投资者情绪对股票收益的影响

——来自股吧评论的分析

刘 倩

(兰州财经大学统计与数据科学学院,兰州 730000)

摘要:以沪深300指数的主要成分股为研究对象,探讨了投资者情绪对股票市场的具体影响。首先提出了四种不同的投资者情绪因子构建方法,并通过分析东方财富股吧的情绪信息,确定了最有效的情绪因子构造方法。接着,将构建的投资者情绪指数与传统的金融指标相结合,运用Fama-French四因子模型和随机森林模型对情绪因子对股票收益率的影响进行了深入分析。结果表明,投资者情绪因子对股票收益率的变化具有显著的预测能力。

关键词:投资者情绪;沪深300指数;Fama-French四因子模型;随机森林;股票收益率

中图分类号:F832.5 **文献标志码**:A **文章编号**:1671-1807(2024)24-0203-07

投资者情绪对股票收益具有重要影响,在股票市场中备受关注。投资者情绪能够影响市场参与者的决策行为,从而影响股票价格的波动。因此,深入理解投资者情绪变化有助于投资者更准确地预测市场走势,提高投资决策的有效性。

情绪金融学研究了情绪对金融决策和市场行为的影响。其中挖掘情绪金融学的理论基础是研究人类情绪如何影响金融市场中的决策和行为,理解市场动态和投资行为的关键。Barberis和Thaler[1]全面总结行为金融学的主要观点和理论,为理解情绪金融学提供背景。Baker和Wurgler[2]探讨投资者情绪对股票收益横截面的影响,为情绪金融学的实证研究提供了重要的支持。刘戈和杨大进[3]在经济研究中对情绪金融学进行了进一步的综述,突出了研究的局限性和未来的发展方向。由情绪金融学在国内的市场研究也发现,情绪对股票市场稳定和投资者行为有着重要影响。刘豪斐[4]在我国股票市场中投资者的心理因素对股票价格和市场波动的影响研究中,突出了情绪金融学在解释投资者情绪如何影响金融决策和市场波动方面的重要性,为投资者提供理性投资和政策制定提供参考价值。姚加权等[5]指出,情绪金融学领域正逐渐从传统的结构化数据分析转向非结构化的文本大数据分析,如新闻、社交媒体和搜索等,以此来探究和解读市场情绪与行为。这一转变预示着情绪金融学研究将在未来更加注重利用文本分析技术来深入理解市场动态。

为深入理解金融市场中情绪的影响,情绪指数的构建至关重要。它通过量化投资者情绪来提供对市场行为的洞察,提供可操作的指标。例如,Bollen等[6]使用Twitter上的情感词汇来构建情绪指数,成功预测股票市场的波动。而Bollen等[7]不仅考虑了情感词汇,还关注了社交网络中的“幸福”传播,为构建更全面的情绪指数提供了一种方法。Curme等[8]通过分析搜索引擎查询的语义内容,构建情绪指数,以预测股票市场的走势。在研究中国股市投资者情绪的连续动态测度方面,王德青等[9]提出了创新的函数型投资者情绪指数,为情绪金融学研究奠定了基础。郭文伟等[10]通过分析股吧股评,创建了新型金融情感词典来测度投资者情绪。在此基础上,刘广和易鸿[11]通过分析金融新闻的标题文本,并结合词嵌入模型和机器学习技术等文本分析方法,构建了考虑新闻累积效应的投资者时闻累积情绪指数,以更精确地追踪投资者情绪的变化。

情绪金融学领域的一个重要研究方向是投资者情绪与股票市场行为之间的关联性研究,该研究旨在揭示情绪对市场的影响机制。张智勇等[12]通过中国股市的经验证据表明,投资者情绪显著影响股票市场的波动。同时大量研究结果表明,情绪指标可用于预测市场的短期变化,例如,Brown和

收稿日期:2024-07-11

作者简介:刘倩(2000—),女,安徽安庆人,硕士研究生,研究方向为自然语言处理。

Cliff[13]在投资者情绪对股票市场短期波动的影响研究;谢世清和唐思勋[14]指出了情绪因素在短期内对市场的影响更为显著;Baker 和 Wurgler[15]发现情绪波动与股票市场的波动存在明显的正相关关系;关筱谨等[16]揭示投资者情绪和媒体关注度对股市波动的影响,指出它们之间的相互作用会放大市场波动,并在不同市场环境下表现出不同的影响力度;曹雄飞等[17]指出在新冠疫情期间,上海板块指数的投资者情绪与市场走势之间存在显著的负相关关系,情绪变化对股市收益率和波动率有显著影响。

