

# 投资者情绪对中国股票市场市场表现的影响研究

何芳 刘千秋

(湖南大学经济管理研究中心, 湖南省长沙市, 410000)

**摘要:** 本文使用上海证券交易所和深圳证券交易所 2003 年 1 月至 2020 年 2 月的变化数据, 讨论了投资者情绪对中国股市表现的影响, 应用了时间序列与岭回归对其影响进行了数据分析。本文主要是通过主成分分析分别以个人和机构投资者情绪指标作为自变量; 股票收益率与股票交易量都是对于股票市场价格波动产生影响的重要因素, 因此本文在采用收益率作为被解释变量的基础上增加成交量, 由成交量和收益率二者作为衡量中国股票市场的表现指标。并在此基础上进一步细化为上海证券交易所和深圳证券交易所, 分别考虑对其的影响。因而分别用上证、深圳综指的收益率和其成分股的交易量的对数形式作为衡量中国股票市场表现的依据; 并用 6 个宏观变量作为控制变量。最后, 使用这些变量来构建多元回归方程模型, 并采用岭回归估计。相关性分析结果显示投资者情绪与宏观经济控制变量指标普遍相关。岭回归进一步显示投资者情绪主要对于交易量有着正相关的促进作用, 而不能对于上证和深成指数产生显著影响。

**关键词:** 投资者情绪; 股市表现; 上海证券交易所; 深圳证券交易所; 岭回归

**中图分类号:** F8      **文献标识码:** A

## 1 引言

标准金融学框架依赖于理性人假设构建, 但在真实的市场环境中, 非理性因素的影响同样显著。伴随着噪声交易模型的研究发展, 投资者情绪的概念界定和衡量方式也发生了较多变化, 现阶段的投资者情绪概念更多的被表述为以来投资者普遍存在的偏好及其所形成的市场反馈。

投资者情绪是对未来现金流和投资风险, 但不为当前相关信息所证明的信念, 由此, 投资者情绪也往往暗示着投资者在没有信息支持的情况下做出投资决策。在经典金融理论中, 投资者情绪对股票价格、实现收益、预期收益的横截面没有任何影响。这些标准模型中, 股价等于没有情绪的投资者的理性预期。但金融危机等极端市场条件下的市场波动往往过于剧烈, 行为金融学试图利用投资者情绪来解释这些差异, 现代股票市场价格波动的解释和判断也往往需要引入投资者情绪的概念予以解释。因此, 投资者情绪是股票市场中一类难以测度而又真实存在的外部影响因素, 它刻画了消费者群体基于现有市场信息和企业经营状况之外的更多股票市场价格影响因素。特别是对于我国股票市场而言, 则不成熟的股票市场之中投资者情绪更为鲜明的表现为惯性效应和相应的价值偏差。

本文研究的意义是, 希望通过设计指标, 实证分析, 建立更准确反应投资者真实情绪的指标, 将投资者情绪指标与传统经济学的分析模型相结合, 能更好的帮助我们解释和预测资本市场中投资者的行为, 以及行为带来的市场影响。阐释和分析了投资者情绪的概念和成因, 选择了主成分分析法, 作为投资者情绪的设计依据, 并进一步区分了投资者情绪概念, 构建了个人投资者情绪指标和机构投资者情绪指标。分别考虑适当水平的投资者情绪数据表现与市场表现的联系, 并利用实证分析对其实际影响展开了进一步分析和论述。从而, 进一步说明不良的投资者情绪影响主要局限于个人投资者之中, 对于现有研究活动中未能进一步进行实证分析的课题进行了探讨, 补充了投资者情绪来源理论框架中的非理性因素来源观点支持

证据，在学术上有一定研究意义。

一方面，我国的投资者情绪影响股票市场价格波动和总体表现的相应作用是相对突出的，这反映出股票市场进一步发展成熟和投资者向理性投资者过度的成长过程，在这一过程中对于相应投资者情绪放大噪声的机制进行探讨能够为此类投资者提供一定积极影响。另一方面，我国的股票市场总体资金流量庞大而市场波动显著，这一背景下对于投资者情绪影响市场波动的相应机制进行研究，具有为投资者提供了解市场波动预期和变化并为相应作出投资决策的相应参考价值与现实意义。

## 2 文献综述

### 2.1 文本挖掘中的投资者情绪

对于投资者情绪的衡量方式主要由封闭基金折价率计算，但也有较为丰富的研究利用了文本信息抓取和分析对于投资者情绪进行衡量。

丁肖丽（2019）利用股吧文本信息研究了投资者情绪与市场波动的关系，显示出投资者情绪有着良好的预测作用<sup>[1]</sup>。石善冲等（2018）基于微信文本挖掘对于投资者情绪进行分析，检验其对上证指数收盘价、成交量的有关机制的影响，研究结果显示其间存在滞后一天的预测作用<sup>[2]</sup>。张信东和原东良（2017）利用微博文本分析对于投资者情绪进行了量化，继而对于这一指数预测股票市场的能力进行检验，结果显示预测效应存在<sup>[3]</sup>。Mahmoud 和 Hazar（2017）使用每日谷歌搜索查询数据趋势建立石油危机散户投资者的注意用于投资者情绪衡量指数，研究结果显示这一搜索数量的提高可以预测在交易日波动增加<sup>[4]</sup>。Turan（2016）考察了由 EGARCH(1, 1) 模型测量的杠杆（坏消息）效应是否会随着新出现的情绪代理——互联网搜索量的加入而改变。该样本包括纳斯达克（NASDAQ）和纽约证交所（NYSE）指数旗下的 14 家美国公司，以及对跨度 9 年、每周收集的 501 份数据观察。实证结果表明，纳入投资者情绪变量对坏消息效应没有明显的影响，但波动性持久性有明显的增强。行为金融理论、可用性和社会证明启发式可以作为这些发现的潜在解释<sup>[5]</sup>。汪昌云和武佳薇（2015）应用了财经媒体正负面词汇数据挖掘构建了投资者情绪，证实了其对于股票数据的影响是同向的<sup>[6]</sup>。

基于上述列举也可以看出，现有研究中关于社交媒体平台和网络平台进行文本抓取和分析活动是构建投资者情绪的一类有效方式，而不同研究者所应用的社交平台选取也有着较大差异。就国内的有关研究而言，微博、微信和股吧是主要的文本信息来源，境外研究者则应用了谷歌和雅虎数据，其中搜索情况是一类较多收到重视的数据来源，标准化水平较高，但相关研究年限较早。

### 2.2 股票市场波动的影响

相较而言，投资者情绪对于股票市场价格波动的影响是较为显著的，这一影响在个股和股票市场整体表现中均较为显著。此方面的研究数量较多，在不同层次上的相关讨论普遍有一定规模，本文试对其进行了一定的列举和对比分析。

#### 2.2.1 国外研究

相较而言，投资者情绪对于股票市场价格波动的影响是较为显著的，这一影响在个股和股票市场整体表现中均较为显著。此方面的研究数量较多，在不同层次上的相关讨论普遍有一定规模，本文试对其进行了一定的列举和对比分析。

Song 等（2019）利用连通性网络方法研究了化石能源市场、投资者对可再生能源的情

绪和可再生能源股票市场之间收益和波动的动态方向性信息溢出。实证结果表明,波动制度的溢出效应普遍强于回报制度,表明风险在市场间的传递更为明显。在两种体系中,化石能源市场,尤其是原油市场对可再生能源股票市场的影响都大于投资者情绪对可再生能源股票市场的影响。这一发现表明可再生能源股票市场与化石能源市场密切相关。此外,还采用了滚动窗口的方法来检测它们之间的时变信息溢出。动态研究结果表明,投资者对可再生能源的情绪可以在一定程度上解释可再生能源股票的收益和波动<sup>[7]</sup>。Chiu 等(2018)考察了2007-2008年金融危机期间投资者情绪如何影响股票流动性和交易行为。利用股票指数和金融 etf 的盘中数据,研究显示了投资者对报价价差、市场深度、非对称深度和净买入压力的显著非对称反应。研究还证明,资金约束会进一步增加投资者情绪对流动性和交易行为的不对称影响<sup>[8]</sup>。Ergun 和 Durukan(2017)分析1997年至2017年金融危机期间投资者情绪对伊斯坦布尔证交所数据的影响,研究者根据 CMAX 方法确定危机时期,并运用回归分析研究投资者情绪对股市回报的影响。结果显示在整个时期和局部危机时期。当反映投资者情绪的封闭式基金折价指数上升时,贝斯特 100 指数的回报率下降,表现出股票市场波动对投资者情绪非常敏感<sup>[9]</sup>。Alajekwu 等(2017)考察了投资者情绪对尼日利亚股市未来回报的影响。数据分析采用 OLS 回归和格兰杰因果关系技术。研究结果表明在控制了工业生产指数、消费价格指数和国库券利率等基本面因素后,投资者情绪对股市收益仍具有显著的正向影响;从投资者情绪的变化到( $\Delta$  CCI)股市回报( $R_m$ )有一个单向的因果关系<sup>[10]</sup>。Mehrani 等(2016)分析了乐观或悲观时期的逆势策略和动量策略,并将它们与正常的市场情绪状况进行比较。继而分析了情绪与股票收益、超额收益、波动率之间的关系。研究结果表明,将正常的市场情绪与行为金融策略相结合可以提高业绩,与动量策略相比,反向策略的效果更为显著<sup>[11]</sup>。Winmore(2015)涵盖2009年2月19日至2012年12月31日66只股票(不包括摘牌及停牌股票)的日收益及成交量数据。引入了一个高低量的情绪指标变量来区分市场的情绪是高还是低。用普通的线性回归来证明投资者情绪指标对股票收益的影响。研究人员发现,约40%的市场走势与市场情绪指标相反。剩下的60%与情绪指标共动,但影响程度大小不同,说明情绪指标对指标有正向影响。然而,使用二项检验,这些结果在统计上并不可靠,因为在5%的区间内,只有6只股票中的5只(大约7%)具有显著性。这小于原假设下预期的显著结果数量。因此,不能肯定地接受或拒绝原假设,也不能完全排除或建立情绪指标对回报的影响<sup>[12]</sup>。

