上扩充了中文金融情感词典。语调信息也能够反映出市场情绪,比如说, Jiang et al. (2019) 基于 10Ks、10-Qs 和电话会议的语调,构建了经理情绪指数。除此之外,以概率主体模型 (LDA) 为代表的自然语言处理方法在构造情绪指标方面也得到了进一步的发展 (Blei et al., 2003; Calomiris and Mamaysky, 2019)。

2.3 社交媒体情绪和新闻情绪

社交媒体所反映的投资者情绪也在资产定价领域发挥着重要的作用 (张维等, 2022). 自从 Antweiler and Frank (2004) 以及 Das and Chen (2007) 利用互联网留言板的数据对股票市场的可预测性进行开创性工作以来, 越来越多的研究人员试图探索网络中的文本信息, 为金融市场提供预测. Bollen et al. (2011) 发现基于美国社交网站 Twitter 的情绪指标能很好地预测道琼斯指数的波动率. 郑志刚等 (2011) 对于媒体的负面报道进行了实证分析, 发现上市公司受到媒体的负面报道后, 公司在下一期的财报状况都会有明显改善, 并提出媒体的报道能够帮助公司实现治理的原因可能是媒体的负面报道会引发投资者的关注, 这种情况也是对于经理人的一种外部约束. 游家兴和吴静 (2012) 基于报道基调、曝光程度、关注水平这三个角度的相关数据, 运用文本分析的方法, 构建了有关媒体情绪的一套评价体系, 并发现当媒体报道的情绪波动较大的时候, 股票也会随之波动. 汪昌云和武佳薇 (2015) 使用了一些主流财经媒体的词汇数据, 构造了相关的情绪指标, 发现媒体的负面语气对于 IPO 抑价率等指标有着较好的解释能力. Renault (2017) 从社交媒体上发布的信息中获得投资者情绪, 分析了在线投资者情绪与日内股票回报之间的关系.

新闻对股票市场的影响是资产定价领域的重要问题 (Fama et al., 1969; Roll, 1984). 不少文献研究了个股或股市指数的表现与新闻报纸文章中相关文本之间的关系 (Tetlock et al., 2007, 2008, 2010, 2011; Garcia, 2013; 张飞鹏等, 2024). Akyildirim et al. (2015) 研究了新闻公告对个股流动性、价格和波动性的作用. Manela and Moreira (2017) 使用 1890 年后华尔街日报的头版文章构建了一个基于新闻文本的不确定性度量, 发现在股市崩盘、政策不确定性时代、世界大战和金融危机期间, 新闻隐含波动率都达到了顶峰. 新闻情绪甚至在行业之间、国家之间存在溢出效应 (Hammoudeh et al., 2009; Audrino and Teterova, 2019). 此外, 根据新闻情绪构建的策略也能带来超额收益, 林建浩等 (2022) 发现小市值、低换手率和低 Beta 的股票对新闻吸收速度较慢, 为基于机器学习方法的新闻情绪策略提供了套利空间.

综上所述, 已有文献大多是关注于市场层面的整体情绪或者根据交易数据构建个股层面的情绪指标, 且大部分的研究对象为月度的情绪指标. 此外, 鲜有文献将新闻媒体情绪和社交媒体情绪进行对比分析. 和以往文献不同, 本文将从日度和月度的角度, 对社交媒体情绪和新闻情绪进行对比研究.

3 数据

3.1 汤森路透市场心理指数 (Thomson Reuters MarketPsych Indices, TRMI)

本文使用的情绪数据来自汤森路透的市场心理指数 (TRMI) 数据库. TRMI 是一种先进的语言指数, 它使用 AI 的自然语言处理技术来解释数据源的变化和词义的变化, 对特定的公司、货币、商品和国家的在线媒体来源进行评分. 与最流行的方法相比, TRMI 对语法结构很敏感, 能够解释单词之间的相关性. 此外, TRMI 还比学术界使用的其他数据具有更好的覆盖范围: 汤森路透每天从 2000 多个新闻媒体源和 800 多个社交媒体源获取并处理 200 万篇英文文本文章, 实时更新 TRMI 指数数据; 汤森路透的 TRMI 数据涵盖了来自 75 个以上国家的 12000 多个活跃公司. TRMI 数据在学术界得到了广泛的认可和使用 (Manela and Moreira, 2017; Audrino and Teterova, 2019; Michaelides et al., 2015, 2019). 本文所使用的

是美国上市公司层面的情绪数据, 即 TRMI 中的 Sentiment 情绪指标. 该指标的更新频率是分钟 (60 秒) 和日度 (24 小时), 本文使用的是日度级别的整合情绪指数. TRMI 分别从新闻的维度和社交媒体的维度对市场情绪进行评估. TRMI 对社交媒体与新闻的解释如下:

1) 新闻: 市场心理数据收集的主流新闻来源是整个路透社新闻的历史新闻数据集. 2005 年, 档案开始包括 Lexis Nexis 收集的互联网新闻内容. Lexis Nexis 包括了来自顶级国际和商业新闻来源、顶级区域新闻来源和领先行业来源的内容. 在 2017 年市场心理数据中还额外增加了加密货币资产类别.

2) 社交媒体: 相对于传统新闻, 社交媒体数据收集过程更加多样. 它始于 1998 年的互联网论坛和留言板内容. 从 2008 年底开始, 数据及也增添了 LexisNexis 社交媒体内容. 从 2009 年底开始, 收录了推特社交网站上的发文与评论. 汤森路透以传入链接来衡量社交论坛的受欢迎程度排名, 选出了一些排名靠前的社交论坛, 通常包括排名前 20% 的博客、微博客和其他金融社交媒体. 此外, 市场心理数据还包括数百个不受欢迎的、与特定资产有关的博客和论坛的内容, 这一特点也使得该数据能够较完整地概括整个社交媒体的市场情绪.

本文所采用的 Sentiment 指标是经过标准化处理的, 其取值范围是 -1 至 1 , 该指标越大说明情绪越积极, 反之则说明消极情绪占主导.

3.2 股价及相关数据

本文所使用的数据详情见表 1. 具体来说, 本文的样本一共包含 2823 家美股上市公司, 这些公司均来自于美国证券交易所、纽约证券交易所或纳斯达克交易所, 样本期为 2010 年至 2019 年. 美股价格数据来自于彭博终端, 其中包括了日度与月度的数据. 上市公司的市值、账面市值比等月度财务数据也来自彭博终端.

此外, 市场异象数据以及 Fama-French 因子数据均来自于 Kenneth R. French 的主页. 因子包括了市场因子 (MKT)、市值因子 (SMB)、价值因子 (HML)、盈利因子 (RMW) 和投资因子 (CMA). 由市场异象构造的多空组合包括了: 1) 5×5 市值-账面市值比多空组合 (即股票根据市值-账面市值比分组, 做多高账面市值比所对应的所有市值的股票组合, 同时做空低账面市值比所对应的所有市值的股票组合, 下同); 2) 5×5 市值-投资多空组合; 3) 5×5 市值-盈利多空组合; 4) 5×5 市值-动量多空组合; 5) 5×5 市值-短期反转多空组合; 6) 5×5 市值-长期反转多空组合, 这些组合统一采用高减低的标准方法来构造¹.

