

基于微博的投资者情绪对股票市场影响研究^{*}

张信东 原东良

(山西大学经济与管理学院 太原 030006)

摘要 [目的/意义]从社交媒体这一角度研究投资者情绪,对完善行为金融领域的研究具有重要意义,并可为舆情信息在金融市场的应用提供新的思路及理论支持。[方法/过程]利用“微指数”构建投资者情绪代理指标,采用向量自回归、神经网络等方法,实证分析了基于微博的投资者情绪对股票市场的影响。[结果/结论]研究结果表明,微博情绪与股票市场表现显著相关,并可以在一定程度上预测股票价格。

关键词 微博 投资者情绪 股票市场 向量自回归 神经网络

中图分类号 F830.91

文献标识码 A

文章编号 1002-1965(2017)08-0081-07

引用格式 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,2017,36(8):81-87.

DOI 10.3969/j.issn.1002-1965.2017.08.015

Research on the Impact of Investor Sentiment on Stock Market Based on Micro-blog

Zhang Xindong Yuan Dongliang

(School of Economics and Management, Shanxi University, Taiyuan 030006)

Abstract [Purpose/Significance] Studying the investor sentiment from the perspective of social media, which is of great significance to improve the research of behavioral finance, can provide new ideas and theoretical support for the application of public opinion information in financial market. [Method/Process] Using “micro-blog index” to construct investor sentiment proxy indicator, based on the methods of vector autoregressive and neural network and so on, the authors made empirical analysis of the impact of investor sentiment on the stock market based on micro-blog. [Result/Conclusion] The findings demonstrate that Micro-blog sentiment is significantly correlated with stock market performance, and can predict the stock price to a certain extent.

Key words micro-blog investor sentiment stock market vector auto regression neural network

0 引言

中国互联网络中心调查显示,截止到2016年12月,我国网民规模较2015年底增加6.25个百分点,达到7.31亿,互联网普及率为53.2%,手机网民用户达到6.56亿。随着互联网及智能手机的普及,人类社会步入了“大数据时代”,互联网已经成为现代社会中表达民情民意的主要载体。近几年,不少学者尝试从互联网数据中挖掘相关信息进行学术研究。

在已有的互联网信息与股市关系研究中,Karabulut研究表明利用Facebook数据构建的GNP指数可以预测股票市场^[1];Bollen,Mao和Zeng利用Twitter构

建的平静类情绪指标能够预测道琼斯工业指数^[2];Da,Engelberg和Gao使用谷歌趋势的数据,解释了股票市场的反转效应^[3]。宋双杰,曹晖和杨坤利用谷歌趋势数据作为投资者关注度代理变量研究了A股市场的IPO异象^[4],俞庆进和张兵研究表明,百度搜索指数对创业板市场表现有着显著影响^[5]。

诞生于互联网Web2.0时代的微博,作为目前国内最具影响力之一的社交平台,拥有庞大的用户群体,2016年12月报告显示,微博月活跃用户已达3.13亿,增幅高达46%,与此同时,日活跃用户数增至1.39亿。由于用户基数庞大,关注领域多样化,微博每天都会产生海量的数据。目前,亦有学者尝试通过挖掘微

收稿日期:2016-12-27

修回日期:2017-03-20

基金项目:国家自然科学基金面上项目“市场微观结构、特质波动率异象与MAX效应”(编号:71371113)研究成果之一。

作者简介:张信东(ORCID:0000-0002-3859-9579),女,1964年生,博士,教授,研究方向:财务管理、金融工程、资产定价;原东良(ORCID:0000-0003-1767-2439),男,1992年生,硕士研究生,研究方向:金融工程、风险管理等。