### 2.2.2 国内研究

杜莉等(2017)研究了投资者情绪对定向增发长期市场表现的影响,结果显示二者呈现负相关,这一特征在财务状况较差的企业中表现更为显著<sup>[13]</sup>。陈鹏程和周孝华(2016)对于机构投资者和个人投资者情绪,这两种不同类别的投资者情绪对 IPO 首日回报率的影响机制进行了实证分析,结果显示两种不同类别的投资者情绪对于 IPO 首日回报率有不同的影响,前者起到积极影响,而后者的私人信息则起到负面影响<sup>[14]</sup>。陆静和周媛(2015)对于投资者情绪影响股价水平进行了分析,研究者采用证券投资基金的损失率来构建投资者情绪指数,即研究结果显示控制宏观经济变量背景下投资者情绪对于 A、H 股票市场价格均有积极影响<sup>[15]</sup>。邵新建(2015)探讨了新闻媒体影响投资者情绪继而对于股票长期收益产生影响的机制。这一框架显示为积极和正面的媒体评价能够带来良好的投资者情绪,继而对于长期收益不利<sup>[16]</sup>。高大良(2015)分析了投资者情绪影响平均相关性与股市收益的机制,并利用实证分析进行了检验,结果显示更高投资者情绪削弱总体风险-收益关系<sup>[17]</sup>。文风华等(2014)探讨

了投资者情绪与股票价格行为的关系,实证分析结果显示积极情绪更为显著的影响了股票价格波动并带来正面影响<sup>[18]</sup>。吴奇峰和宋泽芳(2013)运用结构方程模型检验并证实了投资者情绪与市场表现之间的关联性<sup>[19]</sup>。

综上所述,有关于投资者情绪与股票市场价格波动的研究数量众多,这也是投资者情绪相关领域研究中占比最大,数量较多的一类研究。对于其研究结果进行梳理则可以看出,大多数研究者倾向于投资者情绪在短期内对于股票价格产生积极影响,特别是对于IPO当日股票回报率有积极影响,但对于长期股票授予则有着消极影响。通常而言,个人投资者的积极情绪与熊市背景能够提升这一预测效应,但部分研究者亦认为机构投资者较个人投资者的投资者情绪影响更为显著,预测作用更为有效。从股票市场表现的角度出发,则机构投资者和理性投资者都不能起到稳定市场的作用,剔除噪音投资者、个人投资者和非理性投资者的影响也难以保持价格稳定,价格的波动仍然难以避免。

### 2.3 其他投资者情绪的影响

Ge等(2019)研究了投资者情绪与会计稳健性之间的关系。我们发现,在投资者情绪高涨的时期,管理者更及时地确认收益中的经济损失。此外,情绪-稳健性关系对情绪-价格敏感性较高的公司更强。研究者还发现,情绪-保守主义关联在具有较高诉讼风险和财务专家首席执行官的公司中更强,而在首席执行官退休的公司中则更弱。总体而言,研究结果表明,为了减少潜在的诉讼成本,公司在报告收益时会更加保守,以应对投资者情绪高涨的情况<sup>[20]</sup>。Lee和Ho(2017)考察了投机者的情绪之间的关系,期货市场的价格波动WTI原油、铜、和小麦2003~2014年期间使用格兰杰因果检验。结果表明,投机头寸总体上对每个期货市场的收益没有预测能力。相反,回报率似乎会对投机者的情绪产生影响,尤其是在经济扩张和复苏的时期。与此同时,在经济衰退期间,WTI原油和铜市场投机者情绪指数的变化,为正向回报提供了预测能力,表明投机者的悲观情绪加剧了大宗商品价格的下跌。由于投机头寸对市场价格的影响是模糊的,对投机交易的严格监管是不可取的。然而,在熊市中,监管机构应该考虑提高投机头寸限制,因为大量的投机空头头寸和清算指数交易者的多头头寸可能导致价格急剧下跌<sup>[21]</sup>。黄宏斌等(2016)对于投资者非理性情绪在企业不同的发展阶段的投融资策略的作用机制进行分析。研究结果显示生命周期影响融资方式和融资约束,而不同水平投资者情绪伴随企业生命周期变化对于这一传导机制产生影响<sup>[22]</sup>。俞红海等(2015)对于我国的股票市场企业IPO价格与长期回报之间的偏差进行了分析,引入了投资者情绪和投资者意见分歧等概念对其进行解释,显示出首日投资者情绪对于IPO价格有积极影响而对于长期价格有负面影响<sup>[23]</sup>。李凤羽(2014)对于投资者情绪预测ETF折价率的能力进行了检测,发现两者呈现正相关关系,但在悲观市场中呈现负相关<sup>[24]</sup>。

投资者情绪的影响在常被探讨的股票市场价格波动等影响之外,也有着更多显著影响。一方面,投资者情绪往往将带来企业自身的长期收益表现下降,这是为此前所涉及的有关研究所部分证实了的。另一方面,投资者情绪基于此影响企业治理,迫使企业运用更好的治理表现和更为保守的预测数据进行平衡。基于此,投资者情绪也进一步影响企业的融资约束等外部条件限制。

## 3 实证分析

本文的实证分析主要分为三个部分,一、变量选取。本文选择了主成分分析法,作为投资者情绪综合指标的设计依据,并进一步区分投资者情绪概念,分别以个人和机构投资者情

绪指标作为自变量；上证综指月度收益率及其成分股成交量的对数值、深成指数月度收益率及其成分股成交量的对数值作为衡量中国股市市场波动的因变量；6个宏观变量作为控制变量；并使用描述性统计、相关性分析来描述数据的基本特征。二、时间序列检验。本文选取时间序列数据，因此需要对数据进行稳定性检验，由于一阶差分后平稳，同阶单整，因此再进行协整检验判断数据间是否具有长期均衡关系，再通过格兰杰因果分析。三、岭回归分析。根据岭迹图选择岭参数，得到岭估计值，最后再整理和讨论了实证分析结果。

### 3.1 解释变量的描述与构建

#### 3.1.1 解释变量的描述与构建

为探讨投资者情绪对股票市场影响，考虑到数据可得性和时间限制，本文借鉴魏星集等（2014）所提出的 ICS 综合指数表，选择封闭基金折价率（DCEF）、首次公开募股首日的收益率均值（RIPO）、上月换手率（TURE）、上月新增投资者开户数（NA）、首次公开募股数量（NIPO）、消费者信心指数（CCI）六个指标的月度数据<sup>[25]</sup>，以此作为情绪代理变量。研究区间为 2003 年 1 月至 2020 年 2 月，数据来源为 CSMAR 数据库。

其中，封闭基金折价率（Closed-end fund discount, DCEF）是指封闭经济价格与市场价格之间的差异性，而这一差异显示出封闭基金定价机制未能充分满足市场价格发现机制相应定价方式的特点，因而得到了更为广泛的研究和关注。通常而言，解释封闭基金折价率所使用的是投资者的非理性因素影响，相较于自由市场交易活动，封闭式基金由于噪声投资者的存在而承受更多风险，继而处于风险厌恶形成了较低价格，这一折价率因而可以用于表现市场中的噪音交易者，从而作为投资者情绪的替代指标使用。而本文利用这一解释，将封闭基金折价率应用于展示这一投资者情绪的影响。

首次公开募股首日的收益率均值（Average first-day return, RIPO）展现出 IPO 成功后各类企业股票价格数据的短期波动，反映出股票市场对于特定企业的未来发展逾期情况，是在 IPO 发行价格所展现的市场预期水平之外另一类基于真实的市场反应而提供的的数据变化情况。而首日的收益率均值在 A 股市场的极高表现长期以来都是 A 股 IPO 的显著特征，这一数据也更多的反映出投资者对于市场整体环境的一类倾向性表现。

首次公开募股数（Number of IPOs, NIPO）往往展现出市场对于股票的反馈情况，企业通常选择市场情绪上升和保持在高位水平时进行首次公开募股，因此 IPO 数可以与市场情绪联系起来。但不同于其他指标参考市场表现情况，IPO 数通常是更多的与 IPO 审核过程相联系的，是监管部门基于市场股票价格波动情况特别是 IPO 情况进行调节的一类数据，从而使得此类指标的具体表现更为复杂，其分析结果也更为多样化。