表 1 数据介绍

数据类型	数据频率	时间跨度	数据来源
Sentiment 情绪指标	日度	2010.01.01–2019.12.31	汤森路透 TRMI
美股价格数据	日度/月度	2010.01.01–2019.12.31	彭博终端
上市公司财务数据	月度	2010.01.01–2019.12.31	彭博终端
Fama-French 因子数据	日度/月度	2010.01.01–2019.12.31	Kenneth R. French
市场异象数据	日度/月度	2010.01.01–2019.12.31	Kenneth R. French

¹ 本文所使用的数据和代码请参见科学数据银行 (ScienceDB) 期刊社区, DOI: 10.57760/sciencedb.17831 和 CSTR: 31253.11.sciencedb.17831. 若使用文中数据信息, 请注明引文和数据出处.

4 日度情绪与股票收益率

4.1 基于新闻和社交媒体情绪的单变量分组与双变量分组分析

在每个交易日中, 本文分别根据新闻和社交媒体情绪指标的大小, 将股票平均分成十个投资组合, 持有期为 1 个交易日, 每个投资组合的收益率为组合内所有股票按其市值加权的平均收益率. 本文计算了这十个投资组合在时间序列上的平均收益率 (单位: %) 和平均夏普比率, 结果见表 2.

表 2 的 Panel A 展示了基于新闻情绪分组的结果. 可以看到, 不同新闻情绪组合的平均收益率具有一定的单调性, 多空组合的平均收益率为 0.16%, t 值为 16.61, 在 1% 的水平上显著. 同样地, 新闻情绪越高的组合, 其平均夏普比率越高, 多空组合的平均夏普比率为 0.07, t 值为 12.89, 在 1% 的水平上显著. Panel B 展示了基于社交媒体情绪分组的结果, 不同社交媒体情绪组合的平均收益率也具有一定的单调性, 多空组合的平均收益率为 0.18%, t 值为 19.45, 在 1% 的水平上显著. 社交媒体情绪越高的组合, 其平均夏普比率越高, 多空组合的平均夏普比率为 0.07, t 值为 14.01, 在 1% 的水平上显著. 以上结果表明, 随着新闻和社交媒体的情绪由低到高, 其对应投资组合的收益率和夏普比率都会逐渐增加.

在双变量分组分析中, 本文参考 Fama and French (1993, 2015) 的方法, 分别计算了基于市值和新闻情绪的组合收益率与基于市值和社交媒体情绪的组合收益率, 步骤为: 首先根据 2823 家公司的日度社交媒体情绪和日度新闻情绪的大小, 将股票分成低、中和高三组, 比例分别为 30%、40% 和 30%; 然后根据股票的市值大小将股票分成小、中和大三组, 比例分别为 30%、40% 和 30%; 最后, 本文构建了 3×3 的 9 个投资组合, 每个组合的收益率均按照股票市值进行加权平均, 且每个交易日重新构造. 这 9 个投资组合分别是小低 (即小市值和低情绪的组合, 下同)、小中、小高、中低、中中、中高、大低、大中和大高, 各个投资组合的平均收益率 (单位: %) 与 t 统计量结果如表 3 所示.

表 3 的结果表明, 在基于市值和新闻情绪的组合中, 除了小高、大低和大高三个组合的

表 2 基于新闻和社交媒体情绪十等分的投资组合的日度结果

Panel A: 基于新闻情绪的分组							
	低	3	5	7	9	高	高-低
平均收益率	0.00 (-0.25)	-0.20*** (-9.34)	0.06*** (2.97)	0.17*** (8.50)	0.25*** (11.76)	0.15*** (7.44)	0.16*** (16.61)
夏普比率	0.00 (0.09)	-0.07*** (-8.14)	0.03*** (3.69)	0.07*** (8.13)	0.10*** (10.53)	0.07*** (7.51)	0.07*** (12.89)
Panel B: 基于社交媒体情绪的分组							
	低	3	5	7	9	高	高-低
平均收益率	-0.05*** (-2.70)	-0.15*** (-6.98)	0.05** (2.45)	0.23*** (11.76)	0.21*** (10.74)	0.13*** (6.47)	0.18*** (19.45)
夏普比率	-0.01** (-2.07)	-0.05*** (-6.32)	0.01 (1.47)	0.08*** (11.88)	0.09*** (10.98)	0.06*** (6.70)	0.07*** (14.01)

注: 括号中的数字为 t 值, *, **, *** 分别表示 10%、5% 和 1% 的显著性水平.

表3 基于日度情绪划分的投资组合

	基于市值-新闻情绪的分组			基于市值-社交媒体情绪的分组		
	高	中	低	高	中	低
市值						
小	-0.105** (-2.08)	-0.060 (-1.29)	-0.080 (-1.61)	0.109*** (-2.97)	-0.142*** (-3.86)	-0.200*** (-5.33)
中	0.020 (0.72)	-0.014 (-0.41)	0.004 (0.12)	0.039 (1.36)	0.005 (0.16)	-0.013 (-0.35)
大	0.039* (1.82)	0.033 (1.52)	0.041* (1.90)	0.041* (1.86)	0.040* (1.81)	0.037* (1.67)

注: 括号中的数字为 t 值, *, **, *** 分别表示 10%、5% 和 1% 的显著性水平。

均值显著外, 其他几个投资组合的收益率都不显著, 并且组合的平均收益率也没有明显规律。但是, 在基于市值和社交媒体情绪的组中, 除了中等市值的组合不显著外, 小市值和大市值组合的收益率都较为显著, 同时随着组合的情绪从高到低变化 (积极到消极), 投资组合的收益率也随之减小: 在小市值的组合里, 高情绪的投资组合平均收益率从 -0.109% 减小到了 -0.200%, 而在大市值的组合里, 高情绪的投资组合平均收益率从 0.041% 减小到了 0.037%。社交媒体情绪比新闻情绪展现出了更好的单调性。

4.2 日度情绪与定价能力: 新闻情绪 VS 社交媒体情绪

4.2.1 日度情绪因子的构建

在双变量分组的基础上, 本文根据式 (1) 构建出社交媒体情绪因子和新闻情绪因子:

$$SHML = \frac{SH + MH + BH}{3} - \frac{SL + ML + BL}{3}, \quad (1)$$

其中, SHML 表示情绪因子, SH、MH、BH、SL、ML 和 BL 分别表示小高、中高、大高、小低、中低和大低投资组合的收益率。新闻情绪因子和社交媒体情绪因子的描述性统计见表 4 (单位: %).