综上所述,投资者情绪是股票市场波动的关键驱动因素,它对市场走势和投资者的决策行为产生深远影响。情绪金融学,作为探索情绪如何塑造金融市场决策和行为的学科,为理解这一复杂关系提供了坚实的理论基础,并在实际的市场分析和投资策略中展现了其应用价值。随着大数据和机器学习技术的进步,对投资者情绪的量化分析变得更为精确和细致,这不仅有助于更准确地预测市场动态,而且为引导理性投资和政策制定提供了有力的工具。

基于上述文献,本文提出四种不同的情绪因子构造方法。本文的创新之处在于,比较四种不同的情绪因子构建方法,筛选因子构建的最佳方法,更加精确地捕捉和度量投资者情绪的多维度特性;然后对于情绪指标和技术指标,采用 Fama-French 四因子和随机森林两种不同的模型进行了投资者情绪因子与股票收益率的关联性研究,得出一致结论,强调了情绪在解释股票市场行为中的重要性。

2 研究方法

2.1 情绪指数的构建方法

现有文献[18-21]的基础上,总结提出以下四种情绪指数的构建方法,如表1所示。

表1 情绪指数构建方法

构建方法	变量解释
$MS_t=\frac{M_t^{pos}-M_t^{neg}}{M_t^{pos}+M_t^{neg}}$	M_t^{pos}、M_t^{neg} 分别为 t 时期内积极情绪和消极情绪的用户帖子数量总和
$MS_t=\frac{\ln(1+P_t)}{1+N_t}$	P_t、N_t 分别为 t 日的积极类股评和消极类股评个数
$MS_t=\ln\frac{1+Senti_{t,pos}}{1+Senti_{t,neg}}$	$Senti_{t,pos}$、$Senti_{t,neg}$ 分别为第 t 天的积极词汇和消极词汇总量
$MS_t=\frac{M_t^P-M_t^C}{M_t^P+M_t^C}$	M_t^P、M_t^C 分别表示 t 时期内积极情绪和消极情绪帖子的总评论量

注:MS为特定情绪数量对情绪指数的影响。

2.2 Fama-French 因子模型

2.2.1 Fama-French 三因子模型

除了投资者情绪因子,股票市场还存在着许多影响股票收益率的金融指标,如股票市值、账面市值、市盈率等[22]。Fama 和 French[22]选取多个可能影响股票收益的金融指标进行研究,并通过 Fama-MacBeth 多变量回归模型比较这些指标的拟合效果,拟合效果越好说明给指标对股票收益率的解释能力越强。最终将投资组合的超额收益率用3个因子表示,分别为市场资产组合因子、规模因子和账面市值比因子。Liu 等[23]按照相同的思路构建了中国版的三因子模型,研究结果显示,相对于 Fama 和 French[22]在美国使用的账面市值比(BM)、市盈率(EP)能更好地捕捉中国的价值效应。本文结合上述的 Fama 和 French[22]三因子模型思想最终选取市场价值、市盈率和市场组合收益率构建影响股票收益的3个影响因子——规模因子、价值因子和市场资产组合因子。

规模因子(small minus big,SMB)衡量小市值股票和大市值股票之间的收益差异,当研究对象为多只股票的组合收益率时,可以按照以下步骤计算SMB:首先将所研究的股票按照日个股流通市值的大小平均分为两组,即将股票划分为大市值组和小市值组,再对每组股票再按照市盈率由大到小进行分组,一般将市盈率分为3组,即前30%、中间40%和后30%。最终形成6个组合,分别为大市值高市盈率组(SH)、大市值中市盈率组(SM)、大市值低市盈率组(SL)、小市值高市盈率组(BH)、小市值中市盈率组(BM)和小市值低市盈率组(BL)。