新增开户数（New account, NA）是对于市场中投资者数量的变化进行测量的一种方法，传统上是基于个人用户视角的，但随着我国对于一人一户指标要求的放开，此类数据更多的展现出投资者数量和规模的变化。因此，新增开户数同样对于市场热烈情绪的一种有效反应，是具体的投资者所进行的相应决策。

上月市场换手率（Turnover, TURN）是市场中交易行为所形成的一类数据，展现出股票市场之中转手的频率和不同主体在具体的股票价格发展趋势方面所展现的具体判断差异。此类差异的积累和增大展现为市场换手率的上升，也更多的说明了市场中存在着活跃的投资者群体，大量投资者通过提高换手率的方式扩大了股票的流通性。但对于市场萧条的场景中，则现阶段的价格区间距离目标价格水平，从而使得市场流通性下降，投资活跃程度不足，进

一步带来了交易行为的削减。

上月消费者信心 (Consumer Confidence Index, CCI) 是一类这很对于个人投资者个人情绪的判断指标, 这一指标更多的展现出短期内的市场情绪, 是消费者群体对于当前经济环境的一种综合性评价结果, 而其对于未来市场价格波动的预期也极大地建立在这一当下市场环境和经济发展水平的预期成果基础上, 从而形成了对于市场情绪的有效反应机制。

本文使用主成分分析法, 作为投资者情绪的设计依据。在进行实证前, 首先对数据进行预处理: 在整体数据中, 由于政策原因据统计在 A 股历史上, 曾多次暂停 IPO, 所以首次公开募股数量(NIPO)和首次公开募股首日的收益率均值 (RIPO) 数据缺值较多, 因此在本文实证时间跨度期间缺失的数据, 采取 0 值补充, 尽量保持样本序列的连续性。

表 1 代理变量的描述性统计

指标	最小值	最大值	均值	标准差
DCEF	-0.3321	-0.0228	-0.14802	0.090371
RIPO	-0.0521	6.2674	0.567022	0.67312
NIPO	0	54	13.117	12.3993
NA	2.3	297.47	38.8035	45.06988
TURN	0.0852	0.8247	0.252741	0.132392
CCI	85.7	126.6	103.221	9.8588

在完成描述性统计后, 在对上述指标进行相关性分析。由表 2 可见, 各类指标间的关联性是较为复杂的, 全部的指标都存在着某些指标领域中与其相关性较强的特征, 说明了总体上各类数据呈现出多重共线性特点, 因而适应于应用主成分提取。

表 2 代理变量的相关性分析

指标	DCEF	RIPO	NIPO	NA	TURN	CCI
DCEF	1					
RIPO	-0.038	1				
NIPO	0.448***	0.001	1			
NA	0.071	0.428***	0.243***	1		
TURN	-0.261***	0.194***	-0.005	0.559***	1	
CCI	0.754***	-0.046	0.386***	0.082	-0.133	1

注: \*\*\*, \*\* 分别代表 1%, 5% 的显著性水平。(下同)

继而, 对代理变量进行主成分提取以构建投资者情绪, 结果如下表 3 所示, 由第二行第七列的累计主成分提取百分比可知, 投资者情绪因子的 66.333% 可由第一个主成分和第二个主成分解释, 其中, 第一个主成分能够解释投资者情绪因子的 35.719%, 第二个主成分能够解释投资者情绪因子的 30.614%。方差的累计贡献率为 66.333%, 虽然小于等于 85%, 但属于可以接受的范围, 因此, 可以进行投资者情绪因子合成。

表 3 代理变量的主成分提取

成分	特征根			主成分提取		
	特征根	方差解释率%	累积%	特征根	方差解释率%	累积%
1	2.143	35.719	35.719	2.143	35.719	35.719
2	1.837	30.614	66.333	1.837	30.614	66.333
3	0.835	13.915	80.248	-	-	-

4	0.635	10.582	90.83	-	-	-
5	0.332	5.527	96.357	-	-	-
6	0.219	3.643	100	-	-	-

进而可以得到各指标的载荷系数和公因子方差，内容如下表所示：

表4 代理变量的公因子荷载情况

指标	载荷系数		共同度(公因子方差)
	主成分 1	主成分 2	
DCEF	0.913	-0.027	0.834
RIPO	-0.076	0.633	0.407
NIPO	0.675	0.254	0.52
NA	0.097	0.899	0.817
TURN	-0.279	0.75	0.64
CCI	0.873	0.026	0.762

如表4所示，载荷系数是指变量DCEF、RIPO等指标在主成分1、主成分2等公共因子上的相对重要性，其统计学意义表现为该变量和该因子间的相关性，是该指标依赖于该因子的具体水平。而公共度，也即公因子方差则是这一指标下各个公因子所能够解释的累计贡献度，累积而成为公因子方差。此类累计贡献度衡量了指标的解释能力，其数值越高，说明这一指标之中的更多数据能够为相应的公因子所解释，也说明了公因子对于指标的阐释能力更强，从而展现了公因子作为后续数据之中具有代表性的特殊变量而用于后续分析的基础。这一基础也决定了提取主成分可以作为原始数据之中数据表现的一类突出特征而得到有效表现，从而探讨其与更多指标和因素之间的关联性，更好的从模糊而复杂的原始数据之中提取出具有更强有效性特征的相应数据。

由于Sentiment的计算公式为：

$$Sentiment = \sum ((\text{方差贡献率} / \text{累计方差贡献率}) * \text{主成分}) \quad (1)$$

$$\text{则可以得到 } Sentiment = 0.5281722 * DCEF + 0.2577531 * RIPO + 0.329316 * NIPO + 0.5174061 * NA + 0.405312 * TURN + 0.4825746 * CCI \quad (2)$$

由此可知，DCEF对Sentiment的影响最强，RIPO则最弱。

另外，由于封闭基金折价率(DCEF)和新增开户数(NA)这两个代理变量是机构投资者和个人投资者的最大分野，故而以解释变量进入的形式，可以由上述公式(1)计算得到二者的表达式。如果区分这两个变量，则无法捕捉这一差异性，因而无法进行相关分析。

为了进一步区分整体投资者情绪因子中的个人投资情绪和机构投资者情绪分别对股票收益的影响，本文通过控制变量的方法，分别对变量进行主成分分析，合成个人投资情绪因子和机构投资者情绪因子，其中控制变量为首次公开募股首日的收益(RIPO)、上月换手率(TURN)、首次公开募股个数(NIPO)、消费者信心指数(CCI)。在构建个人投资者情绪时增加代理变量，新增开户数(NA)；构建机构投资者情绪时增加代理变量，封闭基金折价率(DCEF)，其合成方式参考式(2)。其中，个人投资者情绪指标为：Sentiment-I；机构投资者指标为Sentiment-O。

由此我们可以得到

$$\text{Sentiment-I}=0.277321*\text{RIPO}+0.45611148*\text{NIPO}+0.40847752*\text{TURN}+0.45415392*\text{CCI}+0.53245632*\text{NA} \quad (3)$$

$$\text{Sentiment-O}=0.55497792*\text{DCEF}+0.3897504*\text{RIPO}+0.34910976*\text{NIPO}+0.40374144*\text{TURN}+0.5096736*\text{CCI} \quad (4)$$

### 3.1.2 被解释变量的选取

股票收益率指标是反应股票的收益情况,而收益情况又作为投资者投资行为的主要引导性因素和最终成果而相应成为股票市场环境中的重要参考指标。与此同时,股票市场价格波动不仅仅表现为不同股票的股息差异,也同样与股票市场中的交易行为密切联系,而衡量市场中交易情况的重要指标则可以表现为股票交易量。因此,股票收益率与股票交易量都是对于股票市场价格波动产生影响的重要因素,而与之相关的更多指标则可以被纳入数据分析的基础之中。

具体而言,上海证券综合指数简称上证指数,该指数利用加权的方式获得,其有效地反映了上交所全部上市企业的股票价格波动情况,是对于股票价格波动整体态势的一种展示手段。上证综指对于股票价格的整体展现与其他样本规模更小的指数一同组成了较为完整的市场动态指数展现方式。基于此类文献也可以更多的展现出上证综指在展现市场波动整体态势,将之与其他金融市场波动情况进行平行对比时是较好的数据来源。

深圳证券交易所成份股价指数简称深成指数,能够较好的展现出深圳证券交易所股票市场的价格波动情况。深成指更多的作为广安的指数系统之中在深交所具有较强代表性的一支而展示深交所的具体市场情况。而更为广泛的数据表现则大多将之纳入到完整的数据框架之中进行统一处理。

考虑到我国大陆只有上海证券交易所和深圳证券交易所,为了进一步对比分析投资者情绪对中国股票市场市场表现的影响,本文将对对比分析两个投资者情绪指标对两个市场的市场表现的影响,最后再对比分析其结果。因此,本文以上证指数月度收益率为上海证券综合指数收益率(SHR),以成交量的对数结果为股票的成交量(SHV)指标。时间跨度为2003年1月至2020年2月,数据来源为CSMAR数据库。深成指数的月度收益率为深圳证券综合指数收益率(SZR),以成分证券成交量的对数结果为股票的成交量(SZV)指标。时间跨度为2003年1月至2020年2月,数据来源为CSMAR数据库。

