第 7 卷◆第 10 期◆版本 1.0◆2024 年
文章类型：论文|刊号（ISSN）：2630-4759 /（中图刊号）：270GL018

# 投资者情绪对创业板指数收益率的影响研究

江海潮　刘娅琦
湖南科技大学
DOI:10.12238/ej.v7i10.1993

[摘　要] 本文基于情绪-表达范式,利用Python挖掘东方财富股吧股评文本,构建情绪-表达范式的创业板投资者情绪指数,回归发现投资者情绪对创业板指数收益率的影响具有跨期非对称性、情绪类型异质性与时期动态性,而且投资者情绪与市场因素、指数收益率滞后因素一起,非对称影响创业板指数收益率。有效控制投资者情绪对创业板市场收益率的情绪风险溢价,政府部门应高度关注多样化的新兴网络文本情绪的调控引导,构建自由、健康、法治、高效与和谐的文本情绪生态环境,有效疏导投资者情绪,科学引导投资者理性投资,促进创业板市场平稳有序发展。
[关键词] 投资者情绪；创业板指数；文本挖掘；情绪VAR模型
**中图分类号**：F830.59　**文献标识码**：A

## Research on the Influence of Investor Sentiment on Growth Enterprise Market Index Return

Haichao Jiang　Yaqi Liu
Hunan University of Science and Technology

[Abstract] This article is based on the emotion-expression paradigm, using Python to mine the text of stock reviews from the Oriental Wealth Stock Bar to construct an investor sentiment index for the GEM market according to the emotion-expression paradigm. Regression analysis reveals that investor sentiment has a cross-period asymmetry, emotional type heterogeneity, and period dynamism in its impact on the return rate of the GEM index. Moreover, investor sentiment, together with market factors and lagged index return factors, asymmetrically affects the return rate of the GEM index. To effectively control the emotional risk premium of investor sentiment on the GEM market return, the government should pay close attention to the regulation and guidance of diverse emerging network text emotions, construct a free, healthy, legal, efficient, and harmonious text emotional ecological environment, effectively guide investor sentiment, scientifically guide rational investment by investors, and promote the stable and orderly development of the GEM market.
[Key words] investor sentiment; GEM Index;text mining;sentiment VAR model

## 引言

研究投资者情绪对金融资产收益的影响效应,就涉及投资者情绪测度问题,现有文献主要采用三种范式方法测度投资者情绪。一是心理-情绪范式方法。不少研究沿着不同技术方法建立起央视看盘指数(杨阳和万迪昉,2010)[1]、好谈指数(陈其安,2017)[2]、投资者或消费者信心指数(Ciner,2014)[3]等,利用直接调研数据资料,测度投资者情绪。二是情绪-行为范式方法。学者们一方面采用金融市场封闭式基金折价率(靖荣华,2016)[4]、融资融券余额(郭洋,2017)[5]、市盈率(张国胜,2021)[6]等单一行动指标测度投资者情绪；另一方面采用封闭式基金折价率、市场换手率、股市成交量等,构建多维度行为指标体系,以主成分方法测度投资者情绪(Baker & Wurgler,2006；田云珠等,2023)[7-8]。三是情绪-表达范式方法。目前,学者们大多采用词典方法与机器学习技术,进行文本情绪挖掘测度投资者情绪,得到金融文本情绪(姜富伟等,2021)[9]、新闻媒体情绪(张宗新和吴钊颖,2021)[10]、社交媒体情绪(Gao,etal,2020)[11]等,开展多样化文本情绪测度。此外,还有学者利用深度学习技术对文本数据进行情感倾向分类,构建投资者情绪指数,进行情绪测度(张宗新和吴钊颖,2021)[10],把投资者情绪测度引向深入,极大地丰富了投资者情绪测度研究。

值得注意的是,现有研究大多关注投资者情绪对股票价格或股票收益的情绪溢价现象,少有研究关注投资者情绪对创业板股票收益的影响效应,杨晓兰等(2016)[12]发现东方财富网股吧文本情绪对创业板股市收益率有显著影响,李斐(2022)[13]基于情绪-行为范式方法,构建投资者信心指数,发现投资者信心情绪与创业板指数收益率存在滞后相互作用,并且创业板指数

Copyright © This work is licensed under a Commons Attibution-Non Commercial 4.0 International License.