为了度量大规模投资组合和小规模投资组合市值的差异,可以计算 SMB,即计算小市值组合收益率的平均值减去大市值组合收益率的平均值。通过这种方式,可以量化小市值股票和大市值股票之间的收益差异,并用 SMB 作为衡量指标。计算公式为

$$SMB=\frac{SH+SM+SL}{3}-\frac{BH+BM+BL}{3} \quad (1)$$

式中:SH 为每日个股流通市值的前50%且市盈率前30%股票的市值加权收益率;SM 为每日个股流通市值的前50%且市盈率居于中间的40%股票的市值加权收益率;SL 为每日个股流通市值的前50%且市盈率后30%股票的市值加权收益率;BH 为每日个股流通市值的后50%且市盈率前30%股票的市值加权收益率;BM 为每日个股流通市值的后

表 2　因子构建分组标准

每日个股流通市值	市盈率		
	前 30%	中 49%	后 30%
大 50%	SH	SM	SL
小 50%	BH	BM	BL

50%且市盈率居于中间的 40%股票的市值加权收益率;BL 为每日个股流通市值的后 50%且市盈率后 30%股票的市值加权收益率。

价值因子(value minus growth,VMG)用于衡量价值股和成长股之间的收益差异,即上市公司因市盈率不同导致的收益率差异。当研究对象为多只股票的组合收益率时,VMG 的计算方法与 SMB 类似。首先根据上述表格,将股票按市盈率的大小平均分为高市盈率组和低市盈率组,再分别计算高市盈率组合和低市盈率组合的市值加权收益率,最后计算低市盈率组合的市值加权收益率减去高市盈率组合的市值加权收益率。通过计算得到的 VMG 值,可以衡量价值股和成长股之间的收益差异。计算公式为

$$\mathrm{VMG}=\frac{\mathrm{SH}+\mathrm{BH}}{2}-\frac{\mathrm{SL}+\mathrm{BL}}{2} \tag{2}$$

市场资产组合因子(market,MKT)反映了整个市场股票收益的情况。它通常以市场收益率为度量标准,代表了整个股票市场的涨跌情况。在本文中,利用市场收益率与无风险收益率的差异来表示 MKT,即超额收益率,这表示个股相对于无风险资产的额外收益。无风险收益率是指中国人民银行公布的人民币三个月整存整取利率,将其除以 365 天得到每天的平均收益率。MKT 的计算公式为

$$\mathrm{MKT}=R_{\mathrm{m}}-R_{\mathrm{f}} \tag{3}$$

构建 Fama-French 三因子模型:将投资组合的超额收益率用三个因子表示,分别为市场资产组合因子 MKT、规模因子 SMB 和价值因子 VMG。模型表达式为

$$R_t-R_{\mathrm{f}}=\beta_0+\beta_1\mathrm{MKT}_t+\beta_2\mathrm{SMB}_t+\beta_3\mathrm{VMG}_t+\varepsilon_t \tag{4}$$

式中:R_t 为 t 日收盘时股票的收益率;R_{f} 为无风险收益率;R_{m} 为市场收益率;MKT_t、SMB_t、VMG_t 分别为 t 日开盘时市场资产组合因子、规模因子、价值因子的值;$\beta_i(i=0,1,2,3)$ 为待估参数;ε_t 为误差项。

2.2.2　Fama-French 四因子模型

为研究投资者情绪对股票收益率的影响,在上述三因子模型中加入基于股评信息所构建的投资者情绪因子 MS_t,从而构建 Fama-French 四因子模型如式(5)所示:

$$R_t-R_{\mathrm{f}}=\beta_0+\beta_1\mathrm{MKT}_t+\beta_2\mathrm{SMB}_t+\beta_3\mathrm{HML}_t+\mathrm{MS}_t+\varepsilon_t \tag{5}$$

式中:MS_t 为股票在 t 日基于股评信息构建的投资者情绪因子。

3　实证分析

为深入探讨投资者情绪对股票市场的影响,并通过构建有效的情绪指数来评估其影响,以沪深 300 部分成分股作为研究对象进行分析。首先收集沪深 300 部分成分股的股评信息,利用基于情感词典标注的投资者情绪倾向,在上述提出的四种情绪指数构建方法中筛选最有效的情绪指数;然后采用 Fama-French 四因子模型和随机森林模型来研究投资者情绪对股票收益率的影响。