### 3.1.3 控制变量的选取

控制变量主要为宏观变量。本文选择周亮(2017)所选用的指标固定资产完成度同比增速(FA)、房地产开发完成额取对数值(REALE)、广义货币同比增速(M2)、人民币对美元汇率(EX)、国内信贷规模取对数值(CREDIT)<sup>[26]</sup>、工业增加值同比增速(INDUSTRY)<sup>[27]</sup>;。时间跨度为2003年1月至2020年2月,数据来源于国家统计局、CSMAR、RESSET数据库。

工业增加值同比增速(INDUSTRY)展现出制造业发展的基础水平,且保持着长期的稳定上升趋势。而其增速波动情况通过同比方式展现则更好地说明综合因素影响下短期内的制造业发展波动情况。数据来源为国家统计局。

固定资产完成度同比增速(FA)是对于我国的固定资产投资项目和固定定资产处置过程中所形成的固定资产投资进行衡量的一种方式,而固定资产完成度则新一步说明了这一固定投资的实践情况,而固定资产完成度同比增速则展现出基于上一周一进行调节的固定资产完成度对比情况。数据来源为RESSET数据库。

房地产开发完成额取对数值 (REALE) 是对于固定资产中的特殊类型, 仅有开发商等机构参与的一类新建建筑的工作, 而其完成额主要是依托于现有的投资规模所决定的, 直接反映出本地区的高敏感度实时资金流动情况, 因而本文对其取对数值表示。数据来源为 RESSET 数据库。

广义货币同比增速 (M2) 泛指现金与储蓄存款等内容构成的广义货币, 而广义货币是基于央行统计的一类特定指标, 既具有经济学意义又具有一定的标志性而更多的有了广泛的代表性, 其同比增速展现出是能够及发展的整体趋势, 是对于经济走势进行一定预判的有效方式。数据来源为 RESSET 数据库。

人民币对美元汇率 (EX) 为常用的经济指标, 既说明了货币价值的差异和对比, 亦可以利用其数值变化展现国际贸易发展态势。数据来源为 CSMAR 数据库的日度数据加权而成。

国内信贷规模取对数值 (CRIDIT) 指特定时间点上的贷款总余额, 一方面, 这一指标是基于央行的货币政策实现意图而确定的特定贷款控制指标, 另一方面, 信贷规模是基于这一控制指标进一步经由实践发展形成的特定信贷业务规模及其结构特征。数据来源为 RESSET 数据库。

### 3.1.4 数据的基本特征

对各变量进行描述性统计, 可以看到对比房地产开发完成额与国内信贷规模取对数值可以发现前者的标准差相对较大, 而平均值相对较小, 两者均显示出非常显著的快速上升趋势, 但依然可能来自于存在空值的最小值出现于数据中。房地产开发活动与信贷规模的总体快速上升一同展现出了金融市场介入到房地产开发行业之中的更多线索, 但与此同时, 此类金融行业参与是以信贷规模扩大而非证券化的形式进行的。对比不同房地产行业与金融业相结合的案例则可以看出, 此类基于信贷规模扩大也即银行业参与的房地产行业开发活动增长是更多的依赖于银行业承担风险, 提高了银行系统自身的脆弱性的。与之相对的, 则资产证券化等方案通过市场横向分摊了风险, 则更有可能带来经济危机。基于此, 则此类同步上涨趋势给更多地展现出房地产开发完成额增长之中的政府干预和引导机制。

继而汇率的波动特征是相对较为微弱的, 汇率变动的总体变化相对较小, 振幅微弱而标准差保持在中等水平。不同于签署指标, 汇率波动的整体变化情况依赖于国际贸易市场的变化情况, 而本文所使用的汇率数据为人民币与美元汇率, 基于我国的货币政策而得以保持着较为长期的稳定性。这一特征也进一步影响了我国的信贷规模扩大等数据, 从而更多的保持了现有资金水平的变化保持稳定水平。

收益率方面, 上证指数的波动相对较弱, 但其中位数与平均值较低, 显示出剧烈波动的市场环境和整体上发展相对稍缓。相较而言, 深成指则表现出较宽的波动幅度和较高标准差, 但有着更高的预期收益率水平。而对于股票交易市场的成交量而言, 则深圳成指显著地低于上证综指, 显示出两者体量上的显著差异。上指与深指的规模差异对于数据的总体影响也是更为显著的, 由于本文直接使用了上指与深指进入数据进行分析与处理, 则使得其数量规模上的差异性进一步影响了数据波动幅度, 对于后续讨论有一定影响。

表 5 各变量的描述性统计

变量	样本量	最小值	最大值	平均值	标准差	中位数
----	-----	-----	-----	-----	-----	-----

credit	206	5.255	6.349	5.814	0.337	5.821
ex	206	6.097	8.278	7.006	0.742	6.827
fa	206	-24.5	53.02	20.465	11.146	24.35
industry	206	-2.93	29.2	11.269	5.464	10.35
m2	206	8	29.74	15.139	4.833	14.57
reale	206	0	5.121	4.15	0.758	4.256
sentI	206	-2.012	4.459	0	1.269	0.017
sent0	206	-2.441	2.936	0	1.263	0.034
shr	206	-0.246	0.275	0.007	0.077	0.007
shv	206	6.066	8.12	7.218	0.453	7.329
szr	206	-0.256	0.271	0.011	0.087	0.01
szv	206	3.425	4.291	3.898	0.22	3.972

由相关性分析可得，投资者情绪指标与宏观经济控制变量之间高度相关，会使得 OLS 伪回归，因此为了避免多重共线性的影响，本文采用岭回归。

表 6 解释变量与控制变量的相关性分析

指标	credit	ex	fa	industry	m2	Reale	sentI	sent0
credit	1							
ex	-0.766***	1						
fa	-0.843***	0.538***	1					
industry	-0.790***	0.594***	0.787***	1				
m2	-0.704***	0.353***	0.756***	0.635***	1			
reale	0.495***	-0.372***	-0.430***	-0.381***	-0.334***	1		
sentI	0.442***	-0.293***	-0.358***	-0.211***	-0.198***	0.249***	1	
sent0	0.770***	-0.525***	-0.627***	-0.497***	-0.490***	0.377***	0.833***	1

### 3.2 时间序列的检验

本文选取的是时间序列数据，因此需要对数据进行一系列检验，其检验如下：

由于时间序列的平稳性通常是时间序列分析的基础，因此首先对数据的平稳性进行检验，由于自相关函数图检验各指标的序列平稳性不够精确，本文采用单位根检验并选择采用 ADF 检验，同时根据其线形图判断，只有收益率存在时间趋势。如下表 7 所示，原始数据不平稳，从而对其一阶差分后再进行 ADF 检验。其结果如下表 7 所示，由第二行的第二列、第三列和第六列数据，对比其 ADF 值与其在 1% 显著性水平下的临界值， $|-12.70237| > |-3.462574|$ ，且对比其 P 值与 1% 显著性水平， $0.0000 \ll 0.01$ ，因此认为 sent 的一阶差分序列不存在单位根是平稳序列。由第三行的第二列、第三列和第六列数据，对比其 ADF 值与其在 1% 显著性水平下的临界值， $|-13.21943| > |-3.462574|$ ，且对比其 P 值与 1% 显著性水平， $0.0000 \ll 0.01$ ，因此认为 sentI 的一阶差分序列不存在单位根是平稳序列。以此类推可得，各变量的一阶差分序列皆为平稳序列。

表 7 ADF 检验

变量	ADF 值	1% 临界值	5% 临界值	10% 临界值	P 值
sentI	-2.872013	-5.834983	-4.245697	-3.590398	0.2249
dsent	-12.70237	-3.462574	-2.875608	-2.574346	0.0000
dsentI	-13.21943	-3.462574	-2.875608	-2.574346	0.0000
dSHR	-15.45123	-3.462737	-2.87568	-2.574385	0.0000
dSHV	-11.70437	-3.462737	-2.87568	-2.574385	0.0000
dSZR	-12.06606	-3.462901	-2.875752	-2.574423	0.0000
dSZV	-12.20715	-3.462412	-2.875538	-2.574309	0.0000

由于原始数列的不平稳，则说明其在短期内不存在线性关系，但由表 7 ADF 检验可知，各指标同阶单整，因此，其满足进一步协整分析的条件。由下表 8 第一行结果可得，对比其迹统计量值与其在 5% 显著性水平下的临界值， $185.4007 > 125.6154$ ，且对比其 P 值与 5% 显著性水平， $0.0000 << 0.05$ ，由此认为拒绝没有协整向量为“None”的原假设；由下表 3.8 第二行结果可得，对比分析其迹统计量值与其在 5% 显著性水平下的临界值， $120.4122 > 95.75366$ ，且对比其 P 值与 5% 显著性水平， $0.0004 << 0.05$ ，由此认为拒绝最多有一个协整关系为“*At most 1*”的假设；由下表 3.8 第三行结果可得，对比分析其迹统计量值与其在 5% 显著性水平下的临界值， $66.42488 < 69.81889$ ，且对比其 P 值与 5% 显著性水平， $0.0905 > 0.05$ ，由此认为不拒绝最多有两个协整关系为“*At most 2*”的假设；检验到此结束。因此，协整检验通过，也说明了数据间具有长期的均衡关系，可以利用进行因果分析。