从表 4 的 Panel A 中可以看出, 新闻情绪因子的均值为 -0.00%, 小于社交媒体情绪因子的均值 0.05%; 新闻情绪因子的标准差为 1.04%, 高于社交媒体情绪因子的标准差 0.78%, 说明新闻情绪因子具有更高的波动; 新闻情绪因子的偏度为 -3.14, 而社交媒体情绪因子的偏度仅为 0.37, 说明社交媒体情绪因子的分布更对称, 而新闻情绪因子为左偏分布; 新闻情绪因子的峰度数据为 42.14, 远远大于社交媒体情绪因子的峰度 7.01. Panel B 展示的是相关系数, 新闻情绪因子与社交媒体情绪因子的相关系数为 0.06, 说明了日度的社交媒体情绪因子和日度的新闻情绪因子之间的差异性较大. Panel C 展示了均值差异检验的结果. 新闻情绪因子和社交媒体情绪因子之间的均值差异为 -0.05%, t 值为 -2.10, 在 5% 的水平上显著, 说明社交媒体情绪因子的平均收益率比新闻情绪因子的更高。

为了进一步比较社交媒体情绪因子和新闻情绪因子的收益, 本文构造了基于新闻情绪的多空组合和基于社交媒体情绪的多空组合。新闻情绪的多空组合指的是, 在每个交易日中做多高新闻情绪所对应的大中小市值组合, 做空低新闻情绪所对应的大中小市值组合; 社交媒

体情绪的多空组合指的是, 在每个交易日中做多高社交媒体情绪所对应的大中小市值组合, 做空低社交媒体情绪所对应的大中小市值组合. 图 1 展示了新闻情绪和社交媒体情绪的多空组合的累积收益率, 其中黑色粗实线的曲线为社交媒体情绪多空组合的累计收益率, 黑色虚线为新闻情绪多空组合的累计收益率. 在此基础上, 本文还加入了标普 500 指数的累积收益率作为基准, 即图中的灰色虚线. 图 1 表明, 根据新闻情绪构造的多空组合的表现远不如标普 500 指数, 而根据社交媒体情绪构造的多空组合的累积收益率远远高于标普 500 指数, 这说明社交媒体情绪策略要比新闻情绪策略表现得更好.

表 4 基于新闻和社交媒体情绪十等分的投资组合的日度结果

Panel A: 情绪因子收益率描述性统计								
	样本数	均值	标准差	中位数	最小值	最大值	偏度	峰度
新闻情绪因子	2516	−0.00	1.04	0.01	−13.17	8.05	−3.14	42.14
社交媒体情绪因子	2516	0.05	0.78	0.05	−5.02	5.88	0.37	7.01
Panel B: 相关系数矩阵								
	新闻情绪因子				社交媒体情绪因子			
新闻情绪因子	1							
社交媒体情绪因子	0.06***				1			
Panel C: 均值差异检验								
	新闻情绪因子		社交媒体情绪因子		均值差异		<i>t</i> 值	
均值	−0.00		0.05		−0.05**		−2.10	

注: 括号中的数字为 t 值, *, **, *** 分别表示 10%、5% 和 1% 的显著性水平.

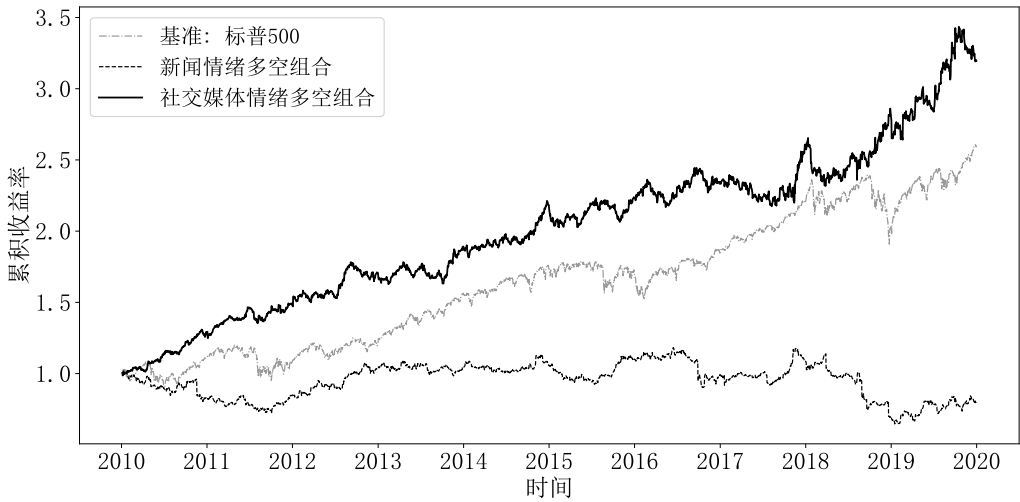


图 1 社交媒体情绪策略与新闻情绪策略

4.2.2 Fama-French 五因子模型对日度情绪因子超额收益的解释能力

Fama and French (2015) 构造了五因子模型,并分别用不同的四个因子组合去解释另外一个因子的组合收益率差,若截距项 α 显著不为零,说明被解释变量这个因子有其他四个解释变量因子所不能够解释的超额收益.依照此思路,本章节利用 Fama-French 模型中的三因子和五因子作为解释变量,用前文所构造的日度社交媒体情绪因子和新闻情绪因子作为被解释变量,探究情绪因子是否有无法被解释的超额收益,结果见表 5 (单位: %).

表 5 展示了新闻情绪因子和社交媒体情绪因子在 Fama-French 三因子模型和五因子模型下的超额收益 α . 可以看到,无论是 Fama-French 三因子模型还是五因子模型,新闻情绪因子的超额收益都不显著,说明新闻情绪因子的超额收益能全部被 Fama-French 三因子模型和五因子模型解释.而社交媒体情绪因子在 Fama-French 三因子模型和五因子模型下的 α 均为 0.05%,且在 1% 的水平上显著,说明 Fama-French 三因子和五因子模型无法全部解释社交媒体情绪因子的超额收益.

表 5 新闻情绪因子和社交媒体情绪因子是否存在超额收益?

	新闻情绪因子	社交媒体情绪因子
Fama-French 三因子模型	-0.00 (-0.19)	0.05*** (3.15)
Fama-French 五因子模型	-0.00 (-0.24)	0.05*** (3.04)

注: 括号中的数字为 t 值, *, **, *** 分别表示 10%、5% 和 1% 的显著性水平.

4.2.3 日度情绪因子对于投资组合的解释能力

Fama and French (2015) 分别用三因子与五因子模型对不同的投资组合进行回归分析,发现五因子模型相比于三因子模型具有更小或更接近于 0 的截距项 α ,表明五因子模型能够解释更多的超额收益.本文参考其方法,构建了五因子 + 新闻媒体情绪因子的六因子模型,来解释市值-社交媒体情绪的 3×3 组合的超额收益,同时也构造了五因子 + 社交媒体情绪因子的六因子模型,来解释市值-新闻情绪的 3×3 组合的超额收益,结果见表 6 和表 7.

表 6 展示的是 Fama 五因子 + 新闻情绪因子的六因子模型对市值-社交媒体情绪的 3×3 组合的解释结果. Panel A 报告的是基准对照组的回归结果,即市值-社交媒体情绪的 3×3 组合在五因子模型下的 α . 可以看到,五因子模型下的 α 都显著为负. Panel B 报告的是市值-社交媒体情绪的 3×3 组合在五因子 + 新闻情绪因子的六因子模型下的 α . 通过对比,本文发现小市值-低情绪组合的 α 略微降低(越不接近 0, 无改善效果),而其他组合的 α 均没有变化. 以上结果说明了加入新闻因子的模型对于市值-社交媒体组合的超额收益的解释能力并没有改善. Panel C 报告的是五因子模型 + 新闻情绪因子下,新闻情绪因子对市值-社交媒体情绪组合的解释能力 β , 可以看到,新闻情绪因子的 β 在小市值-低情绪、中市值-中情绪、中市值-低情绪和大市值-高情绪四个组合中显著,而在其他不同类型的组合中不显著,说明了新闻情绪因子对市值-社交媒体情绪的 3×3 组合的收益率的解释能力有限.