3.1　基于股评信息构建投资者情绪指数

3.1.1　数据来源

选取东方财富网股吧作为文本信息的来源,以沪深 300 部分成分股作为研究对象,对沪深 300 权重靠前的 10 只成分股在东方财富股吧中的股评信息进行文本挖掘,并选取不同指数构建方法构建投资者情绪指数进行对比分析。其中,沪深 300 成分股的权重信息来源于国泰安数据库,表 3 给出沪深 300 权重靠前的 10 只成分股。

表 3　沪深 300 权重排名前 10 只股票

股标代码	股票名称	权重/%
600519	贵州茅台	6.010
601318	中国平安	2.797
300750	宁德时代	2.781
600036	招商银行	2.162
000858	五 粮 液	1.720
000333	美的集团	1.477
600900	长江电力	1.255
601166	兴业银行	1.246
002594	比亚迪	1.228
600276	恒瑞医药	1.216

3.1.2　文本数据收集和预处理

(1)数据收集。通过八爪鱼采集软件爬取东方财富股吧中沪深 300 权重靠前的 10 只成分股自 2022 年 3 月 1 日至 2023 年 2 月 28 日的发帖数据。每条帖子发帖内容、发布日期、阅读量、评论量、作者等字段。

(2)数据预处理。对于通过爬虫获取到的文本数据,可以进行以下预处理操作。①删除重复数据和无用数据:通过比较文本数据的内容,去除重复的数据和无意义的贴子,以确保数据的准确性和可

靠性；②处理缺失数据和非交易日数据：对于存在缺失数据或非交易日的数据，可以选择填充缺失值或删除非交易日数据，以保持数据的完整性和一致性；③中文分词：使用Python语言中的jieba库进行中文分词，将文本数据按照词语进行切分，便于后续的文本分析和处理；④去停用词：为提高文本分析的准确性，可以使用中文停用词表对分词结果进行去除停用词的操作。除了常见的中文停用词表，还可以根据具体需求，根据10只成分股文本数据集的特点，加入一些特定的停用词，构建适用于本文的较为完备的中文停用词表。

通过以上预处理操作，可以提高爬取到的文本数据的质量，减少无用信息的干扰，为后续的文本分析和挖掘提供更准确和可靠的数据基础。

3.1.3 基于词典的股吧信息分类

使用姚加权等[24]构建的适用于金融领域研究的中文情绪词典，对经过预处理后的数据集进行情感倾向标注，并按照三分类原则将情感倾向分为三类，分别为正面情绪、中性情绪和负面情绪，同时将正面情绪标记为1，中性情绪标记为0，负面情绪标记为－1[19]。使用的中文情感词典的部分情感词库如表5所示。

表4 中文情感词典部分情感词库

类型	部分词语
积极	安全可靠、拔尖、保持稳定、充分、大幅提高、飞速发展、高效、极好、叫好、名列前茅
消极	弊端、不平衡、差强人意、倒下、高消耗、积压、急跌、落后、大势已去、滞销

3.1.4 指数构建方法对比研究

经过预处理和股吧信息分类后的非结构化数据集，利用上述构建的四个投资者情绪指数方法，相对应得到四个随时间变化的投资者情绪序列。

为比较四种构建方法的优劣，对10只股票收益率以其在沪深300中的相对权重进行加权得到10只股票的加权收益率序列，其中10只股票的收益率数据在国泰安数据库获取，分别计算这10只股票的加权收益率序列与上述的投资者情绪序列之间的皮尔逊相关系数r的大小，r越大，所构建的情绪指数与收益率之间的线性相关性越强。四种情绪指数与股票收益率之间的相关系数如表5所示。

结果表明，在本文所提出的四种因子构建方法中，由方法(a)构建的情绪因子与股票收益率之间的相关系数最大，故方法(a)效果最佳，后续研究将选用此方法所构建的情绪指数进行分析。

表5 四种情绪指数与股票收益率之间的相关系数汇总

方法	相关系数(r)
方法(a)	0.667 5
方法(b)	0.662 2
方法(c)	0.654 7
方法(d)	0.190 4

当相关系数为0.668时，表明投资者情绪指数与股票收益率序列之间存在较强的正相关关系，这意味着当投资者情绪较高时，股票收益率往往也会较高，反之亦然。换句话说，投资者情绪的波动很可能会引起股票市场收益率的波动。