表 8 Johansen 协整检验

协整向量 原假设	特征根	迹统计量	5% 临界值	P 值
None*	0.276262	185.4007	125.6154	0.0000
<i>At most 1</i> *	0.235546	120.4122	95.75366	0.0004
<i>At most 2</i>	0.116035	66.42488	69.81889	0.0905
<i>At most 3</i>	0.080696	41.63392	47.85613	0.1692
<i>At most 4</i>	0.068453	24.72201	29.79707	0.1716
<i>At most 5</i>	0.040757	10.46943	15.49471	0.2463
<i>At most 6</i>	0.010421	2.105692	3.841466	0.1468

只有平稳序列或者有协整关系的单位根过程才可用于进行格兰杰因果分析。由表 8 结果可知，各指标通过协整检验，因此可以进行格兰杰因果检验。对投资者情绪、股票交易量和股票收益率进行因果关系分析，由表 9 可以知，各类数据直接影响投资者情绪，而投资者情绪未能对于任何数据提供预测作用。各项指标普遍表现为投资者情绪的格兰杰因，而投资者情绪不为任何指标的格兰杰因，对于前者而言，p 值保持在极低水平，充分证明了这一影响。

表 9 格兰杰因果检验

原假设	F 检验	P 值
-----	------	-----

原假设	F 检验	P 值
SENTI does not Granger Cause SHR	0.50068	0.6069
SHR does not Granger Cause SENTI	34.6667	1.00E-13
SENT does not Granger Cause SHR	0.58125	0.5601
SHR does not Granger Cause SENT	33.7954	2.00E-13
SENTO does not Granger Cause SHR	0.58523	0.5579
SHR does not Granger Cause SENTO	19.0961	3.00E-08
SENTI does not Granger Cause SHV	0.88268	0.4153
SHV does not Granger Cause SENTI	10.8293	3.00E-05
SENT does not Granger Cause SHV	0.26379	0.7684
SHV does not Granger Cause SENT	10.6761	4.00E-05
SENTO does not Granger Cause SHV	0.39515	0.6741
SHV does not Granger Cause SENTO	14.7401	1.00E-06
SENTI does not Granger Cause SZR	0.34351	0.7097
SZR does not Granger Cause SENTI	32.4415	6.00E-13
SENT does not Granger Cause SZR	0.47608	0.6219
SZR does not Granger Cause SENT	29.5202	6.00E-12
SENTO does not Granger Cause SZR	0.57883	0.5615
SZR does not Granger Cause SENTO	19.4024	2.00E-08
SENTI does not Granger Cause SZV	1.4545	0.236
SZV does not Granger Cause SENTI	35.5377	6.00E-14
SENT does not Granger Cause SZV	1.20082	0.3031
SZV does not Granger Cause SENT	31.7646	1.00E-12
SENTO does not Granger Cause SZV	0.44663	0.6404
SZV does not Granger Cause SENTO	21.3293	4.00E-09

### 3.3 岭回归

考虑到解释变量和各宏观控制变量间可能存在的多重共线性,本文选择使用岭回归进行数据分析,其基本原理如下:

对于最小二乘法的改善有着多种方法,概括而言有两种主要思路。其一为直接处理解释变量自身的独立性,此类方法在变量分析中常被应用,也是经典的统计学分析方法,主成分回归则是其中应用较多和较为成熟的一类例证。而另一类解决方案是寻找有偏估计,通过损失无偏性而保障较好的稳定性,岭回归能够有效应对相关性过强的自变量关系及其共线性问题。

继而引入损失函数一正则化项,再使其平方和最小,则有:

$$\beta(K) = (X^T X + KI)^{-1} X^T Y \quad (5)$$

公式(5)我们称为 $\beta$ 的岭估计回归, $K$ 为岭参数。而伴随着岭参数的增长,各元素 $\beta(K)_i$ 的绝对值趋小,当岭参数趋于无穷大,则 $\beta(K)_i$ 趋于0,此过程中 $\beta(K)$ 的变化轨迹称为岭迹。

对于正交模型而言, $K$ 是其 $P \times P$ 的对角矩阵,对角矩阵对角线上元素为 $k_1, k_2, \dots, k_p > 0$ ,即偏置参数,其余元素均为0,共同组成矩阵。因而 $K$ 值的选取影响着岭估计的结果,而为了降低岭估计的均方误差,需要根据实际的问题和数据情况进行选取。岭参数 $K$ 值的选择主要为岭迹图法;方差膨胀因子法;控制残差平方和法。通常而言,在不考虑更为复杂的数学模型框架和特殊情况时,通常使用岭迹图法选取 $K$ 值。本文采用岭迹图法。

## 3.3.1 投资者情绪对上交所股票交易量影响研究

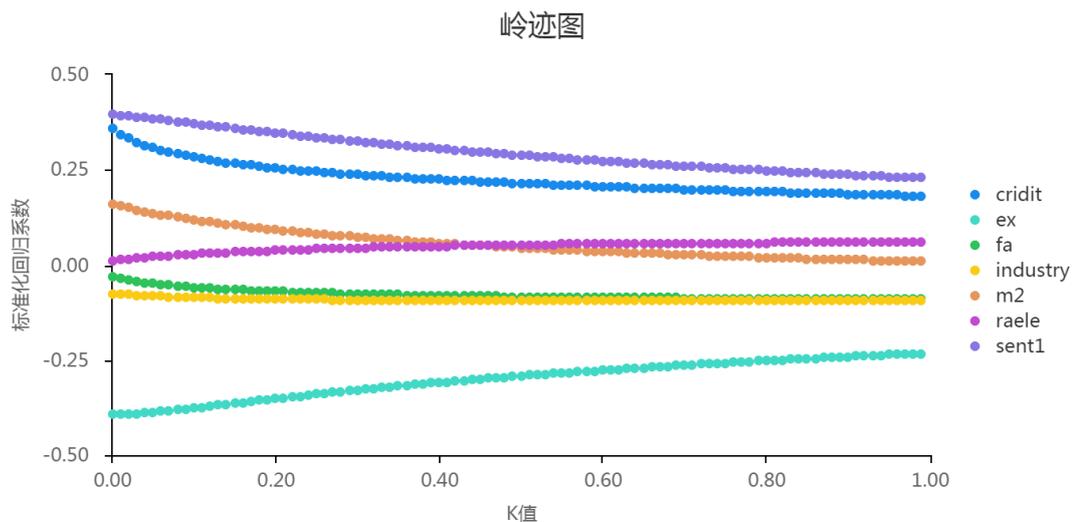


图 1 个人投资者情绪对上交所股票交易量影响岭回归岭迹图

根据岭迹图 1，当  $K=0.40$  时岭迹大体达到稳定。依岭迹法取  $K$  值为 0.40，以 Senti 为解释变量，6 个宏观指标作为控制变量，上证股票交易量的对数值 (SHV) 为被解释变量，进行岭回归。

由岭回归结果表 10 可得， $F$  值为 192.578， $p$  值为 0.0000，模型成立。固定资产完成度同比增速 (fa)，房地产开发完成额 (reale) 对上交所的股票交易量数据影响不显著，除此，其他变量对对上交所的股票交易量数据有影响显著。由  $R^2=0.872$  可知，该模型具有很好的拟合优度。相较于个人投资者情绪，对上交所的股票交易量数据影响最大的是信贷规模数据，其系数为 0.423，其次为汇率数据，其系数为 -0.238，更多的显示出股票交易量依托于国际范围内的资金跨国流动，而与投资者自身判断关联较弱。由投资者情绪的回归系数 0.138 可知，投资者情绪对上交所股票交易量有正向相关的关系，同时投资者情绪表现出数据波动对于当下股市情绪和股票交易价格惯性效应的影响，与之相比，则场外因素的干扰更少的影响投资者情绪，从而造成了其影响与投资者情绪的脱节。而对于此类交易量上升密切相关的则是场外因素中的资金流动情况，货币供应量的变化难以在短期内造成鲜明影响，更为定向提供资金支持的信贷规模数据增长影响则能更好的显示出了我国的金融也发展走向和相应企业经营特征。此后，则汇率数据的波动显示出国际市场资金流动的显著影响，这是较单项的市场价格波动而言有着更为突出影响力的一类资金来源和规模。

表 10 个人投资者情绪对上交所股票交易量影响岭回归分析

指标	回归系数
constant	6.289*** (11.804)
senti1	0.138*** (14.111)
industry	-0.007** (-2.063)
Fa	-0.002 (-1.046)
Reale	0.012 (0.724)
m2	0.013*** (3.882)

指标	回归系数
Ex	-0.238*** (-11.010)
Credit	0.423*** (6.024)
Obs	206
$R^2$	0.872
Adjusted $R^2$	0.867
F Value	F (7, 198)=192.578 , p=0.000