第 7 卷◆第 10 期◆版本 1.0◆2024 年
文章类型：论文|刊号（ISSN）：2630-4759 /（中图刊号）：270GL018

收益率对投资者信心情绪影响更大,但情绪-表达范式的投资者情绪对创业板指数收益率的影响效应问题很少有关注。所以投资者文本情绪对创业板指数收益率的影响作用如何,值得深入研究。

本文借鉴已有研究成果,以中国创业板指数为研究对象,采用情绪-表达范式方法,挖掘东方财富网股吧文本,构建投资者网吧文本情绪指数,讨论投资者情绪对创业板指数收益率的影响,以确定中国创业板指数收益的投资者情绪风险溢价效应,检验投资者对东方财富吧等网络文本信息响应的理性水平,为投资者投资创业板指数决策,提供有益的参考。

## 1 投资者情绪测度

1.1测度方法与样本数据。本文采用情绪-表达范式方法测度投资者情绪,借鉴姜富伟(2021)[9]做法,基于情感词典的情感分析技术,构建投资者情绪综合指数,具体做法是从中文通用情感词典中筛选“积极”词性和“消极”词性的词语,构建投资者情绪词典,用Python软件爬取文本中的情绪词典用词语情况,进而评价投资者情绪。

为测度投资者情绪,根据姜富伟(2021)[9]中文金融情感词典,本文使用Python软件进行代码编写,对东方财富创业板指股吧文本内容进行提取,样本文本包括东方财富创业板指股吧2022年1月5日至2024年3月20日的股吧评论,本文选取的时间段内通过删除无效评论以及剔除节假日,总共有530天有效交易日文本,共提取260393条文本数据,提取内容包括评论标题和正文、时间和阅读数,操作时对东方财富创业板指股吧的每一条正文分别提取“积极”词性和“消极”词性的词语数,以此测度投资者情绪。

1.2投资者情绪水平。为了测度投资者情绪水平,本文参照Antweiler和Frank(2004)[14]的做法,计算东方财富创业板指股吧530天中每条评论的情感倾向。测度投资者股吧文本情绪,一是要确定文本中积极与消极情绪词语数量；二是要确定不同情绪词语的权重。由于股吧帖子的投资者意见文本阅读量能够传播,而不能通过评论量传播(熊熊等,2017)[15],所以本文引入阅读数作为情感倾向加权。具体计算公式如下：

$$W_{it}=\frac{D_{it}}{D_t} \tag{1}$$

$$IS_{it}=W_{it}*\frac{P_{it}-N_{it}}{P_{it}+N_{it}} \tag{2}$$

$$IS_t=\sum_{i=1}^{n}IS_{it}(\mathrm{n}=1,2,3,4,...) \tag{3}$$

其中,$W_{it}$ 为权数,$D_{it}$ 为第t日第i条评论的阅读数,$D_t$ 为第t日总阅读数,$IS_{it}$ 为第t日第i条评论的情绪指数,$P_{it}$ 为第t日第i条评论的积极词语数,$N_{it}$ 为第t日第i条的消极词语数,$IS_t$ 为测度的情绪水平。

## 2 实证研究

2.1实证模型。投资者情绪是影响创业板指数收益率的重要因素(李媛和刘芳,2023)[16],而且由于市场信息非对称性,创业板指数收益率对投资者情绪等信息的吸收具有时滞性,为确定投资者情绪对创业板指数收益率的影响,本文参照带有情绪因素的罗斯多因素模型,建立如下向量自回归模型(VAR),回归出投资者情绪对创业板指数收益率的贝塔系数,从而发现其影响规律。回归基准模型如下：

$$R_t=a_0+\sum_{i=0}^{g}\beta_{si}IS_{t-i}+\sum_{j=1}^{k}a_jR_{t-j}+bX+\varepsilon_t \tag{4}$$

其中 $R_t$ 为 $t$ 期创业板指数收益率、$IS_{t-i}$ 为 $t-i$ 期投资者情绪、$X$ 为控制变量向量、$R_{t-j}$ 为滞后 $j$ 期创业板指数收益率。此外,模型中 $a_0$ 为常数,$\beta_{si}$ 为情绪贝塔系数,$a_j$ 为滞后期回归系数,$b$ 为控制变量回归系数向量,$\varepsilon_t$ 为回归残差。