3.2 投资者情绪对股票收益率的影响

上述分析中，由于相关系数接近1，可以合理地推断该投资者情绪指数具有一定的预测股票收益率的能力。这种情绪指数可以被视为股票市场的一个重要因子，可能在一定程度上预示着未来股票收益的走势。

然而，单个相关系数并不能说明一切，因为股票市场受到多种因素的影响。为了更全面地评估这种情绪指数作为预测因子的有效性，还需要进一步的研究来确认这种关系的稳定性和可靠性。同时，结合其他经典的预测模型和因子进行综合分析，可以更好地理解股票收益率背后的机制和影响因素。因此采用Fama-French四因子模型和随机森林模型对投资者情绪和股票收益率之间的相互影响进行分析。

3.2.1 投资者情绪对股票收益率的线性影响

(1)模型设计。构建适用于研究对象的Fama-French三因子模型，即

$$R_t - R_f = \beta_0 + \beta_1 \mathrm{MKT}_t + \beta_2 \mathrm{SMB}_t + \beta_3 \mathrm{VMG}_t + \varepsilon_t \tag{6}$$

为研究投资者情绪对股票收益率的影响，在上述三因子模型中加入投资者情绪因子MS_t，构建Fama-French四因子模型，即

$$R_t - R_f = \beta_0 + \beta_1 \mathrm{MKT}_t + \beta_2 \mathrm{SMB}_t + \beta_3 \mathrm{HML}_t + \mathrm{MS}_t + \varepsilon_t \tag{7}$$

(2)变量选择。选定沪深300权重靠前的10只股票作为研究的投资组合，以沪深300指数作为市场投资组合，无风险收益采用中国人民银行公布的人民币3个月整存整取利率除以365，以沪深300权重靠前的10只股票的加权收益率-无风险收益率($R_t - R_f$)即综合风险溢价因子作为因变量，市场资产因子即沪深300指数日收益率-无风险收益率($\mathrm{MKT} = R_m - R_f$)、规模因子(SMB)、价值因子

(VMG)、情绪因子(MS)作为自变量,建立 Fama-French 因子模型。

首先计算模型中所需变量的值,从国泰安数据库中获取沪深 300 权重靠前的 10 只成分股的日收盘价、市盈率和日个股流通市值。由方法(a)构建这 10 只股票的投资者情绪因子,再依据各股票在沪深 300 指数成分股中所占的相对权重加权得到综合投资者情绪因子值 MS;由式(1)～式(3)得规模因子(SMB)、价值因子(VMG)和市场资产组合因子(MKT)的值;由沪深 300 权重靠前的 10 只股票的加权收益率-无风险收益率得到因变量的值。

(3)结果分析。在开始拟合模型之前,为了防止伪回归现象,首先对各个变量进行了平稳性检验。通过绘制时序图(图 1),初步评估了各个因子的时间序列数据是否平稳。接着,进行 ADF(augmented Dickey-Fuller test)检验,结果如表 6 所示。由表 6 可知,5 个因子的 P 都小于 0.01,这意味着都通过了 ADF 检验,表现出平稳性。