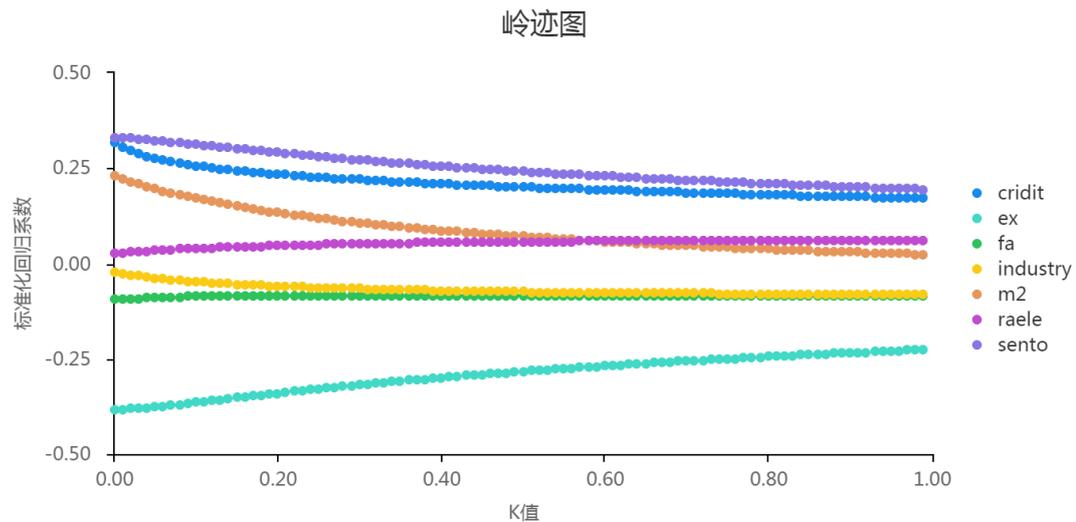


图2 机构投资者情绪对上交所股票交易量影响岭回归岭迹图

根据岭迹图 2, 当  $K=0.37$  时岭迹大体达到稳定。依岭迹法取  $K$  值为 0.37, 以 SENTO 为解释变量, 6 个宏观指标作为控制变量, 上证股票交易量的对数值(SHV)为被解释变量, 进行岭回归。

由岭回归结果表 11 可得,  $F$  值为 4.365,  $p$  值为 0.0000, 模型成立。各变量对上交所股票交易量均有显著影响, 相较于个人投资者, 机构投资者情绪回归系数仅为 0.007, 其对上证股票交易量的影响更为微弱, 机构投资者情绪相较而言对上交所股票交易量的影响力下降, 各宏观经济变量对其的影响虽然显著, 但是影响非常小。

表 11 机构投资者情绪对上交所股票交易量影响岭回归分析

指标	回归系数
constant	7.161*** (231.406)
sent0	0.007*** (10.938)
credit	0.026*** (12.257)
ex	-0.012*** (-11.710)
fa	-0.001*** (-9.195)
industry	-0.001*** (-8.715)
reale	0.001*** (5.925)

指标	回归系数
m2	-0.001*** (-5.225)
obs	206
$R^2$	0.134
Adjusted $R^2$	0.103
F Value	F (7, 198)=4.365, p=0.000

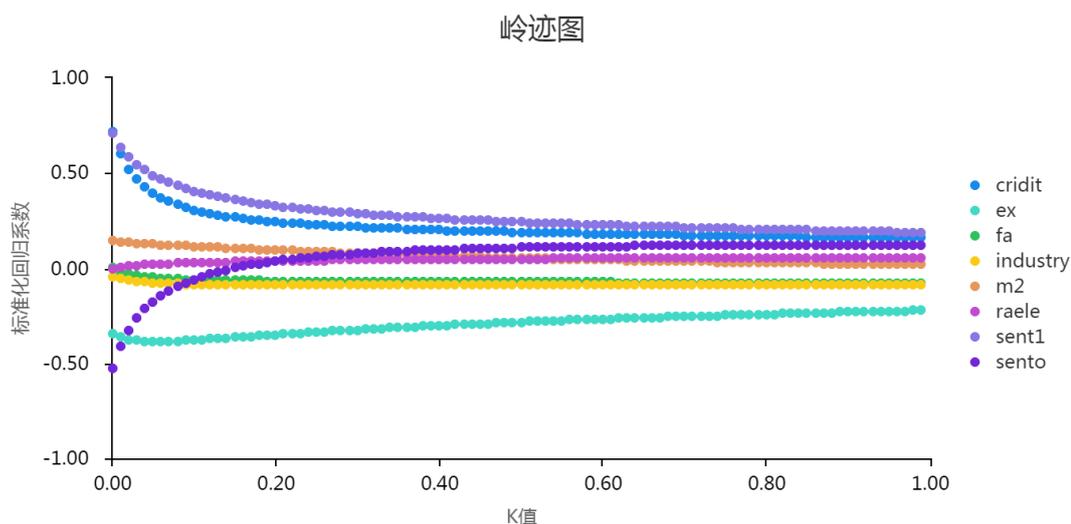


图3 机构投资者情绪对上交所股票交易量影响岭回归岭迹图

根据岭迹图3，当  $K=0.30$  时岭迹大体达到稳定。依岭迹法取  $K$  值为 0.30，以 SENTI、SENTO 为解释变量，6 个宏观指标作为控制变量，上证股票交易量的对数值(SHV)为被解释变量，进行岭回归。

由岭回归结果表 12 可得， $F$  值为 203.672， $p$  值为 0，模型成立。同时考虑 SENTI、SENTO 两个投资者情绪指标以后，机构投资者对上交所股票交易量的影响转为负值且较个人投资者情绪影响显著较弱。由此表明，个人投资者的行为所带来的风险，会阻碍理性机构投资者的套利行为。信贷规模和汇率等指标影响显著增强，货币供应量影响则较弱且为负值。

表 12 机构投资者情绪对上交所股票交易量影响岭回归分析

指标	回归系数
constant	5.078*** (9.273)
credit	0.626*** (8.408)
ex	-0.232*** (-11.193)
fa	-0.002 (-0.901)
industry	0.011 (1.035)
m2	0.012*** (3.808)
reale	0.008 (0.530)
senti1	0.195*** (14.130)
sento	-0.094*** (-5.333)

obs	206
$R^2$	0.892
Adjusted $R^2$	0.888
F Value	F (8, 197)=203.672 , p =0.000

### 3.3.2 投资者情绪对深交所股票交易量的影响

岭迹图

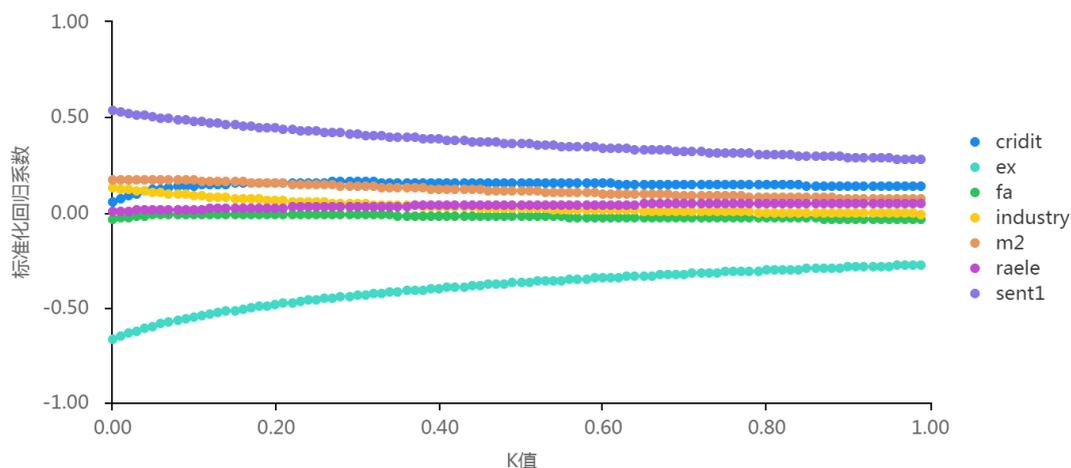


图4 个人投资者情绪对深交所股票交易量影响岭回归岭迹图

根据岭迹图4，当  $K=0.50$  时岭迹大体达到稳定。依岭迹法取  $K$  值为 0.50，以 SENTI 为解释变量，6 个宏观指标作为控制变量，深圳成分证券交易量的对数值(SZV)为被解释变量，进行岭回归。

由岭回归结果表 13 可得， $F$  值为 62.548， $p$  值为 0，模型成立。其中较为突出的是汇率对于深成指股票交易量数据有着更为显著的影响，这一影响是显著高于上交所的相应影响的，显示出深交所所在资金流动方面有着更多国际资金流动的参与和影响。这一差异性可能来自于上交所和深交所在现有资金规模方面的差异性，这一差距也在相应的指标规模之中得以相应体现出来。而此类现有资金池与外来资金比例差一些则可能更多的影响后续数据分析之中的比例差异性和讨论结果。