表 6 新闻情绪因子解释基于市值和社交媒体情绪的投资组合 (单位: %)

社交媒体情绪		高	中	低
Panel A: 五因子模型下的 alpha				
市值	小	-0.148***	-0.184***	-0.237***
		(-5.30)	(-6.64)	(-7.98)
		-0.015	-0.048**	-0.065***
	中	(-1.52)	(-2.55)	(-2.65)
		-0.015***	-0.016***	-0.018***
		(-3.29)	(-3.69)	(-4.25)
Panel B: 五因子 + 新闻情绪因子模型下的 alpha				
市值	小	-0.148***	-0.184***	-0.238***
		(-5.29)	(-6.64)	(-8.03)
		-0.015	-0.048**	-0.065***
	中	(-1.52)	(-2.56)	(-2.64)
		-0.015***	-0.016***	-0.018***
		(-3.29)	(-3.68)	(-4.26)
Panel C: 五因子 + 新闻情绪因子模型下, 新闻情绪因子的解释能力 beta				
市值	小	4.265	1.416	-12.187***
		(1.58)	(0.53)	(-4.28)
		0.422	-3.454*	4.533*
	中	(0.44)	(-1.91)	(1.93)
		0.924**	0.579	-0.359
		(2.10)	(1.39)	(-0.86)

注: 括号中的数字为 t 值, *, **, *** 分别表示 10%、5% 和 1% 的显著性水平. alpha 是因子模型回归的截距项, beta 是因子模型回归中新闻情绪因子的系数.

表 7 展示的是 Fama 五因子 + 社交媒体情绪因子的六因子模型对市值-新闻媒体情绪 3×3 组合的解释结果. Panel A 报告的是基准对照组的回归结果, 即市值-新闻情绪的 3×3 组合在五因子模型下的 alpha. 可以看到, 在所有的市值-社交媒体情绪的组合中, 五因子模型下的 alpha 都显著为负, 且 alpha 呈现了一定的单调性. Panel B 是加入了社交媒体因子后的回归结果, 可以看出, 除了小市值-高情绪的组合和大市值-高情绪的组合没有改善之外, 其他的组合都有不同程度的改善, 即 alpha 更接近于 0. 具体来说, 小市值-中情绪和小市值-低情绪组合中的 alpha 分别从 -0.093 和 -0.125 增加到了 -0.089 和 -0.119; 中市值-高情绪、中市值-中情绪和中市值-低情绪组合中的 alpha 分别从 -0.032、-0.068 和 -0.050 增长到了 -0.029、-0.059 和 -0.045, 这说明无论情绪的高低, 社交媒体因子对于中等市值的公司股票的超额收益率都有着较好的解释能力; 同样地, 在大市值的组合中, alpha 也得到了不同程度的改善, 大市值-中情绪和大市值-低情绪组合中的 alpha 分别从 -0.023 和 -0.013 增加到了 -0.022 和 -0.012. Panel C 报告的是五因子模型 + 社交媒体情绪因子下, 社交情绪因子对市值-新闻情绪组合收益率的解释能力 beta, 可以看到, 社交媒体情绪因子的 beta 在小市值-中情绪、小市值-低情绪、中市值-低情绪、中市值-中情绪、中市值-高情绪、大市值-中

表 7 社交媒体情绪因子解释基于社交媒体情绪的投资组合 (单位: %)

新闻媒体情绪		高	中	低	
Panel A: 五因子模型下的 alpha					
市值	小	-0.152*** (-3.48)	-0.093** (-2.35)	-0.125*** (-2.85)	
		中	-0.032*** (-3.14)	-0.068*** (-3.22)	-0.050*** (-3.23)
			大	-0.016*** (-3.77)	-0.023*** (-5.82)
	Panel B: 五因子 + 社交媒体情绪因子模型下的 alpha				
	市值	小		-0.154*** (-3.53)	-0.089** (-2.23)
			中	-0.029*** (-2.88)	-0.059*** (-2.79)
大				-0.016*** (-3.79)	-0.022*** (-5.65)
		Panel C: 五因子 + 社交媒体情绪因子模型下, 新闻情绪因子的解释能力 beta			
		市值	小	5.130 (0.91)	-9.555* (-1.87)
中				-5.670*** (-4.35)	-20.623*** (-7.65)
	大			0.223 (0.40)	-1.442*** (-2.87)

注: 括号中的数字为 t 值, *, **, *** 分别表示 10%、5% 和 1% 的显著性水平. alpha 是因子模型回归的截距项, beta 是因子模型回归中社交媒体情绪因子的系数.

情绪、大市值-低情绪的七个组合中显著, 说明了社交媒体情绪因子对市值-新闻情绪的 3×3 组合的收益率有着较强的解释能力.

以上结果表明, 相对于 Fama-French 五因子模型来说, 五因子 + 新闻情绪因子模型无法很好地解释市值-社交媒体情绪的 3×3 组合的超额收益率, 而五因子 + 社交情绪因子模型能够进一步改善对市值-新闻情绪的 3×3 组合超额收益率的解释能力. 此外, 从情绪因子的回归系数 beta 来看, 社交媒体情绪因子系数的显著个数和大小均远高于新闻情绪因子, 这进一步说明了日度社交媒体情绪因子对股票收益率的解释强度高于日度新闻情绪因子.

4.2.4 日度情绪因子对于市场异象的解释能力

为进一步检验社交媒体情绪因子和新闻情绪因子对市场异象的解释能力, 本文在 Fama-French 五因子模型上加入了新闻情绪因子和社交媒体情绪因子, 模型如式 (2) 所示, 并尝试用该因子模型去解释账面市值比异象、投资异象、盈利异象、动量异象、短期反转异象和长期反转异象, 比较社交媒体情绪因子和新闻情绪因子的强弱, 结果见表 8.

$$R_t = \alpha + \beta_1 \text{SHML_Social}_t + \beta_2 \text{SHML_News}_t + X_t^T \theta + \varepsilon_t, \quad (2)$$

表 8 日度情绪因子对市场异象的解释能力

	账面市值比	投资	盈利	动量	短期反转	长期反转
社交媒体情绪因子	-0.58 (-1.52)	-0.10 (-0.28)	0.80* (1.81)	4.06** (2.53)	6.40*** (3.15)	4.75*** (4.58)
新闻情绪因子	-0.33 (-1.16)	0.23 (0.83)	0.42 (1.28)	1.63 (1.36)	1.31 (0.87)	0.24 (0.31)
控制 FF5 因子	是	是	是	是	是	是

注：括号中的数字为 t 值，*、**、*** 分别表示 10%、5% 和 1% 的显著性水平。

其中, R_t 是市场异象的收益率, $SHML_Social_t$ 是社交媒体情绪因子, $SHML_News$ 是新闻情绪因子, X_t^T 对应着 Fama-French 五因子.