2.2模型变量与数据来源。(1)被解释变量。被解释变量为创业板指数收益率。具体采用东方财富Choice金融终端数据,获得创业板指数(399006)的每日收盘价,参考陈其安学者[2]计算创业板指数收益率数据：

$$R_t=ln(SZCEI_t)-ln(SZCEI_{t-1}) \tag{5}$$

$SZCEI_t$ 表示为在t日的创业板指数的每日收盘指数,$SZCEI_{t-1}$ 表示为在t-1日的创业板指数的每日收盘指数,$R_t$ 表示创业板指数的对数收益率。

(2)解释变量。解释变量 $IS_{t-i}$ 为投资者情绪指数,本文直接利用前面估计的投资者情绪综合指数数据。

(3)控制变量。控制变量参照魏卉(2023)[17]等学者的做法与罗斯多因素模型,回归时主要控制了市场层面因素,具体选择创业板指数换手率($TO_t$)和成交量($TAV_t$)两个控制变量,数据来源于东方财富Choice金融终端数据库。

2.3平稳性检验。在回归之前,对模型变量进行平稳性检验,防止出现“伪回归”现象以及保证变量的平稳性。本文将对创业板指收益率($R_t$)、投资者情绪($IS_t$)、换手率($TO_t$)和成交量($TAV_t$)进行平稳性检验,采用ADF方法对变量进行单位根检验。经检验发现,$R_t$、$IS_t$、$TO_t$ 以及 $TAV_t$ 均为平稳序列,能进行模型的构建。

2.4回归结果。采用VAR模型回归时,首先要确定最优滞后阶数,这里使用赤池信息准则(AIC)、施凡兹准则(SC)和最终预测误


Copyright © This work is licensed under a Commons Attibution-Non Commercial 4.0 International License.

差(FPE)方法确定最优滞后阶数,发现该模型最优滞后阶数为3阶,为此构建滞后期数为3的VAR模型。通过计算AR特征根的倒数发现,VAR(3)模型的均值均小于1(均落在单位圆内),说明模型的构建是稳定的,可进一步分析。VAR模型回归结果如方程(6)所示,根据回归结果,滞后一期和滞后两期的投资者情绪对创业板指数收益率产生非对称负向影响,滞后三期的投资者情绪对创业板指数收益率产生正向影响,说明投资者情绪对创业板指数收益率影响十分复杂,一是不同期投资者情绪对创业板指数收益率有显著影响；二是不同期的投资者情绪对创业板指数收益率的影响方向不一致,不同期情绪贝塔存在跨期对冲效应,影响具有跨期动态性。同时,创业板指数收益率存在三期非对称滞后效应,市场换手率与成交量也对创业板指数收益率也存在三期非对称作用效应,说明创业板指数收益率不仅受投资者情绪非对称影响,也受市场等因素非对称影响,这符合罗斯多因素模型预期。

$$R_t = 0.006 + 0.032R_{t-1} + 0.007R_{t-2} - 0.029R_{t-3} - 0.045IS_{t-1} - 0.096IS_{t-2} + 0.081IS_{t-3} - 0.743TO_{t-1} + 1.411TO_{t-2} - 0.855TO_{t-3} + 0.913TAV_{t-1} - 1.38TAV_{t-2} + 0.688TAV_{t-3} \quad (6)$$

2.5模型稳健性与异质性检验。

2.5.1稳健性检验。宏观经济政策变化会对市场投资者情绪进行回应,因此宏观政策一定程度地包含有投资者情绪信息,因而理论上可以利用若干宏观政策变化测度的投资者情绪,替代投资者股吧文本情绪,进行稳健性检验。为检验模型稳健性,本文选取期限利差($TSR_t$)[18],短期无风险利率($SBR_t$)[18]以及人民币兑美元汇率($EX_t$)[19]构建投资者宏观政策情绪指标,检验投资者情绪对创业板指数收益的影响效应。其中,期限利差($TSR_t$)选取十年期国债收益率与一年期国债收益率的差值,短期无风险利率($SBR_t$)选取SHIBOR隔夜拆借利率,人民币兑美元汇率($EX_t$)选取日度即期汇率,数据均来源于Choice金融终端。稳健性检验采取如下思路,首先回归三个宏观变量对创业板指数收益率的影响,发现短期无风险利率($SBR_t$)对创业板指数收益率影响并不显著,于是剔除变量短期无风险利率($SBR_t$)；然后再次进行回归,得到稳定方程；最后利用回归后得到的残差作为投资者宏观情绪指标,检验前文的结论。通过构建VAR(3)模型得到回归方程(7),可以发现,模型中的系数与上文结论一致,说明模型有较好的稳健性。

$$R_t = 0.006 + 0.033R_{t-1} + 0.008R_{t-2} - 0.028R_{t-3} - 0.043IS_{t-1} - 0.089IS_{t-2} + 0.075IS_{t-3} - 0.720TO_{t-1} + 1.425TO_{t-2} - 0.811TO_{t-3} + 0.889TAV_{t-1} - 1.392TAV_{t-2} + 0.642TAV_{t-3} \quad (7)$$