博信息研究我国股票市场。周胜臣等通过使用微博搜索数据,并借助 SVM 模型实证表明微博投资者情绪短期内可以预测股市^[6]。程琬芸和林杰研究显示,利用微博数据构建的投资者涨跌情绪指数与股票市场收益及交易量都存在正相关关系^[7]。赖凯声、陈浩、钱卫宁,等发现,微博情绪综合指数与同期及下一交易日的上证综指之间存在着显著的长期均衡关系^[8]。黄润鹏、左文明和毕凌燕验证了在 SVM 模型中加入微博情感,可以提高对股价的预测效果^[9]。前述文献研究全部采用目前较为成熟的文本挖掘及语义分析技术,并取得了一定的研究成果,但所构建指标为社会化大众互联网情绪,并非投资者情绪,研究周期基本为半年到一年,存在代表性较差、解释力度不强的问题。

根据前述背景,本文以微博数据为基础,构建直接衡量中小投资者情绪的代理指标,以行为金融理论为依据,提出假设,即基于微博信息的投资者情绪能够影响并预测股票市场。

1 理论分析与研究假设

1.1 微博情绪预测股票市场的合理性 作为投资者,经常会关注一些提供股市消息、投资建议等信息的微博账号,微博为投资者获取市场消息提供了极大的便利。且具有全民公开、观点和情感集中,以及及时更新的特性^[10]。这就决定了微博情绪的特性:一是自发性。股市信息在微博的传递过程中,不断有投资者根据自己掌握的消息发表对于股票市场的观点,这种自发的行为使得投资者通过微博表露的情绪更为纯粹。二是交互性。微博作为开放的平台,在与其他用户进行交流过程中,投资者流露出的对股票市场现状及未来预期的看法,在主导投资者自身决策行为的同时也会影响到其他投资者,这种交互过程使得投资者情绪趋于一致。三是匿名性。微博普通用户无需实名认证,匿名性的情绪表露给了投资者更为宽松的表达氛围,因此,从微博获取的情绪信息也更接近投资者的真实意愿。四是实时性。微博作为即时社交媒体,每时每刻都有海量的数据产生,相对于传统的情绪指标,基于微博信息的投资者情绪可以比较准确及时的反映出用户的心理和行为变化,这也为科学研究提供了良好的数据来源。《2015 微博用户发展报告》指出,截止到2015年上半年,财经类微博认证账号达到了12.4万个,财经类博文的日均阅读量达到了5.92亿次,其中与股市有关博文的提及度和上证综合指数走势之间又存在一定的正相关关系。《2015 微博财经白皮书》研究表明,在财经类领域中,有超过一半的用户关注的是股票市场。

基于前述分析,本文认为投资者基于对股票市场

的关注,通过微博表达的相关情绪是可以预测股票市场的。

1.2 微博情绪影响股票市场的渠道 微博作为接收信息资源与发布信息之间的桥梁,在满足投资者信息需求的同时也记录了投资者的心里变化过程,在一定程度上可以代表投资者的心里感知趋向,由这种感知趋向所主导的投资决策行为会影响股票市场交易活动。