表 13 个人投资者情绪对深交所股票交易量影响岭回归分析

指标	回归系数
constant	3.857*** (69.473)
cridit	0.089*** (11.138)
sentI	0.048*** (14.301)
ex	-0.082*** (-15.635)
fa	-0.001** (-2.411)
industry	-0.000 (-0.525)

指标	回归系数
reale	<b>0.013**</b> (2.337)
m2	<b>0.003***</b> (3.901)
obs	206
$R^2$	0.689
Adjusted $R^2$	0.678
F Value	F (7, 198)=62.548 , p=0.000

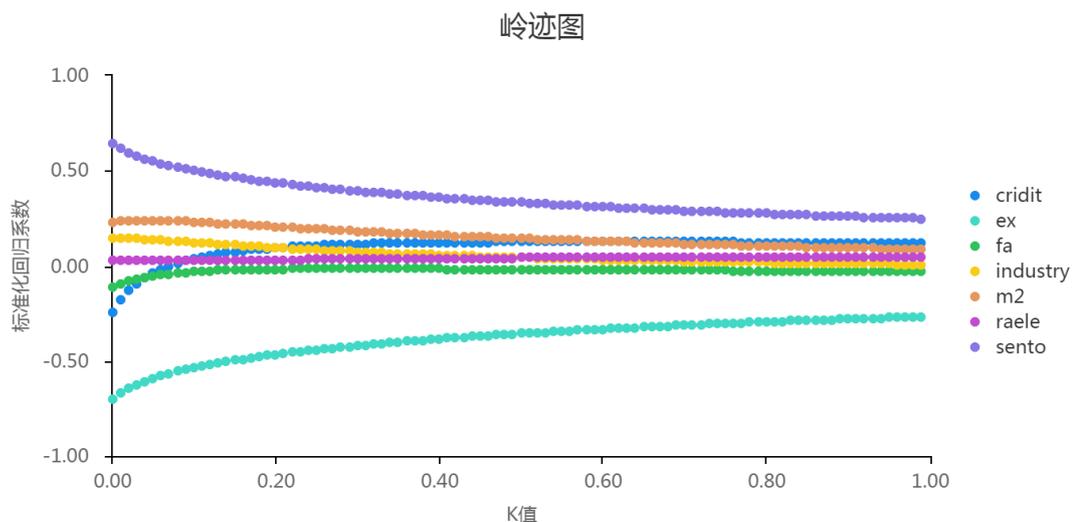


图5 机构投资者情绪对深交所股票交易量影响岭回归岭迹图

根据岭迹图 5，当  $K=0.99$  时岭迹大体达到稳定。依岭迹法取  $K$  值为 0.99，以 SENTO 为解释变量，6 个宏观指标作为控制变量，深圳成分证券交易量的对数值(SZV)为被解释变量，进行岭回归。

由岭回归结果表 14 可得，F 值为 49.531，p 值为 0，模型成立。对于机构投资者情绪而言，深成指的股票交易量受到汇率影响更为显著，此类资金流动极大地影响着股票市场，而机构投资者情绪的波动影响与个人投资者相比则变化较小，基本保持同一水平的影响。

表 14 机构投资者情绪对深交所股票交易量影响岭回归分析

指标	回归系数
constant	<b>3.903***</b> (61.786)
credit	<b>0.080***</b> (9.442)
ex	<b>-0.084***</b> (-13.795)
fa	-0.001 (-1.623)
industry	0.000 (0.622)
m2	<b>0.004***</b> (4.766)
reale	<b>0.013**</b> (2.059)

指标	回归系数
sent0	<b>0.046***</b> (12.721)
obs	206
$R^2$	0.637
Adjusted $R^2$	0.624
F Value	F (7, 198)=49.531 , p=0.000

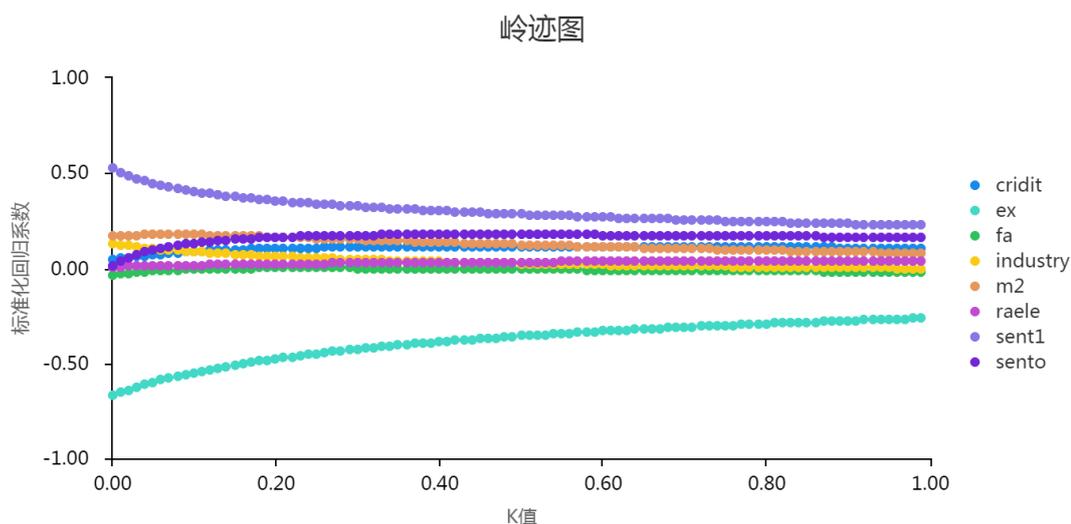


图6 机构、个人投资者情绪对深交所股票交易量影响岭回归岭迹图

根据岭迹图6, 当  $K=0.99$  时岭迹大体达到稳定。依岭迹法取  $K$  值为 0.99, 以 SENTI、SENTO 为解释变量, 6 个宏观指标作为控制变量, 深圳成分证券交易量的对数值(SZV)为被解释变量, 进行岭回归。

由岭回归结果表 15 可得,  $F$  值为 60.383,  $p$  值为 0, 模型成立。同时考虑 SENTI、SENTO 两个投资者情绪指标以后, 则能够显示出相对于机构而言个人投资者情绪对深交所交易量的影响略高且都呈现出正向的相关关系。其他指标中汇率、信贷规模、货币供应量与房地产开发完成额等指标均产生一定影响, 而其中又主要以货币供应量的影响显著。

表 15 机构、个人投资者情绪对深交所股票交易量影响岭回归

指标	回归系数
constant	<b>3.951***</b> (75.887)
credit	<b>0.069***</b> (9.696)
ex	<b>-0.078***</b> (-15.357)
fa	-0.000(-1.351)
industry	-0.000(-0.075)
m2	<b>0.004***</b> (4.747)
reale	<b>0.011**</b> (2.032)
sentI	<b>0.039***</b> (13.753)

sento	0.028*** (12.116)
obs	206
$R^2$	0.71
Adjusted $R^2$	0.699
F VALUE	F (8, 197)=60.383 , p =0.000

### 3.3.3 个人投资者情绪对上证综指的影响

依岭迹法取 K 值为 0.99, 以 SENTI 为解释变量, 6 个宏观指标作为控制变量, 上证综合指数 (SHR) 为被解释变量, 进行岭回归。F 值为 0.892, p 值为 0.514, 模型不成立, 机构投资者情绪对上证综合指数影响有限。

当 K=0.99 时岭迹大体达到稳定。依岭迹法取 K 值为 0.99, 以 SENTO 为解释变量, 6 个宏观指标作为控制变量, 上证综合指数 (SHR) 为被解释变量, 进行岭回归。F 值为 0.477, p 值为 0.851, 模型不成立, 个人投资者情绪对上证综合指数影响有限。

当 K=0.99 时岭迹大体达到稳定。依岭迹法取 K 值为 0.99, 以 SENTI、SENTO 为解释变量, 6 个宏观指标作为控制变量, 上证综合指数 (SHR) 为被解释变量, 进行岭回归。F 值为 0.793, p 值为 0.609, 模型不成立, 个人、机构投资者情绪对上证综合指数影响有限。

### 3.3.4 机构投资者情绪对深圳综指的影响

当 K=0.99 时岭迹大体达到稳定。依岭迹法取 K 值为 0.99, 以 SENTI 为解释变量, 6 个宏观指标作为控制变量, 深成指数 (SZR) 为被解释变量, 进行岭回归。F 值为 1.332, p 值为 0.237, 模型不成立, 个人投资者情绪对深成指数影响有限。

当 K=0.99 时岭迹大体达到稳定。依岭迹法取 K 值为 0.99, 以 SENTO 为解释变量, 6 个宏观指标作为控制变量, 深成指数 (SZR) 为被解释变量, 进行岭回归。F 值为 0.771, p 值为 0.613, 模型不成立, 机构投资者情绪对深成指数影响有限。

当 K=0.99 时岭迹大体达到稳定。依岭迹法取 K 值为 0.99, 以 SENTI、SENTO 为解释变量, 6 个宏观指标作为控制变量, 深成指数 (SZR) 为被解释变量, 进行岭回归。F 值为 1.163, p 值为 0.323, 模型不成立, 个人、机构投资者情绪对深成指数影响有限。