为确定投资者情绪影响创业板指数收益率的稳健性,对投资者情绪影响创业板指数收益率效应进行脉冲分析。投资者情绪冲击创业板指数收益率的效应见图1,不难发现,当投资者情绪施加一个正向冲击时,创业板指数收益率在第一期后产生负向响应,在第三期达到负向峰值后不断增强,在第四期变为正向效应并达到正向峰值,之后冲击效应快速回落收敛,效应大概持续八期。总的来说,投资者情绪对创业板指数收益率在短期内会形成波动性较小的冲击,但是长期会随着时间的推移而逐渐收敛,说明投资者情绪对创业板指数收益率的影响一是动态的,二是短期稳健与长期收敛的。

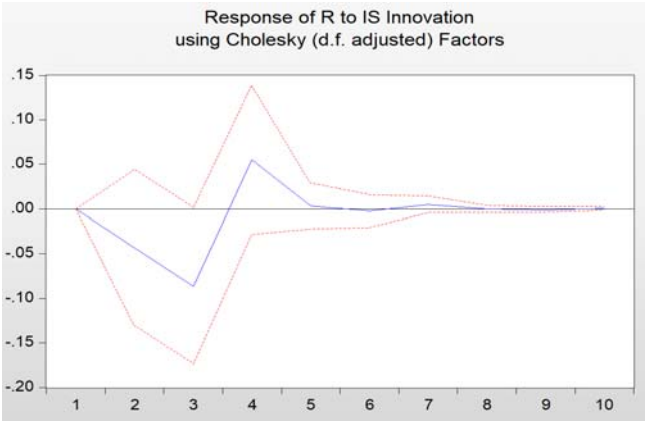


图1 投资者情绪变动对创业板指数收益率脉冲响应图

2.5.2异质性检验。理论上讲,投资者不同情绪对金融资产收益率的影响是非对称。为了探究不同情绪对创业板指数收益率的影响,本文将投资者情绪分为积极性情绪($PIS_{t-i}$)与消极性情绪($NIS_{t-i}$)来进行异质性检验。为检验投资者情绪影响异质性,首先建立VAR(3)模型,分别回归投资者积极性情绪和消极性情绪对创业板指数收益率的影响,发现两类情绪滞后第一期和滞后第二期均为负向影响,滞后第三期均为正向影响见公式(8)(9),而且投资者消极性情绪对创业板指数收益率影响的情绪贝塔系数大于积极性情绪对创业板指数收益率影响的情绪贝塔系数,说明投资者消极性情绪对创业板指数收益的情绪溢价大于积极性情绪的风险溢价,可见投资者情绪对创业板指数收益率的影响存在情绪异质性与时期异质性。其次,为确认投资者情绪影响创业板指数收益率的情绪异质性,回归检验同时加入投资者积极性情绪和消极性情绪见公式(10),发现投资者积极性情绪的滞后一期和滞后三期会对创业板投资者情绪产生负向影响,而积极性情绪滞后二期会对创业板投资者情绪产生正向影响,而消极投资者情绪在滞后一期和滞后三期会对创业板投资者情绪产生正向影响,在滞后二期会对创业板投资者情绪产生负向影响。这进一步说明,投资者情绪对创业板指数收益率的影响具有投资者情绪类型异质性,而且这种影响效应具有时期对冲性与动态性。

$$R_t = 0.007 + 0.034R_{t-1} + 0.008R_{t-2} - 0.025R_{t-3} - 0.049PIS_{t-1} - 0.088PIS_{t-2} + 0.074PIS_{t-3} - 0.704TO_{t-1} + 1.385TO_{t-2} - 0.883TO_{t-3} + 0.874TAV_{t-1} - 1.353TAV_{t-2} + 0.712TAV_{t-3} \quad (8)$$

Copyright © This work is licensed under a Commons Attibution-Non Commercial 4.0 International License.