部分学者已经研究证实微博情绪对股票市场存在影响,但在理论上还未形成共识。结合现有研究和本文的思考,这里尝试用图1来阐述微博情绪影响股票市场的渠道。

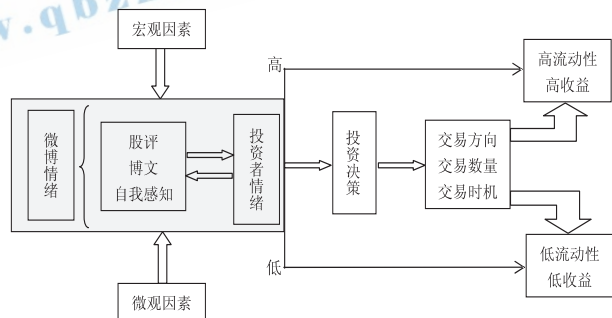


图1 微博情绪影响股票市场的渠道

投资者情绪是指投资者基于对未来现金流和投资风险的预期而形成的一种信念,但这种信念并不能完全反应当前既有的事实^[11];在有限关注下,投资者也不能充分理解市场上的信息。由于信息不对称及其他不确定因素的影响,投资者很难对资产进行准确的定价,只能对其形成某种预期并以此来指导自己的投资行为。在图1中,宏观因素包括宏观经济状况、行业信息、国家政策等,微观因素包括股市表现、公司业绩、发展前景等,在宏观因素和微观因素的共同作用下,存在一个针对投资者在微博上接收、发布信息相互作用的不可观测的“灰箱”。在灰箱内部,投资者会通过微博表达自己对股市的看法及投资预期,而各类财经类账号、股评家、其他投资者发布的股评、财经类博文又会反作用于投资者本人,两者的相互作用构成了“微博情绪”。当微博情绪高涨(低沉)时,市场交易活动更为活跃(低迷),收益率更高(更低)。

对于灰箱是如何影响投资者决策行为进而影响交易方向、交易数量及交易时机,并最终影响股票市场表现的过程,也有大量文献予以支持^[12-13]。

1.3 研究假设 股市投资者通过微博发布的信息,在很大程度上代表自己的投资预期,并由此决定自己的投资决策,其所发布的信息也会在一定程度上影响其他投资者的投资预期,进而影响股票市场,最终表现就是通过微博表露的看法(情绪)与股票市场指标存在很强的相关性。同时,股票市场的表现又会引发投

投资者情绪的变化,从另外一个方向加强了两者的相关性。这种相关性的存在,是研究微博情绪与股票市场之间关系的基础。故提出如下假设:

H1:微博情绪与股票市场指标具有相关性。

流动性是证券市场的生命力所在,是保证市场健康运行的基础。在理论上,投资者情绪越高涨,市场流动性越高;情绪越低沉,市场流动性越差。在2014年后半年到2015年上半年A股市场的一波大牛市中,市场投资者情绪高涨,市场换手率及股票成交量居高不下,市场流动性充分,市场指数一路上行。而在2015年6月出现前所未有的股灾,2016年初短短一周之内市场四次触发“熔断”,“千股跌停”的局面在很大程度上冲击了市场流动性,受此影响,整个市场的投资者情绪极度低迷,证券市场出现了流动性枯竭,投资者信心严重不足,悲观的情绪又在一定程度上加深了市场流动性的紧缩。据此,提出本文的第二个假设:

H2:在控制其他因素的条件下,微博情绪与股票市场流动性正相关。

国内基于微博情绪与股市的研究,基本都将关注点放在微博情绪与市场指标的相互关系之上,并无相关文献对后续的反转进行研究。本文所构建的微博情绪,在一定程度上也代表投资者对股市的关注,在投资者有限关注视角下的股票市场研究中,饶育蕾、彭叠峰和成大超实证表明,媒体对上市公司的关注度越高,该公司的股票在接下来的一个月内的平均收益率越低^[14]。王勇和杨庆运利用和讯网数据研究发现,网络关注与当日股票收益存在显著正相关关系,但这种关系在次日就发生了反转^[15]。

在我国的投资构成类型中,中小投资者的比例要远远高于机构投资者,由于信息的不对称,中小投资者更容易受到“小道消息”的影响,诸多学者均证实我国股票市场存在明显的“羊群效应”。当投资者情绪高涨时,会促使其在短期内完成买入行为,这种行为导致股价在短期内上涨,但这种靠高情绪而不是基本面信息支撑的投资行为,在投资者情绪恢复理性之后,股票价格将有所回落,回归其正常价值。图2可以描述该过程。

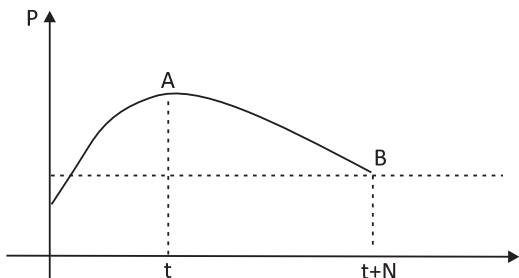


图2 股票市场收益反转

在图2中,横轴代表时间,纵轴代表股价。假设在

t期,投资者情绪高涨,股票市场交易活跃,股价大幅上涨到达A点,随着时间的推移,市场逐步趋于理性,股票已无法持续产生正收益,从而股价在t+N期回落到B点,重新回到其真实价值 P_1 。因此,提出本文的三个假设:

H3:在控制其他因素的条件下,微博情绪会对股票收益产生正向影响,该影响会在后期发生反转。

股市于周五收盘之后,经过周六、周日和法定节假日,市场层面和上市公司层面的消息可能发生一定的改变。Louis和Sun研究发现,美国上市公司管理层在公告股权并购事件时,择机性很明显,通常会选择周五投资者注意力不集中的时候进行公告^[16]。谭伟强在研究我国上市公司公布年报和半年报的择机选择时,发现坏消息更倾向于在周六被公布^[17]。权小峰和吴世农发现管理层更倾向于在周五公布利空消息。这些市场信息的变化会引起投资者情绪的变化,进而影响投资者后期的投资决策^[18]。此外,董颖红、陈浩和赖凯声等证实了微博客周末的快乐情绪要显著高于工作日^[19]。因此我们提出如下假设:

H4:交易日和非交易日微博情绪对股票市场收益率的预测能力不同。

2 研究设计

2.1 样本选择和数据来源 本文选择2013年3月1日到2016年9月30日之间上证综指作为研究样本,在数据频率上使用日度和周度数据。相关微博数据来自微指数(<http://data.weibo.com/index>),微指数是新浪微博开发的数据分析工具,于2013年3月1日正式上线,它是通过关键词的热议度,以及行业、类别的平均影响力,来反映微博舆情的发展走势。

数据的筛选:本文仅对日数据进行异常值探测,具体方法是按照拉依达准则剔除异常日收益率,并将微博数据及其他股票市场数据与之匹配,最终得到854组日数据,186组周数据。不作特别说明,本文所使用的股票市场数据均来自Wind金融数据库。

2.2 变量说明

2.2.1 微博情绪 借鉴Mao,Scott和Johan的做法^[20],我们定义微博情绪(Micro-blog Sentiment)为:

$$MS = \log(1 + N) - \log(1 + X) \quad (1)$$

其中,N和X分别代表“牛市”和“熊市”每天在微博上出现的次数,计算周微博情绪时,N和X分别代表“牛市”和“熊市”每周交易日在微博出现的次数总和。在关键词的选取上,参考已有研究确定了49个股票评论常用术语,然后邀请10名股票投资者逐一甄别,剔除有争议的词语,最终确认包括“牛市”“熊市”在内的共计36个股票术语。按照式(1)构建微博情

绪并与市场收益率进行比较,发现通过“牛市”“熊市”构建的情绪指标与收益率的走势高度一致。此外,Zhou 等也证明了与股票市场高度相关的金融词汇在微博中出现的频率可以在短期内预测上证指数的变化^[21]。对微博原始数据进行简单统计之后,发现:微博中“牛市”和“熊市”在交易日平均出现的次数要明显高于周六周日(图3),说明微博用户在交易日对股票市场的关注度要明显高于非交易日,这与我国股市周一至周五交易,周六和周日闭市的实际情况相符。

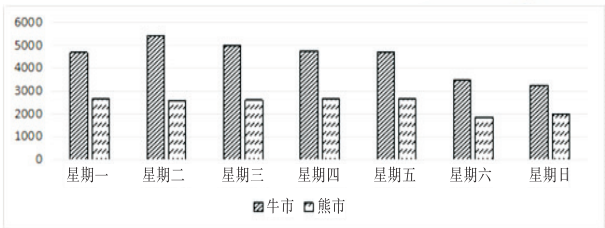


图3 “牛市”和“熊市”日均统计量

2.2.2 收益率、换手率和成交量 对股票市场表现的检验,本文主要选择了以下指标:日/周收益率(Ret),流通股日/周换手率(Dtrd)和日/周成交量(Tvol)。换手率和交易量的数据直接来自Wind数据库,其中,交易量取自然对数,收益率采用以下公式计算:

$$Ret = \text{LogClindex}_t - \text{LogClindex}_{t-1} \tag{2}$$

式中,Clindex_t代表当期日/周收盘价,Clindex_{t-1}代表上一期日/周收盘价。主要变量的描述性统计见表1。

表1 主要变量描述性统计

	日				周			
	极小值	极大值	均值	标准差	极小值	极大值	均值	标准差
MS	-3.680	5.071	0.630	0.913	-2.045	4.576	0.612	0.824
Ret	-4.366	4.803	0.159	1.349	-14.291	9.073	.130	3.581
Dtrd	0.830	7.770	2.677	1.366	0.950	7.770	2.629	1.383
Tvol	22.451	5.174	23.603	0.639	1.996	5.849	4.425	0.709

2.2.3 控制变量 在控制变量的选取上,选择对投资者情绪、股市收益和市场流动性有关的变量,具体包括市盈率(PE)、账市比(BM)和规模(Size)。本研究还将“周一效应”(Mon)“一月效应”(Jan)及季节性情绪紊乱(SAD)等因素取哑变量纳入模型^[22]。

2.3 模型设计 模型1:为了检验微博情绪是否和股票市场存在相关关系,构建单变量回归模型,设计如下:

$$MarInd_t = \beta_0 + \beta_1 MS_t + \varepsilon_t \tag{3}$$

其中,MarInd 指收益率、换手率和成交量。

模型2:在实际市场中,可能存在股票交易量放大,但是股价却在横盘整理甚至下降的情况,因而收益率指标不能很好的描述市场交易的活跃程度。因此,

选择换手率和交易量作为因变量来研究微博情绪与股票市场交易活动之间的关系。借鉴俞庆进和张兵的做法,构建以下模型:

$$MarInd_t = \beta_0 + \beta_1 MS_t + \beta_2 BM_t + \beta_3 Size_t + \varepsilon_t \tag{4}$$

MarInd 为市场交易指标,在这里为股票市场换手率和成交量。

模型3:微博情绪对股票市场收益的预测能力

为了检验微博情绪是否可以预测股票市场收益,构建式(5),借鉴 Mao, Scott 和 Johan 等学者的做法,在式(6)中加入滞后期的换手率,并控制同期的账市比、规模、日历效应等因素做进一步研究。

$$Ret_t = \alpha + \sum_{i=1}^n \beta_i MS_{t-i} + \sum_{i=1}^n \chi_i Ret_{t-i} + \varepsilon_t \tag{5}$$

$$Ret_t = \alpha + \sum_{i=1}^n \beta_i MS_{t-i} + \sum_{i=1}^n \chi_i Ret_{t-i} + \sum_{i=1}^n \delta_i Dtrd_{t-i} + \varphi PE_t + \varphi BM_t + \gamma Size_t + \eta Mon_t + \kappa Jan_t + \lambda SAD_t + \varepsilon_t \tag{6}$$

3 实证结果分析

3.1 线性回归分析

3.1.1 平稳性检验 为了避免出现伪回归,回归检验之前,先对所有变量进行单位根检验。检验结果显示,所有时间序列数据均是平稳的,不存在单位根,可以进行回归检验。

3.1.2 微博情绪与市场指标的相关性分析 表2是式(3)的回归结果,本小节,以市场交易指标为因变量,微博情绪指数为自变量,分别考察日度和周度频率上两者的关系,结果显示,所有的市场交易指标与微博情绪均是显著正相关,除周微博情绪与周股票市场收益在5%的水平上显著正相关以外,其他的均在1%