表 8 报告了社交媒体情绪因子和新闻情绪因子的系数大小. 可以看到, 当被解释变量为账面市值比异象时, 社交媒体情绪因子和新闻情绪因子都不显著, 说明情绪因子对此异象没有较好的解释能力; 当被解释变量为投资异象时, 社交媒体因子和新闻因子不显著, 说明情绪因子对投资异象没有较好的解释能力; 当被解释变量为盈利异象时, 社交媒体情绪因子的系数显著为 0.80, t 统计量为 1.81, 在 10% 的水平上显著, 但新闻情绪因子的系数不显著, 这说明社交媒体情绪因子能够很好地解释盈利异象, 而新闻情绪因子对此异象无法解释; 当被解释变量为动量异象的时候, 社交媒体因子的系数 4.06, t 统计量为 2.53, 在 5% 的水平上显著, 而新闻情绪因子不显著, 这说明社交媒体情绪因子能够较好地解释动量效应, 而新闻情绪因子则无法解释; 当被解释变量为短期反转异象时, 社交媒体情绪因子在五因子的基础之上仍然有强的解释能力, 其回归系数为 6.40, t 统计量为 3.15, 在 1% 的水平上显著, 说明社交媒体情绪因子是短期反转异象的一个很好的补充解释, 但新闻情绪因子对于此类异象没有显著的解释能力; 当被解释变量为长期反转异象时, 社交媒体情绪因子依旧有较显著的回归结果, 其回归系数为 4.75, t 统计量为 4.58, 在 1% 的水平上显著, 而新闻情绪因子不显著, 说明了在解释长期反转异象时, 社交媒体情绪因子依旧比新闻情绪因子有更大的优势. 总的来说, 日度社交媒体情绪因子能够对盈利异象、动量异象、短期反转异象和长期反转异象具有一定的解释能力, 而新闻情绪因子无法解释其中的任何一个异象, 表明日度社交媒体情绪因子对异象的解释能力强于新闻情绪因子, 是 Fama-French 五因子模型的一个较好的补充.

4.3 日度社交媒体情绪因子解释能力更强的可能原因

本文认为, 在日度层面, 社交媒体情绪因子的表现之所以比新闻情绪因子好, 其中可能的原因在于社交媒体的发布速度和反应速度要比新闻媒体快, 这一点主要体现在社交媒体发文的便捷性和低门槛性: 社交媒体给人们提供了一个可以及时表达对当前信息看法和情绪的平台, 而新闻媒体的发文推送需要校对、核验、编辑、审核等操作, 导致其中的情绪信息会相对滞后. 具体来说, 人们在某个时间点接收到某一个事件的消息时, 可以立即在社交媒体上发表自己的看法、意见, 并且不止是投资者可以发表意见, 管理层也可以发表意见, 这样就做到了人群的大面积覆盖, 可以很有效地度量出股票在当天投资者情绪, 并及时传导到股价上. 而新闻则完全不同, 对于新闻报道来说, 首先要对消息进行核对, 接着对消息进行整合、撰写、编辑等相关工作, 最后到审核阶段, 审核通过才能发表, 这些虽然能够保证新闻的严谨性、准确

性以及权威性,但是效率较低,无法实时地反应出当下的市场情绪,相对于社交媒体的反应要慢得多,所以在日度层面无法获得超额收益。

为了研究日度社交媒体情绪因子和日度新闻情绪因子之间的领先滞后关系,本文建立如式(3)和(4)所示的时间序列模型,并进行格兰杰因果检验,结果见表9。

$$\text{SHML_Social}_t = \delta_1 + \sum_{i=1}^s \lambda_{1,i} \text{SHML_Social}_{t-i} + \sum_{i=1}^s \gamma_{1,i} \text{SHML_News}_{t-i} + \varepsilon_{1,t}, \quad (3)$$

$$\text{SHML_News}_t = \delta_2 + \sum_{i=1}^s \lambda_{2,i} \text{SHML_Social}_{t-i} + \sum_{i=1}^s \gamma_{2,i} \text{SHML_News}_{t-i} + \varepsilon_{2,t}, \quad (4)$$

其中, SHML_Social_t 是社交媒体情绪因子, SHML_News_t 是新闻情绪因子, s 是最大滞后期。

从表9中的格兰杰因果检验的结果可以看出,当滞后期数为1期(即滞后1个交易日,下同)和2期时,社交媒体情绪因子不是新闻情绪因子的格兰杰原因,反之亦然。当滞后期数为3期时,社交媒体情绪因子和新闻情绪因子互为格兰杰原因。当滞后期数为4期和5期时,社交媒体情绪因子对新闻情绪因子的预测能力更加显著,而新闻情绪因子对社交媒体情绪因子不存在显著的预测能力。以上结果说明社交媒体情绪的反应速度比新闻情绪的反应速度更快,社交媒体情绪因子领先新闻情绪因子大约3至4个交易日。

表9 新闻和社交媒体情绪因子格兰杰检验结果

	滞后1期	滞后2期	滞后3期	滞后4期	滞后5期
情绪:					
社交媒体 → 新闻	0.20	1.79	2.33*	2.33**	2.18**
新闻 → 社交媒体	1.33	1.68	2.32*	1.71	1.56

注:表中数值为格兰杰因果检验的 F 统计量,*,**,*** 分别表示 10%、5% 和 1% 的显著性水平。

5 进一步研究

本文研究表明,在日度层面上,社交媒体情绪的优势在于反应迅速,从而对股票价格具有一定的解释能力。然而,在月度层面,日度的信息会被市场较为充分地消化,新闻媒体的可靠性、专业性以及权威性可能会使得新闻媒体情绪对股票收益率具有更强的解释能力。为了继续探究新闻媒体情绪和社交媒体情绪在月度层面的定价差异,本文将每家公司的日度情绪数据进行算数平均,得到月度层面的新闻情绪变量和社交媒体情绪变量。

5.1 月度新闻情绪因子和社交媒体情绪因子的构建

本文分别计算了基于市值和月度新闻情绪的 3×3 组合收益率,与基于市值和月度社交媒体情绪的 3×3 组合收益率,步骤为:首先根据 2823 家公司的月度社交媒体情绪和月度新闻情绪的大小,将股票分成低、中和高三组,比例分别为 30%、40% 和 30%;然后根据股票的市值大小将股票分成小、中和大三组,比例分别为 30%、40% 和 30%;最后,本文构建了 3×3 的 9 个投资组合,每个组合的收益率均按照股票市值进行加权平均,且在每个月重新构

造. 这 10 个投资组合分别是小低、小中、小高、中低、中中、中高、大低、大中和大高, 各个投资组合的平均收益率 (单位: %) 与 t 统计量结果如表 10 所示 (单位: %).

表 10 的结果表明, 在基于市值和月度新闻情绪分组的组合中, 除了中高、中中、中低和大低四个组合的平均收益率外, 其他组合的收益率均显著. 我们发现, 在大市值的组合里, 收益率也随着新闻情绪的增加而增加, 且大市值组合的收益率普遍比小市值组合的平均收益率高得多, 这一点与日度情绪的结果类似. 与日度情绪结果不同的是, 日度的 9 个组合里只有 3 个显著, 而月度层面的组合里有 5 个组合显著, 其原因可能是新闻的相对滞后性导致其不能够快速地反应日度情绪的影响, 但是在月度层面, 新闻则能够更好地捕捉市场的股价信息. 在基于市值和月度社交媒体情绪分组的组合中, 存在着相似的结果. 在分组的结果上, 本文进一步根据式 (1) 构建出月度社交媒体情绪因子和月度新闻情绪因子, 用于下文对异象的检验.