第 7 卷◆第 10 期◆版本 1.0◆2024 年
文章类型：论文|刊号（ISSN）：2630-4759 /（中图刊号）：270GL018

$$R_t = -0.250 + 0.030\,R_{t-1} + 0.003\,R_t - 0.032\,R_{t-3} - 0.504\,NIS_{t-1} - 1.265\,NIS_{t-2} + 1.094\,NIS_{t-3} - 0.770\,TO_{t-1} + 1.429\,TO_{t-2} - 0.825\,TO_{t-3} + 0.938\,TAV_{t-1} - 1.395\,TAV_{t-2} + 0.659\,TAV_{t-3} \quad (9)$$

$$R_t = -0.119 + 0.035\,R_{t-1} + 0.002\,R_{t-2} - 0.031\,R_{t-3} - 0.048\,PIS_{t-1} + 0.062\,PIS_{t-2} - 0.048\,PIS_{t-3} + 0.017\,NIS_{t-1} - 1.966\,NIS_{t-2} + 1.620\,NIS_{t-3} - 0.752\,TO_{t-1} + 1.378\,TO_{t-2} - 0.814\,TO_{t-3} + 0.913\,TAV_{t-1} - 1.338\,TAV_{t-2} + 0.647\,TAV_{t-3} \quad (10)$$

为确定不同类型投资者情绪影响创业板指数收益率的异质性与动态收敛性，对回归模型进行脉冲动态分析，从图2可以发现投资者积极性情绪施加一个正向冲击时，创业板指数收益率呈现出先负向后正向的变动效应，最后作用效应逐渐趋向于零，影响具有较快的动态收敛性。从图3可以看出，消极投资者情绪变动不会给创业板收益率立刻带来影响，在第三期时会带来负向影响后又减弱，在第四期呈现正向影响后逐渐趋于零，影响同样具有动态收敛性。

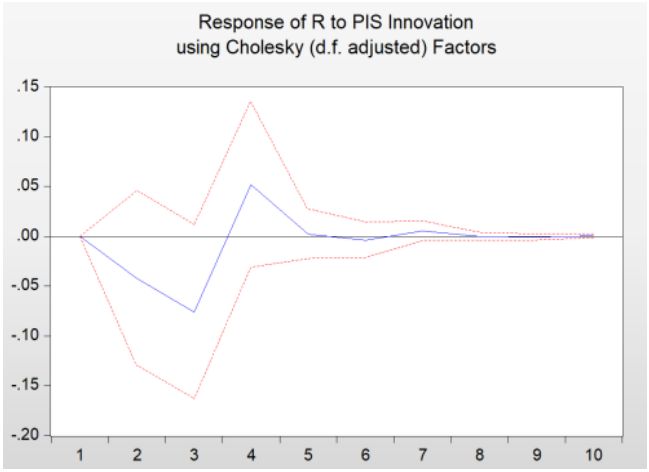


图2 投资者积极性情绪变动对创业板指数收益率脉冲响应图

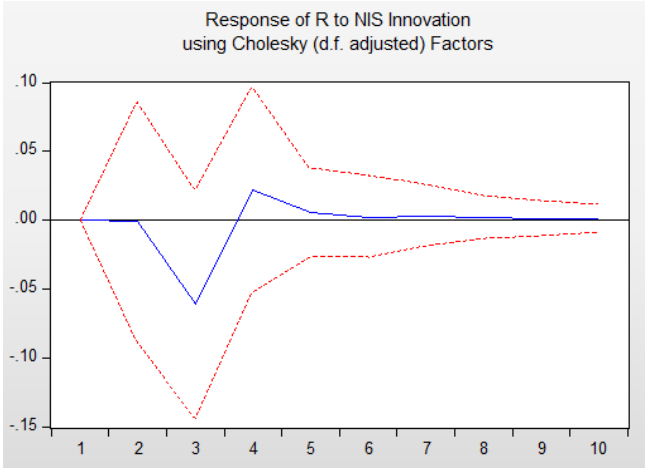


图3 投资者消极性情绪变动对创业板指数收益率脉冲响应图

总的来说，投资者积极性情绪和消极性情绪对创业板指数收益率的影响是非对称与动态收敛的，有着显著的情绪类型异质性。短期内积极性情绪变动会使创业板指数收益率呈现先负向后正向的变动，消极性情绪的变动会使创业板指数收益率呈现相反的影响，且长期内均具有收敛性。