表2 微博情绪与收益率、换手率和成交量单变量回归结果

	Cons	MS	Adj-R ²	F
Panel A 微博情绪与市场指标回归(日)				
Ret	0.035 (0.628)	0.197 *** (3.932)	0.017	15.463 ***
Dtrd	2.425 *** (44.274)	0.400 *** (8.094)	0.070	65.516 ***
Tvol	23.457 *** (934.928)	0.231 *** (10.203)	0.108	104.103 ***
Panel B 微博情绪与市场指标回归(周)				
Ret	-0.355 (-1.099)	0.792 ** (2.516)	0.028	6.328 **
Dtrd	2.358 *** (19.283)	0.443 *** (3.713)	0.065	13.786 ***
Tvol	4.276 *** (68.623)	0.244 *** (4.009)	0.075	16.074 ***

注:①*、**、*** 分别表示10%、5%和1%的显著性水平。
②括号内是t值(下文同)。

的水平上相关。表明微博情绪越高,市场表现越活跃,市场收益越高,这一结果验证了假设 1。

3.1.3 微博情绪对股票市场交易活动的影响
表 3 显示,式(4)回归的 F 值都在 1% 的水平上显著,且其 P 值均为 0(文中未列示),表明参数整体上是显著的。在控制了账市比(BM)和规模(Size)之后,短期微博情绪(本文定义日微博情绪为短期微博情绪,周微博情绪为中期微博情绪)对市场流动性指标均有显著正向影响,MS 的系数分别达到了 0.081、0.092。但是,在周微博情绪和市场交易活动指标的回归中,因

表 3 微博情绪与换手率、成交量多变量回归结果

	Cons	MS	BM	Size	Adj-R ²	F
Panel A 微博情绪与换手率、成交量回归结果(日)						
Dtrd	-0.751	0.081 ***	-7.056 ***	1.186 ***	0.737	795.843 ***
	(-0.655)	(2.720)	(-16.310)	(8.696)		
Tvol	21.259 ***	0.092 ***	-2.988 ***	0.632 ***	0.754	874.140 ***
	(41.044)	(6.832)	(-15.299)	(10.267)		
Panel B 微博情绪与换手率、成交量回归结果(周)						
Dtrd	-1.182	0.110	-6.975 ***	1.231 ***	0.733	170.468 ***
	(-0.455)	(1.479)	(-7.181)	(3.993)		
Dtrd ₁	0.617		-7.712 ***	1.038 ***	0.731	252.966 ***
	(0.268)		(-9.216)	(3.074)		
Tvol	0.718	0.108 **	-2.713 ***	0.810 ***	0.670	126.280 ***
	(0.485)	(2.548)	(-4.902)	(4.612)		

3.1.4 微博情绪对股票市场收益的解释 前文已经证实本文所使用的的数据是平稳的,均不存在单位根,可以直接构建 VAR 过程。本文借鉴国内外学者广泛使用的 AIC 信息准则、HQ 信息准则等方法确定了短期和中期 VAR 模型中的最优滞后阶数均为 4 阶。
VAR 模型回归的结果如表 4 所示,无论是短期还是中期(列(1)和列(3)),滞后一期的 MS 系数均在 5% 的水平上显著,说明滞后一期的短期和中期情绪均对股票市场收益具有正向预测能力。结合表 2 的结果,发现微博情绪高(低)会在一定程度上促使同期股票市场收益率的提高(降低),而且这种影响会持续到下一期,这在一定程度上说明,在我国,股票市场收益具有惯性效应的倾向。在控制了换手率、市盈率、账市比、规模、“日历效应”等因素之后(列(2)和列(4)),前述结论依然成立,而且模型的 R²显著提高。此外,还需要注意的是,中期微博情绪对股票市场的影响会在第三期发生反转(列(3)和列(4)),与 Da 等学者的研究结论一致。假设 3 得以证实。在表 4 中可以看出,市场收益除了受到微博情绪的影响之外,还受到历史信息的影响。