表 10 基于月度情绪划分的投资组合

	基于市值-新闻情绪的分组			基于市值-社交媒体情绪的分组		
	高	中	低	高	中	低
市值						
小	-2.013*** (-3.00)	-3.379*** (-4.74)	-2.596*** (-4.06)	-1.552*** (-2.24)	-4.276*** (-7.39)	-3.951*** (-7.71)
中	0.484 (0.87)	0.274 (0.47)	0.209 (0.38)	1.716*** (3.62)	-0.291 (-0.53)	-0.717 (-1.49)
大	1.004** (2.400)	0.767* (1.79)	0.679 (1.64)	1.920*** (4.88)	0.549 (1.31)	-0.113 (-0.27)

注: 括号中的数字为 t 值, *, **, *** 分别表示 10%、5% 和 1% 的显著性水平.

5.2 月度情绪因子对于市场异象的解释能力

为检验月度社交媒体情绪因子和月度新闻情绪因子对市场异象的解释能力, 本文在 Fama-French 五因子模型上加入了新闻情绪因子和社交媒体情绪因子, 模型如式 (2) 所示, 并尝试用该因子模型去解释账面市值比异象、投资异象、盈利异象、动量异象、短期反转异象和长期反转异象, 比较月度社交媒体情绪因子和月度新闻情绪因子的强弱, 结果见表 11.

表 11 报告了月度社交媒体情绪因子和月度新闻情绪因子的系数大小. 可以看到, 在月度层面, 社交媒体情绪因子仅能解释投资异象, 说明了月度社交媒体情绪因子的解释能力相比日度来说, 出现了大幅的下降. 但是月度新闻情绪因子能够显著解释长期反转异象, 和日度情

表 11 月度情绪因子对市场异象的解释能力

	账面市值比	投资	盈利	动量	短期反转	长期反转
社交媒体情绪因子	0.01 (0.50)	-0.07** (-2.19)	0.04 (1.11)	-0.10 (-0.90)	0.23 (1.64)	-0.08 (-1.01)
新闻情绪因子	0.01 (0.41)	0.01 (0.28)	0.03 (0.64)	-0.09 (-0.72)	-0.26 (-1.51)	-0.26*** (-2.72)
控制 FF5 因子	是	是	是	是	是	是

注: 括号中的数字为 t 值, *, **, *** 分别表示 10%、5% 和 1% 的显著性水平.

绪因子比起来(日度新闻情绪因子无法解释任意一个异象), 月度新闻情绪因子对异象的解释能力略有提高。以上结果表明, 从日度到月度, 新闻情绪因子对异象的解释能力有所改善, 而社交媒体情绪因子对异象的解释能力大幅下降。

5.3 月度情绪因子与波动率异象

在月度层面, 波动率异象和特质波动率异象是十分普遍的, 并且具有很强的显著性和持续性 (Ang et al., 2006, 2009)。本文将进一步检验月度新闻情绪因子和月度社交媒体情绪因子对波动率异象的解释能力, 回归模型如式 (2) 所示。表 12 展示了月度新闻情绪因子和月度社交媒体情绪因子对股票波动率异象的解释结果。在 Panel A 中, 被解释变量为根据股票波动率的大小而十等分的十个投资组合的收益率, 解释变量为月度新闻情绪因子、月度社交媒体情绪因子, Fama-French 五因子作为控制变量。可以看到, 社交媒体因子在所有组合的回归结果都不显著, 说明在月度层面, 社交媒体因子对波动率组合的收益率没有解释能力。然而, 月度新闻情绪因子能够解释大部分波动率组合的收益率。在波动率的多空组合中, 月度新闻情绪因子的系数为 0.59, t 值为 3.56, 在 1% 的水平上显著, 而月度社交媒体情绪因子的系数则不显著, 说明月度新闻情绪因子对波动率异象具有一定的解释能力, 而月度社交媒体情绪因子则无法解释波动率异象。在 Panel B 中, 被解释变量为根据股票特质波动率的大小而十等分的十个投资组合的收益率。同样地, 社交媒体因子在所有组合中的回归结果都不显著, 说明在月度层面, 社交媒体因子对股票特质波动率组合的收益率没有解释能力。相比之下, 月度新闻情绪因子能够解释高特质波动率组合的收益率。在特质波动率的多空组合中, 月度新闻

表 12 月度情绪因子对波动率异象的解释能力

Panel A: 月度情绪因子对波动率异象的解释能力							
波动率	低	3	5	7	9	高	高-低
α	0.24 (1.26)	-0.08 (-0.59)	0.13 (0.74)	-0.08 (-0.33)	-0.18 (-0.59)	-0.64* (-1.66)	-0.88* (-1.89)
新闻	-0.12* (-1.79)	-0.09* (-1.76)	0.01 (0.17)	0.19** (2.09)	0.19* (1.80)	0.47*** (3.40)	0.59*** (3.56)
社交媒体	0.02 (0.36)	-0.01 (-0.15)	-0.09* (-1.77)	0.01 (0.18)	-0.07 (-0.75)	-0.06 (-0.51)	-0.08 (-0.57)
控制 FF5 因子	是	是	是	是	是	是	是
Panel B: 月度情绪因子对特质波动率异象的解释能力							
特质波动率	低	3	5	7	9	高	高-低
α	0.28* (1.81)	-0.11 (-0.77)	0.12 (0.64)	-0.12 (-0.52)	-0.08 (-0.30)	-0.54 (-1.49)	-0.82* (-1.96)
新闻	-0.09 (-1.71)	-0.01 (-0.22)	-0.03 (-0.40)	0.21** (2.60)	0.25** (2.43)	0.37*** (2.87)	0.47*** (3.13)
社交媒体	-0.03 (-0.65)	0.04 (0.93)	-0.04 (-0.81)	-0.09 (-1.39)	-0.10 (-1.24)	-0.02 (-0.21)	0.01 (0.06)
控制 FF5 因子	是	是	是	是	是	是	是

注: 括号中的数字为 t 值, *, **, *** 分别表示 10%、5% 和 1% 的显著性水平。

情绪因子的系数为 0.47, t 值为 3.13, 在 1% 的水平上显著, 而月度社交媒体情绪因子的系数则不显著, 说明月度新闻情绪因子对股票特质波动率异象具有一定的解释能力, 而月度社交媒体情绪因子则无法解释特质波动率异象。