## 3 结论

本文基于情绪-表达范式，利用Python文本挖掘方法，挖掘东方财富股吧2022年1月5日至2024年3月20日的创业板指数有效交易日股评文本，通过姜富伟情感词典进行词性分类，构建情绪-表达范式的创业板投资者情绪指数，建立情绪VAR模型，探讨投资者情绪对创业板指数收益率的影响效应，发现投资者情绪对创业板指数收益率产生负向冲击影响并具有时期收敛性与稳健性，异质性检验发现不同类型文本情绪对创业板指数收益率影响效应存在显著差异，消极性情绪对创业板指数收益率的影响大于积极性情绪的影响，说明投资者情绪对创业板指数收益率存在显著的、动态的情绪溢价效应，并且投资者情绪与市场因素、指数滞后因素一起，非对称影响创业板指数收益率。为有效控制投资者情绪对创业板市场收益率的情绪风险溢价，政府部门应合理注重引导投资者情绪，高度关注多样化文本情绪的承载与传递，特别要关注东方财富网股吧、网络文本等新兴文本情绪的调控引导，有效疏导投资者情绪，科学引导投资者理性投资，抑制非理性投资行为对创业板市场的影响，促进创业板市场的平稳有序发展。

## [参考文献]

[1]杨阳,万迪昉.不同市态下投资者情绪与股市收益、收益波动的异化现象——基于上证股市的实证分析[J].系统工程,2010,28(1):19-23.

[2]陈其安,雷小燕.货币政策、投资者情绪与中国股票市场波动性:理论与实证[J].中国管理科学,2017,25(11):1-11.

[3]Ciner,C.The Time Varying Relation Between Consumer Confidence and Equities[J].Journal of Behavioral Finance,2014,15(4):312-317.

[4]靖荣华,贺晓玲.我国投资者情绪指数的构建[J].当代经济,2016(34):43-47.

[5]郭洋.融资融券业务、投资者情绪以及股票价格[J].金融经济,2017(6):97-98.

[6]张国胜,林宇.结构突变下投资者情绪与股市收益间的非线性溢出效应研究[J].数理统计与管理,2021,40(1):148-161.

[7]Baker,M.,&Wurgler,J.Investor Sentiment and the Cross-section of Stock Returns[J].Journal of Finance,2006,61(4):1645-1680.

[8]田云珠,李亚文,王馨悦.投资者情绪与股票市场收益之间的关系[J].全国流通经济,2023(10):156-159.

[9]姜富伟,孟令超,唐国豪.媒体文本情绪与股票回报预测[J].经济学(季刊),2021,21(4):1323-1344.

[10]张宗新.媒体情绪传染与分析师乐观偏差——基于机器学习文本分析方法的经验证据[J].管理世界,2021(1):170-185+11.


Copyright © This work is licensed under a Commons Attibution-Non Commercial 4.0 International License.

[11]Gao Z.,Ren,H. & Zhang,B. Googling Investor Sentiment Around the World[J].Journal of Financial and Quantitative Analysis,2020,55(2):549-580.

[12]杨晓兰,沈翰彬.本地偏好、投资者情绪与股票收益率:来自网络论坛的经验证据[J].金融研究,2016(12):143-158.

[13]李斐.低频投资者情绪指数的构建及对创业板指数收益率影响研究[J].产业创新研究,2022(21):133-135.

[14]Antweiler W,Frank M Z.Is All that talk just noise?The information content of internet stock message boards[J].The Journal of Finance,2004,59(03):1259-1294.

[15]熊熊,罗春春,张烨.股吧和交易:股吧中的信息内容研究[J].系统科学与数学,2017,37(12):2359-2374.

[16]李媛,刘芳.股市收益、个股情绪与市场监管[J].金融监管研究,2023(1):100-114.

[17]魏卉,陈勇吉.个人投资者情绪与权益资本成本—基于行为金融视角的分析[J/OL].新疆农垦经济,1-20[2024-09-04].

[18]谢婷婷,董敏慧.投资者情绪与股票收益率变化——基于沪深300指数[J].青海金融,2024(1):56-64.

[19]周亮.投资者情绪、宏观经济变量及股票市场的关系研究[J].金融发展研究,2017(11):68-75.

**作者简介：**

江海潮(1970--),男,汉族,湖南邵阳人,博士,教授,研究方向：金融学,博弈论与经济增长。

刘娅琦(2000--),女,汉族,四川南充人,硕士,研究方向：金融理论与政策。

Copyright © This work is licensed under a Commons Attibution-Non Commercial 4.0 International License.