## 4 研究结论

### 4.1 主要结论

通过上述分析和讨论可以看出, 基于外部宏观数据指标、投资者情绪和股票交易量之间的关联性表现为宏观数据的争相表现带来了投资者情绪的积极性增长, 继而带来了股票交易量上升。从国内外先行研究的成果进行分析, 投资者情绪对于股票市场特别是股票价格的短期上涨得到了进一步验证, 而本文更引入大量控制变量, 通过区分机构投资者情绪与个人投资者情绪等方式对于这一课题进行了细化。

最投资者情绪对于股票交易量的促进作用是极为显著的, 这一判断自身能够得到数据的有力支持, 也同样为大量先行研究所认可。通过现有理论和实证分析结果则可以看出, 相较于机构投资者情绪 SENTO 而言, 个人投资者情绪 SENTI 对上交所和深交所股票交易量的影响略高。同时考虑 SENTI、SENTO 两个投资者情绪指标以后, 机构投资者对上交所股票交易量的影响转为负值且较个人投资者情绪影响显著较弱。由此表明, 个人投资者的行为所带

来的风险,会阻碍理性机构投资者的套利行为。但是该现象在深交所并未体现。虽然 SENTI、SENTO 两个投资者情绪指标与交易量之间具有显著的正向相关关系,但却并未对收益率产生影响,笔者认为虽然 SENTI、SENTO 两个投资者情绪指标与交易量之间具有显著的影响,但影响系数太小以致于并未反映在股价的波动上,因而二者未对收益率产生显著的影响。

## 4.2 创新点

本文分析和讨论了 SENTI、SENTO 二者对中国股票市场市场表现的影响,并应用了文献综述,理论分析和实证分析等方法对于这一课题进行分析。

本文的主要创新点在于运用了实证分析的方法对于股票交易量数据的有关影响因素进行了罗列和分析,对于机构投资者和个人投资者情绪数据的区分和讨论,个人、机构投资者情绪对于中国股票市场表现的影响,相应的这可以为进一步分析和讨论相应的投资者情绪提供一定的价值,本文也即基于机构投资者和个人投资者在投资者情绪方面的表现差异提供相应政策性建议。对比分析二者对中国股票市场市场表现的影响;由于投资者情绪直观影响为对交易量的影响,因此本文在采用收益率作为被解释变量的基础上增加成交量,由成交量和收益率二者作为衡量中国股票市场的表现指标。并在此基础上进一步细化为上海证券交易所和深圳证券交易所,分别考虑对其的影响。并基于数据分析方法而排除了其他因素影响。因此,本文可以更为准确的描述机构投资者对于股票市场交易量水平变化的实际影响,并将之进一步量化为具体的表达形式。

## 4.3 不足

而本文目前的主要不足和缺憾仍然是来自于数据分析手段的,现有的数据分析基于较小规模的数据,这是由于股票市场的总体时间跨度相对较短等原因而造成的,而各类股指体系也未能得到充分和有效的建立,从而使得相应的数据分析往往由于原始数据的规模相对不足而缺乏进一步深入探讨的空间。

## 参考文献

- [1] 丁肖丽.网络讨论、情绪分化与股票市场表现[J].湘潭大学学报(哲学社会科学版),2019, 43 (02): 75-81.
- [2] 石善冲, 朱颖楠, 赵志刚, et al.基于微信文本挖掘的投资者情绪与股票市场表现[J].系统工程理论与实践,2018, 38 (6): 1404-1412.
- [3] 张信东, 原东良.基于微博的投资者情绪对股票市场影响研究[J].情报杂志,2017, 036 (008): 81-87.
- [4] Qadan M, Nama H. Investor sentiment and the price of oil[J]. Energy Economics,2018, 69 (jan.): 42-58.
- [5] Son-Turan S. The Impact of Investor Sentiment on the "Leverage Effect"[J]. International Econometric Review,2016, 8 (1): 4-18.
- [6] 汪昌云, 武佳薇.媒体语气、投资者情绪与 IPO 定价[J].金融研究,2015 (09): 174-189.
- [7] Song Y, Ji Q, Du Y J, et al. The dynamic dependence of fossil energy, investor sentiment and renewable energy stock markets[J]. Energy Economics,2019, 84 (Oct.): 104564.104561-104564.104515.
- [8] Chiu J, Chung H, Ho K Y, et al. Investor sentiment and evaporating liquidity during the financial crisis[J]. International Review of Economics Finance,2018, 55 (MAY): 21-36.
- [9] Durukan M B, Ergun Z C. Investor Sentiment in the Crisis Periods: Evidence from Borsa

- Istanbul[J]. Pressacademia,2017, 6 (4): 309-317.
- [10] Alajekwu U B, Okoro C O, Obialor D M C, et al. Effect of Investor Sentiment on Future Returns in the Nigerian Stock Market[J]. Journal of Trend in Scientific Research and Development,2017, 1 (5): 141-155.
- [11] Mehrani K, Roodposhti F R, Nekomaram H, et al. Behavioral trading strategies and investor sentiment: Empirical research in Tehran stock exchange (TSE)[J]. Indonesian Capital Market Review,2016, 8 (2): 94-109.
- [12] Winmore M B. Measuring Investor Sentiment on the Zimbabwe Stock Exchange[J]. Asian Journal of Economic Modelling,2015, 3 (2): 21-32.
- [13] 杜莉, 范洪辰, 李思飞.投资者情绪对定向增发长期市场表现影响的实证研究[J].中国经济问题,2017 (06): 98-109.
- [14] 陈鹏程, 周孝华.机构投资者私人信息、散户投资者情绪与IPO首日回报率[J].中国管理科学,2016, 24 (04): 37-44.
- [15] 陆静, 周媛.投资者情绪对股价的影响——基于AH股交叉上市股票的实证分析[J].中国管理科学,2015, 23 (11): 21-28.
- [16] 邵新建, 何明燕, 江萍, et al.媒体公关、投资者情绪与证券发行定价[J].金融研究,2015 (09): 190-206.
- [17] 高大良, 刘志峰, 杨晓光.投资者情绪、平均相关性与股市收益[J].中国管理科学,2015, 23 (02): 10-20.
- [18] 文凤华, 肖金利, 黄创霞, et al.投资者情绪特征对股票价格行为的影响研究[J].管理科学学报,2014 (03): 60-69.
- [19] 吴奇峰, 宋泽芳.基于结构方程模型的投资者情绪与市场表现关系的实证研究[J].沿海企业与科技,2013 (06): 6-7+11.
- [20] Ge R, Seybert N, Zhang F. Investor Sentiment and Accounting Conservatism[J]. Accounting Horizons,2019, 33 (1): 83-102.
- [21] Bok L H, Ho P C. Effects of Investors' Sentiment on Commodity Futures Prices[J]. Journal of the Korea Convergence Society,2017, 8 (11): 383-391.
- [22] 黄宏斌, 翟淑萍, 陈静楠.企业生命周期、融资方式与融资约束——基于投资者情绪调节效应的研究[J].金融研究,2016 (07): 96-112.
- [23] 俞红海, 李心丹, 耿子扬.投资者情绪、意见分歧与中国股市IPO之谜[J].管理科学学报,2015, 18 (03): 78-89.
- [24] 李凤羽.投资者情绪能够解释ETF的折溢价吗?——来自A股市场的经验证据[J].金融研究,2014 (02): 180-192.
- [25] 魏星集, 夏维力, 孙彤彤.基于BW模型的A股市场投资者情绪测度研究 [J].管理观察,2014 (33): 71-73+76.
- [26] 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.
- [27] 张宗新, 王海亮.投资者情绪、主观信念调整与市场波动[J].金融研究,2013 (04): 142-155.

## Research on the Impact of Investor Sentiment on the Performance of China's Stock Market

Fang He

(Center for Economics, Finance and Management Studies,  
Hunan University

Changsha, 410000)

**Abstract:** this article analyzes and discusses the impact of investor sentiment on the performance of China's stock market, using data on changes in the Shanghai Stock Exchange and Shenzhen Stock Exchange from January 2003 to February 2020, and this article uses time series and ridge regression to analyze its influence. This article mainly constructs individual and institutional investor sentiment indicators as independent variables through principal component analysis; Both the stock yield and stock trading volume are important factors that have an impact on the stock market price fluctuations. Therefore, the logarithmic form of the trading volume of the Shanghai Stock Exchange and the Shenzhen Composite Index and the return of its constituent stocks are used as the basis for measuring the performance of the Chinese stock market; And use 6 macro variables as control variables. Finally, these variables are used to construct a multiple regression equation model, and ridge regression is used to estimate. The results of correlation analysis demonstrate that investor sentiment is generally correlated with six macroeconomic indicators, but there is no significant correlation with yield. Ridge regression further shows that investor sentiment is mainly positively correlated with trading volume, but there is no significant impact on the yield of the Shanghai Stock Exchange and Shenzhen Component Index.

**Keywords:** Investor sentiment; Stock market performance; The Shanghai stock exchange; The Shenzhen stock exchange; Ridge regression