3.1.5 不同状态下的微博情绪对股票市场收益的解释 本节所使用的微博情绪数据与前文稍有不同。表 5 中,模型(1)使用的是交易日微博情绪,模型

变量为成交量时结论与 Panel A 的一致;因变量为换手率时,MS 的系数并不显著,但依然为正,为了进一步确认周微博情绪是否对周换手率存在影响,在模型 2 中剔除 MS 重新回归,发现调整后的 R²由 0.733 降低到 0.731,说明在模型中纳入 MS,可以提高回归的拟合优度。综合前述分析,假设 2 得到验证,微博情绪与股票市场流动性正相关。微博情绪越高,投资者对市场的预期越乐观,市场交易活动越活跃,市场流动性得以提高。

表 4 微博情绪对股票市场收益预测回归结果

解释变量	日		周	
	(1)	(2)	(3)	(4)
Cons	0.098 *	-1.121	0.273	11.494
	(1.700)	(-0.406)	(0.809)	(0.760)
MS(-1)	0.229 **	0.228 **	1.303 **	1.09 **
	(2.454)	(2.444)	(2.464)	(2.129)
MS(-2)	-0.075	-0.072	-0.060	0.192
	(-0.699)	(-0.679)	(-0.101)	(0.348)
MS(-3)	-0.028	-0.019	-2.593 ***	-2.363 ***
	(-0.260)	(-0.176)	(-4.404)	(-4.385)
MS(-4)	-0.042	-0.069	1.118 **	0.607
	(-0.455)	(-0.746)	(2.218)	(1.230)
Ret(-1)	0.057 *	0.044	0.133 *	0.020
	(1.650)	(1.225)	(1.744)	(0.239)
Ret(-2)	-0.050	-0.064	-0.014	-0.130
	(-1.440)	(-1.653 *)	(-0.175)	(-1.601)
Ret(-3)	0.018	-0.037	-0.007	-0.119
	(0.526)	(-0.963)	(-0.089)	(-1.484)
Ret(-4)	0.039	0.030	0.120	0.148 **
	(1.140)	(0.807)	(1.597)	(2.061)
控制变量	No	Yes	No	Yes
Adj-R ²	0.008	0.033	0.107	0.272
F	1.892 *	2.626 ***	3.703 ***	4.977 ***

(2)使用的是周六、周日以及节假日微博情绪,模型(3)使用的是同时包括交易日和非交易日的微博情绪,例如第 t 周股票市场交易五天,闭市两天,则模型

(1)、(2)、(3)分别使用的是五天微博情绪、两天微博情绪和七天微博情绪,具体的计算方法参见变量说明。在比较回归结果之后,我们发现,与交易日的回归结果不同,在非交易日状态下,滞后二期的MS系数在1%的水平上显著为正,达到了1.591,假设4得证。造成这种原因的结果可能是,投资者在交易日受工作时间限制,无暇时刻关注股市,所以会在周末及节假日对股市有一个较为长期的预期与判断。模型(3)中,使用交易日和非交易日微博数据进行预测,模型的拟合优度最高,但无论是哪种情况,中期微博情绪对股票市场收益的影响均会在第三期发生反转。

表 5 不同状态下微博情绪对股票市场收益预测回归结果

解释变量	模型(1)	模型(2)	模型(3)
Cons	11.494 (0.760)	10.109 (0.636)	8.087 (0.526)
MS(-1)	1.09 ** (2.129)	-0.140 (-0.267)	1.024 * (1.752)
MS(-2)	0.192 (0.348)	1.591 *** (2.829)	0.791 (1.243)
MS(-3)	-2.363 *** (-4.385)	-1.066 * (-1.929)	-2.784 *** (-4.517)
MS(-4)	0.607 (1.230)	-0.780 (-1.620)	0.580 (1.054)
Ret(-1)	0.020 (0.239)	0.058 (0.667)	0.023 (0.266)
Ret(-2)	-0.130 (-1.601)	-0.182 ** (-2.142)	-0.157 * (-1.924)
Ret(-3)	-0.119 (-1.484)	-0.165 * (-1.895)	-0.114 (-1.395)
Ret(-4)	0.148 ** (2.061)	0.163 ** (2.232)	0.154 ** (2.143)
控制变量	Yes	Yes	Yes
Adj-R ²	0.272	0.239	0.277
F	4.977 ***	4.316 ***	5.059 ***