6 结论与展望

6.1 结论

本文基于汤森路透市场心理指数中的情绪指标, 分别研究了日度层面和月度层面的新闻情绪和社交媒体情绪对股票收益率的影响。在日度层面, 本文发现基于日度社交媒体情绪构建的策略能够获得比标普 500 指数更高的累积收益率, 而日度新闻情绪的相关策略却无法获得超过标普 500 指数的表现。本文构建了日度层面的新闻情绪因子和社交媒体情绪因子, 发现在 Fama-French 五因子模型下, 日度社交媒体因子仍具有显著的超额收益率, 而日度新闻情绪因子不存在超额收益率。在定价能力上, 本文发现 Fama-French 五因子 + 社交情绪因子模型能够进一步提高对市值-新闻情绪的 3×3 组合超额收益率的解释能力, 并且日度社交媒体情绪因子能够解释盈利异象、动量异象、短期反转异象和长期异象, 而日度新闻情绪因子则无法解释这些异象, 说明了在日度层面, 社交媒体情绪因子对股票收益率的解释能力远远强过新闻情绪因子。进一步地, 本文通过格兰杰因果检验发现社交媒体情绪因子领先新闻情绪因子大约 3 至 4 个交易日。最后, 本文构建了月度层面新闻情绪因子和社交媒体情绪因子, 发现新闻情绪因子对异象的解释能力有所改善, 而社交媒体情绪因子对异象的解释能力大幅下降。此外, 对于波动率异象和特质波动率异象来说, 月度新闻情绪因子有较为显著的解释能力, 而月度社交媒体情绪因子的解释能力不再显著。

6.2 研究展望

本文在研究中仅使用了美国的情绪数据, 这是由于中国还没有个股层面的情绪数据, 且美国的资本市场发展比较早, 相对于中国较为成熟, 所涉及到的数据质量较高, 噪声更小, 结果真实可靠。对于中国 A 股市场来说, 本文具有一定的借鉴和参考意义。对于美国市场的研究有助于我们更清晰、深刻地认识情绪对资本市场的影响, 从而将其运用到中国市场。中国的互联网迅速发展, 数据类型和总量庞大, 新闻和社交媒体方面的情绪研究有着巨大的潜力。

Brown and Cliff (2005) 认为投资者情绪对未来股票收益的影响在短期内是正向的, 而在长期是负向的, 这可能与投资者的过度自信和外推信念偏误 (Da et al., 2021) 有关。在未来的研究中, 本文计划从投资者行为的角度入手, 深入剖析情绪的产生机制、外推过程以及传染效应, 分别对比新闻情绪和社交媒体情绪对股票价格的长期影响, 并挖掘它们对股票联动性的具体作用。通过这样的研究, 本文期望能够更全面、更深入地理解媒体情绪影响股市的机制, 为投资决策提供更为科学的参考依据。

参 考 文 献

- 韩立岩, 伍燕然, (2007). 投资者情绪与 IPOs 之谜——抑价或者溢价 [J]. 管理世界, (3): 51-61.
- Han L Y, Wu Y R, (2007). Investment Sentiment and IPOs Puzzle — Discount or Premium[J]. Management World, (3): 51-61.
- 何诚颖, 陈锐, 薛冰, 何牧原, (2021). 投资者情绪、有限套利与股价异象 [J]. 经济研究, 56(1): 58-73.

- He C Y, Chen R, Xue B, He M Y, (2021). Investment Sentiment, Limited Arbitrage and Stock Price Anomaly[J]. *Economic Research Journal*, 56(1): 58–73.
- 姜富伟, 孟令超, 唐国豪, (2021). 媒体文本情绪与股票回报预测 [J]. *经济学 (季刊)*, 21(4): 1323–1344.
- Jiang F W, Meng L C, Tang G T, (2021). Media Textual Sentiment and Chinese Stock Return Predictability[J]. *China Economic Quarterly*, 21(4): 1323–1344.
- 林建浩, 张一帆, 陈良源, 邓益萌, (2022). 基于新闻情绪的机器学习交易策略 [J]. *计量经济学报*, 2(4): 881–908.
- Lin J H, Zhang Y F, Chen L Y, Deng Y M, (2022). News Sentiment and Machine Learning Investment Strategy[J]. *China Journal of Econometrics*, 2(4): 881–908.
- 唐国豪, 姜富伟, 张定胜, (2016). 金融市场文本情绪研究进展 [J]. *经济学动态*, (11): 137–147.
- Tang G H, Jiang F W, Zhang D S, (2016). Research Progress on Text Sentiment in Financial Markets[J]. *Economic Perspectives*, (11): 137–147.
- 汪昌云, 武佳薇, (2015). 媒体语气、投资者情绪与 IPO 定价 [J]. *金融研究*, (9): 174–189.
- Wang C Y, Wu J W, (2015). Media Tone, Investor Sentiment and IPO Pricing[J]. *Journal of Financial Research*, (9): 174–189.
- 游家兴, 吴静, (2012). 沉默的螺旋: 媒体情绪与资产误定价 [J]. *经济研究*, 47(7): 141–152.
- You J X, Wu J, (2012). Spiral of Silence: Media Sentiment and the Asset Mispricing[J]. *Economic Research Journal*, 47(7): 141–152.
- 张飞鹏, 徐一雄, 陈曦, 周勇, (2024). 基于新闻文本情绪的区间值股票回报预测研究 [J]. *计量经济学报*, 4(1): 204–230.
- Zhang F P, Xu Y X, Chen X, Zhou Y, (2024). Forecasting Interval Valued Stock Returns Based on News Media Sentiments[J]. *China Journal of Econometrics*, 4(1): 204–230.
- 张维, 李奕, 王鹏飞, (2022). 社交媒体与资本市场研究进展与展望: 基于文献计量的方法 [J]. *计量经济学报*, 2(1): 32–57.
- Zhang W, Li Y, Wang P F, (2022). Social Media and Capital Markets: A Bibliometric Analysis[J]. *China Journal of Econometrics*, 2(1): 32–57.
- 郑志刚, 丁冬, 汪昌云, (2011). 媒体的负面报道、经理人声誉与企业业绩改善——来自我国上市公司的证据[J]. *金融研究*, (12): 163–176.
- Zheng Z G, Ding D, Wang C Y, (2011). Negative Media Reporting, Manager Reputation, and Corporate Performance Improvement: Evidence from Listed Companies in China[J]. *Journal of Financial Research*, (12): 163–176.
- Aboudy D, Even-Tov O, Lehavy R, Trueman B, (2018). Overnight Returns and Firm-specific Investor Sentiment[J]. *Journal of Financial and Quantitative Analysis*, 53(2): 485–505.
- Akyıldırım E, Altarovici A, Ekinci C, (2015). Effects of Firm-specific Public Announcements on Market Dynamics: Implications for High-frequency Traders[M]// *Handbook of High Frequency Trading*: 305.
- Ang A, Hodrick R J, Xing Y, Zhang X, (2006). The Cross-section of Volatility and Expected Returns[J]. *The Journal of Finance*, 61(1): 259–299.
- Ang A, Hodrick R J, Xing Y, Zhang X, (2009). High Idiosyncratic Volatility and Low Returns: International and Further US Evidence[J]. *Journal of Financial Economics*, 91(1): 1–23.
- Antweiler W, Frank M Z, (2004). Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards[J]. *The Journal of Finance*, 59(3): 1259–1294.
- Audrino F, Tetereva A, (2019). Sentiment Spillover Effects for US and European Companies[J]. *Journal of Banking and Finance*, 106: 542–567.