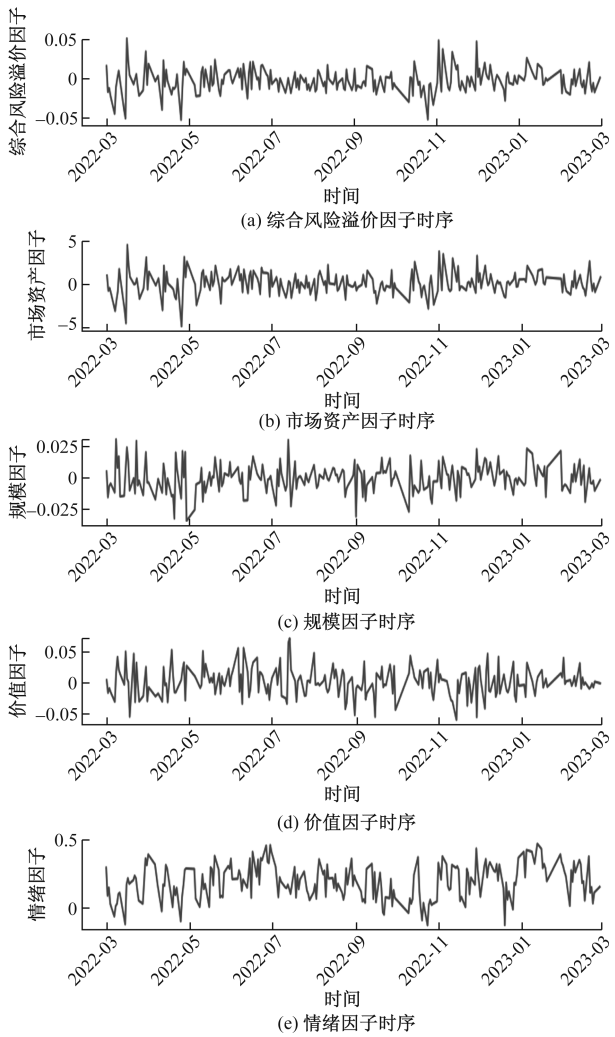

图 1　各变量时序

选取沪深 300 权重靠前的 10 只股票构成投资组合,并应用多元线性回归模型分别对式(4)和式(5)进行线性拟合,表 7 和表 8 展示了拟合结果。可以发现,两个回归方程的调整 R^2 和 F 统计量均显示出了统计显著性($P<0.05$),这说明回归方程在统计学上是有着显著意义的。当投资者情绪因子被纳入模型后,调整 R^2 从 0.893 提升至 0.903,这表明投资者情绪因子增强了模型对股票收益率的解释力度。这一提升表明,投资者情绪因子对股票收益率有着显著的线性作用。

表 6　ADF 检验结果

变量	ADF 检验统计量	P
$R-R_f$	−15.489 8	2.49×10^{-28}
MKT	−15.610 6	1.78×10^{-28}
SMB	−16.578 9	1.84×10^{-29}
VMG	−15.587 7	1.90×10^{-28}
MS	−6.578 5	7.62×10^{-9}

表 7　三因子与股票收益率的线性拟合结果

变量	系数	标准差	t	P	F 检验-P	调整 R^2
常数	−0.003	0.000	−7.744	0.000	6.07×10^{-14}	0.893
MKT	0.011	0.000	40.109	0.000		
SMB	0.219	0.031	6.964	0.000		
VMG	−0.061	0.014	−4.217	0.000		

表 8　四因子与股票收益率的线性拟合结果

变量	系数	标准差	t	P	F 检验-P	调整后 R^2
常数	−0.006	0.001	−8.191	0.000	8.85×10^{-18}	0.903
MKT	0.01	0.000	30491	0.000		
SMB	0.211	0.03	7.058	0.000		
VMG	−0.047	0.014	−3.357	0.001		
MS	0.016	0.003	5.067	0.000		

3.2.2　投资者情绪对股票收益率的非线性影响

上述结论表明投资者情绪因子对股票收益率具有一定的线性影响。为了进一步探究投资者情绪与股票收益率之间可能存在的非线性关系,下面将采用随机森林模型进行分析。

(1)参数选择。首先数据集按 8∶2的比例划分训练集和测试集,然后在随机森林中选择 $R-R_f$ 作为因变量,MKT、SMB、VMG 和 MS 作为自变量,随机森林中树的数量设定为 100,决策树的最大深度设定为 None,表示树会一直扩展,直到每个叶子节点包含的样本数小于'min_samples_split',决策树节点分裂所需的最小样本数设定为 2,表示任何节点都需要至少包含 2 个样本才能分裂,随机数种子设定为 42,用于确保每次运行代码时都能得到相同的随机结果。

(2)结果分析。在python中进行代码实现,结果显示,模型的均方误差数值很小,为 2.444×10^{-5},模型的决定系数为0.898,非常接近1,这说明模型在训练集上表现出色,能够很好地解释股票收益率的方差,具有较好的拟合效果。