- Baker M, Stein J C, (2004). Market Liquidity as A Sentiment Indicator[J]. *Journal of Financial Markets*, 7(3): 271–299.
- Baker M, Wurgler J, (2006). Investor Sentiment and the Cross-section of Stock Returns[J]. *The Journal of Finance*, 61(4): 1645–1680.
- Birru J, Young T, (2022). Sentiment and Uncertainty[J]. *Journal of Financial Economics*, 146(3): 1148–1169.
- Blei D, Ng A Y, Jordan M, (2003). Latent Dirichlet allocation[J]. *The Journal of Machine Learning Research*, 3: 993–1022.
- Bochkay K, Dimitrov V, (2014). Qualitative Management Disclosures and Market Sentiment[R]. Working Paper.
- Bollen J, Mao H, Zeng X, (2011). Twitter Mood Predicts the Stock Market[J]. *Journal of Computational Science*, 2(1): 1–8.
- Brown G W, Cliff M T, (2004). Investor Sentiment and the Near-term Stock Market[J]. *Journal of Empirical Finance*, 11(1): 1–27.
- Calomiris C W, Mamaysky H, (2019). How News and Its Context Drive Risk and Returns around the World[J]. *Journal of Financial Economics*, 133(2): 299–336.
- Chen J, Tang G, Yao J, Zhou G, (2023). Employee Sentiment and Stock Returns[J]. *Journal of Economic Dynamics and Control*, 149: 104636.
- Da Z, Engelberg J, Gao P, (2015). The Sum of All Fears Investor Sentiment and Asset Prices[J]. *The Review of Financial Studies*, 28(1): 1–32.
- Da Z, Huang X, Jin L J, (2021). Extrapolative Beliefs in the Cross-section: What Can We Learn from the Crowds[J]. *Journal of Financial Economics*, 140(1): 175–196.
- Das, S R, Chen M Y, (2007). Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web[J]. *Management science*, 53(9): 1375–1388.
- De Long J B, Shleifer A, Summers L H, Waldmann R J, (1990). Noise Trader Risk in Financial Markets[J]. *Journal of Political Economy*, 98(4): 703–738.
- Fama E F, Fisher L, Jensen M C, Roll R, (1969). The Adjustment of Stock Prices to New Information[J]. *International Economic Review*, 10(1): 1–21.
- Fama E F, French K R, (1993). Common Risk Factors in the Returns on Stocks and Bonds[J]. *Journal of Financial Economics*, 33(1): 3–56.
- Fama E F, French K R, (2015). A Five-factor Asset Pricing Model[J]. *Journal of Financial Economics*, 116(1): 1–22.
- Garcia D, (2013). Sentiment During Recessions[J]. *The Journal of Finance*, 68(3): 1267–1300.
- Hammoudeh S M, Yuan Y, McAleer M, (2009). Shock and Volatility Spillovers among Equity Sectors of the Gulf Arab Stock Markets[J]. *The Quarterly Review of Economics and Finance*, 49(3): 829–842.
- Henry E, (2008). Are Investors Influenced by How Earnings Press Releases are Written?[J]. *International Journal of Business, Communication*, 45(4): 363–407.
- Huang D, Jiang F, Tu J, Zhou G, (2015). Investor Sentiment Aligned: A Powerful Predictor of Stock Returns[J]. *The Review of Financial Studies*, 28(3): 791–837.
- Jiang F, Lee J, Martin X, Zhou G, (2019). Manager Sentiment and Stock Returns[J]. *Journal of Financial Economics*, 2019, 123(1): 126–149.
- Kumar A, and Lee C M C, (2006). Retail Investor Sentiment and Return Comovements[J]. *Journal of Finance*, 61(5): 2451–2486.

- Lee C M C, Shleifer A, Thaler R H, (1991). Investor Sentiment and the Closed-End Fund Puzzle[J]. *The Journal of Finance*, 1991, 46(1): 75–109.
- Lee E, (2013). *Individual Stock Investor Sentiment, Stock Issuance, and Financial Market Anomalies*[D]. Philadelphia: University of Pennsylvania.
- Liu J, Stambaugh R F, Yuan Y, (2019). Size and Value in China[J]. *Journal of Financial Economics*, 134(1): 48–69.
- Ljungqvist A, Wilhelm W J, (2003). IPO Pricing in the Dot-com Bubble[J]. *Journal of Finance*, 58(2): 723–752.
- Loughran T, McDonald B, (2016). Textual Analysis in Accounting and Finance: A Survey[J]. *Journal of Accounting Research*, 54(4): 1187–1230.
- Manela A, Moreira A, (2017). News Implied Volatility and Disaster Concerns[J]. *Journal of Financial Economics*, 123(1): 137–162.
- Michaelides A, Milidonis A, Nishiotis G P, Papakyriakou P, (2015). The Adverse Effects of Systematic Leakage ahead of Official Sovereign Debt Rating Announcements[J]. *Journal of Financial Economics*, 116(3): 526–547.
- Michaelides A, Milidonis A, Nishiotis G P, (2019). Private Information in Currency Markets[J]. *Journal of Financial Economics*, 31(3): 643–665.
- Price S M, Doran J S, Peterson D R, Bliss B A, (2012). Earnings Conference Calls and Stock Returns: The Incremental Informativeness of Textual Tone[J]. *Journal of Banking & Finance*, 36(4): 992–1011.
- Renault T, (2017). Intraday Online Investor Sentiment and Return Patterns in the US Stock Market[J]. *Journal of Banking & Finance*, 84: 25–40.
- Roll R, (1984). Orange Juice and Weather[J]. *The American Economic Review*, 74(5): 861–880.
- Scheinkman J A, Xiong W, (2003). Overconfidence and Speculative Bubbles[J]. *Journal of Political Economy*, 111(6): 1183–1219.
- Seybert N, Yang H I, (2012). The Party's Over: The Role of Earnings Guidance in Resolving Sentiment-driven Overvaluation[J]. *Management Science*, 58(2): 308–319.
- Shen J, Yu J, Zhao S, (2017). Investor Sentiment and Economic Forces[J]. *Journal of Monetary Economics*, 86: 1–21.
- Stambaugh R F, Yu J, Yuan Y, (2012). The Short of It: Investor Sentiment and Anomalies[J]. *Journal of Financial Economics*, 104(2): 288–302.
- Swaminathan B, (1996). Time-varying Expected Small Firm Returns and Closed-End Fund Discounts[J]. *Review of Financial Studies*, 9(3): 845–887.
- Tetlock P C, (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market[J]. *The Journal of Finance*, 62(3): 1139–1168.
- Tetlock P C, Saar-Tsechansky M, Macskassy S, (2008). More than Words: Quantifying Language to Measure Firms' Fundamentals[J]. *The Journal of Finance*, 63(3): 1437–1467.
- Tetlock P C. (2010). Does Public Financial News Resolve Asymmetric Information?[J]. *The Review of Financial Studies*, 23(9): 3520–3557.
- Tetlock P C. (2011). All the News that's Fit to Reprint: Do Investors React to Stale Information?[J]. *The Review of Financial Studies*, 24(5): 1481–1512.
- Yu J, Yuan Y, (2011). Investor Sentiment and the Mean-variance Relation[J]. *Journal of Financial Economics*, 100(2): 367–381.