模型各因子的特征重要性数组为[0.898 189 66 0.044 602 09 0.032 162 68 0.025 045 58]。特征重要性数组显示,情绪因子(MS)在模型中具有相对较高的重要性,虽然不及沪深300指数收益率(MKT),但仍对模型的预测有一定的贡献。考虑到随机森林模型的非线性特性,以及情绪因子对股票收益率的贡献,可以得出投资者情绪因子对股票收益率存在一定的非线性影响。这意味着投资者情绪的波动会以非线性的方式影响着股票市场的表现,而随机森林模型能够较好地捕捉这种复杂关系。

综上所述,通过两种模型的分析,无论是线性拟合还是非线性随机森林模型,都表明投资者情绪因子对股票收益率具有一定的解释作用,而且在加入情绪因子后模型的拟合效果都显著改善。这说明市场情绪对股票价格具有显著影响,并且情绪因子可能存在着一定的非线性影响,这种复杂关系随机森林模型能够较好地捕捉。综合两个模型的结果,可以进一步确认投资者情绪因子在解释股票市场行为中的重要性。

4 结论与建议

4.1 结论

结合以往文献的研究方法,本文首先提出四种不同的情绪因子构建方法,通过实证分析筛选出最能代表股票市场投资者情绪的指数构建方法,然后结合股票的金融指标数据,研究投资者情绪对股票收益率的影响。主要结论如下。

(1)方法(a)构建的情绪因子最具代表性,最能反映投资者的情绪状态。通过比较四种因子构建方法的相关系数的大小,发现方法(a)构建的投资者情绪因子与股票市场的相关系数最高,这表明通过方法(a)构建的投资者情绪因子在捕捉投资者情绪与股票市场变动之间的强关联性方面表现最为有效。

(2)投资者情绪因子对股票收益率具有一定解释作用。首先在使用Fama-French因子模型研究投资者情绪与股票市场收益率的线性影响关系时,通过在模型中加入投资者情绪因子,观察到可决系数的增加。这表明投资者情绪因子对股票收益率的变动具有一定的解释作用,也暗示了投资者情绪在解释股票市场收益率方面可能具有一定的线性关系;其次通过使用随机森林模型研究投资者情绪与股票市场收益率的非线性影响关系,关注模型中各因子的特征重要性占比,观察到投资者情绪因子在随机森林模型中具有一定的特征重要性,说明该因子对股票收益率的解释在模型中具有一定意义;最后通过上述两个模型的分析,表明投资者情绪因子对股票收益率具有一定的解释作用,进一步确认投资者情绪因子在解释股票市场行为中的重要性。

4.2 建议

基于上述结论,可以提出以下建议。

(1)情绪因子的选择:选择与股票市场相关性最高的情绪因子构建方法,有望提高模型的预测准确性,这对于投资决策和风险管理具有积极的影响。

(2)模型选择:研究投资者情绪影响股票收益率时,可以采用多种模型进行分析,包括传统的线性回归模型和非线性模型(如随机森林模型)。通过综合分析不同模型的结果,可以更全面地理解投资者情绪对股票市场的影响机制。

(3)风险管理:考虑到投资者情绪对股票市场的影响,投资者应当在制定投资策略时充分考虑市场情绪因素,以降低投资风险并提高投资收益。

(4)市场监测:监测投资者情绪指标可以作为一种辅助手段,帮助投资者更好地了解市场走势和预测未来市场变化,从而做出更明智的投资决策。

综上所述,研究投资者情绪对股票市场行为的重要性提示我们在投资决策和风险管理中应充分考虑市场情绪因素,以提高投资效益和降低风险。

参考文献

[1] BARBERIS N, THALER R. A survey of behavioral finance[J]. Handbook of the Economics of Finance, 2003(1): 1053-1128.

[2] BAKER M, WURGLER J. Investor sentiment and the cross-section of stock returns[J]. The Journal of Finance, 2006, 61(4): 1645-1680.

[3] 刘戈, 杨大进. 情绪金融学研究述评与展望. 经济研究, 2011(9): 43-54.

[4] 刘豪斐. 基于行为金融学的投资者情绪与中国股票收益关系的研究[D]. 济南: 山东大学, 2018.

[5] 姚加权, 张锟澎, 罗平. 金融学文本大数据挖掘方法与研究进展[J]. 经济学动态, 2020(4): 143-158.

[6] BOLLEN J, MAO H, ZENG X. Twitter mood predicts the stock market[J]. Journal of Computational Science, 2011, 2(1): 1-8.

[7] BOLLEN J, GONÇALVES B, RUAN G, et al. Happiness is assortative in online social networks[J]. Artificial Life, 2011, 17(3): 237-251.
[8] CURME C, PREIS T, STANLEY H E, et al. Quantifying the semantics of search behavior before stock market moves[J]. Proceedings of the National Academy of Sciences, 2014, 111(32): 11600-11605.
[9] 王德青, 田思华, 朱建平, 等. 中国股市投资者情绪指数的函数型构建方法研究[J]. 数理统计与管理, 2021, 40(1): 162-174.
[10] 郭文伟, 卢立晔, 钟袁妍. 投资者情绪如何影响股市泡沫——兼议对股市泡沫监管的建议[J]. 金融监管研究, 2024(1): 61-78.
[11] 刘广, 易鸿. 融合媒体信息和信号分解的股票市场深度学习预测[J]. 计算机科学, 2024, 51(S1): 1104-1115.
[12] 张智勇, 张毅, 李兴平. 投资者情绪对股票市场的影响: 来自中国股市的经验证据. 经济研究, 2011(10): 71-82.
[13] BROWN G W, CLIFF M T. Investor sentiment and the near-term stock market[J]. Journal of Empirical Finance, 2004, 11(1): 1-27.
[14] 谢世清, 唐思勋. 投资者情绪与宏观经济波动对股票市场收益率的影响[J]. 宏观经济研究, 2021(2): 99-107.
[15] BAKER M, WURGLER J. Investor sentiment in the stock market[J]. Journal of Economic Perspectives, 2007, 21(2): 129-151.
[16] 关筱谨, 张骏, 刘彦迪. 媒体关注度、投资者情绪与股票市场波动[J]. 统计与决策, 2022, 38(24): 143-148.
[17] 曹雄飞, 杨双会, 范炜彬. 基于文本挖掘的疫情情绪对股票市场的影响[J]. 金融理论与教学, 2023(5): 44-52.
[18] JING N, WU Z, WANG H. A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction[J]. Expert Systems with Applications, 2021, 178: 115019.
[19] 孟雪, 李艳茹, 胡锋. 基于股票评论的投资者情绪对持有股票收益影响的研究[J]. 数学的实践与认识, 2022, 52(8): 78-96.
[20] ZHANG W, GONG X, WANG C, et al. Predicting stock market volatility based on textual sentiment: a nonlinear analysis[J]. Journal of Forecasting, 2021, 40(8): 1479-1500.
[21] 樊鹏英, 杨音, 张正平, 等. 个股投资者情绪与股票收益率的关系——基于股评信息视角的研究[J]. 数学的实践与认识, 2021, 51(16): 305-320.
[22] FAMA E F, FRENCHK R. Common risk factors in the returns on stocks and bonds[J]. Journal of Financial Economics, 1993, 33(1): 3-56.
[23] LIU J, STAMBAUGH R F, YUAN Y. Size and value in China[J]. Journal of Financial Economics, 2019, 134(1): 48-69.
[24] 姚加权, 冯绪, 王赞钧, 等. 语调、情绪及市场影响: 基于金融情绪词典[J]. 管理科学学报, 2021, 24(5): 26-46.

Impact of Investor Sentiment on Stock Returns: An Analysis from Stock Bar Comments

LIU Qian
(School of Statistics and Data Science, Lanzhou University of Finance and Economics, Lanzhou 730000, China)

Abstract: Taking the main component stocks of the CSI 300 index as the research object, the specific impact of investor sentiment on the stock market was explored. Firstly, four different methods for constructing investor sentiment factors were proposed, and the most effective method by analyzing the sentiment information from the East Money stock bar was determined. Next, the constructed investor sentiment index was combined with traditional financial indicators, and the Fama-French four-factor model and random forest model were used to conduct in-depth analysis of the impact of sentiment factors on stock returns. The results show that investor sentiment factors have a significant predictive ability for changes in stock returns.

Keywords: investor sentiment; CSI 300 index; Fama-French four-factor model; random forest; stock returns