3.2 神经网络分析 为了进一步检验微博情绪对股票市场的影响,通过建立非线性模型对上证综指的收盘价格进行预测。

本节采用 2016a 版的 MATLAB 中的神经网络工具模块进行分析。结合前文研究及相关学者的做法,本文选取前三期的上证综指的开盘价、最高价、最低价和收盘价共计 12 个变量作为当期收盘价格的输入变量,并以此构建基础预测模型。为了检验加入微博情绪是否可以提高对股价的预测精度,在基础模型中加入前一期的微博情绪(本文亦尝试在基础模型中同时加入滞后两期和滞后三期的微博情绪,结果显示,加入微博情绪的模型对股价的预测精度均有所提高)。共设计四个以收盘价为输出变量的预测模型,各个模型包含的输入变量具体如下:

MD0:前 3 日的上证综指开盘价、最高价、最低价和收盘价;

MD1:在 MD0 的基础上加入前一日的微博情绪;
MW0:前 3 周的上证综指开盘价、最高价、最低价和收盘价;

MW1:在 MW0 的基础上加入前一周的微博情绪。

在确定了输入变量和输出变量之后,进行网络参数的设定,选择 tansig 函数为传递函数,LM 算法为训练算法,其他基本参数都采用默认值。特别地,因为隐含层节点数的不同对网络的影响也存在很大的不同,因此,最佳隐含层节点数的设置利用经验公式(7)通过多次实验来确定:

$$l = \sqrt{m + n} + a$$

(7)

其中,l 代表隐含层节点个数,m 和 n 分别代表输入变量和输出变量的个数,a 为 0 到 10 之间的整数。结合本文构建的预测模型,隐含层节点数的范围为[4,14],经过多次比较网络的均方误差,确定模型 MD0、MD1、MK0 和 MK1 的最佳隐含层节点数分别为 6、11、9 和 12。

确定模型参数之后,选择 2013 年 3 月 1 日至 2016 年 9 月 30 日的上证综指数据作为训练样本,并对 2016 年 10 月 1 日到 2016 年 12 月 31 日之间的上证综指日/周收盘价进行预测,得到预测结果。并采用均方误差(MSE: Mean Squared Error)和平均绝对百分误差(MAPE: Mean Absolute Percentage Error)两个误差指标对预测结果进行评价,误差越小,代表预测值更为接近实际值,模型的预测能力则越好,结果如表 6。

表 6 预测结果评价

预测模型	MD0	MD1	MW0	MW1
MSE	620.72	501.78	3113.94	2311.03
MAPE	0.64%	0.60%	1.43%	1.29%

由表 6 的误差评价指标可知:加入微博情绪的模型,预测精度均有所提高,证明了短期内微博情绪对股价具有预测能力,与前文研究结果一致。

4 稳健性检验

为了保证本文结论的可靠性,对实证方法进行调整。调整方法后的实证结果与前文结果基本一致,说明本文研究结论基本稳健。具体的调整方法有以下几个方法:一是使用深证综指数据代替上证综指数据;二是调整微博情绪的计算方法为: $MS = \frac{1 + N}{1 + N + X}$; $MS = \frac{1 + N - X}{1 + N + X}$;三是在微博情绪对股票市场收益解释部分,在控制变量的选取上,用成交量代替换手率,结果显示,微博情绪系数的方向和显著性均未发生变化,但拟合优度却明显提高;四是考虑闭市之后的市场消息变化会在当期微博情绪反映,但此时微博情绪对股票

市场收益的影响只会在下一期反映,因此调整市场收益的计算方法为: $Ret = \log Closedex_t - \log Opendex_t$ 。

5 结论与建议

通过前文分析,本研究发现:微博情绪与股票市场收益、换手率及成交量均显著相关;微博情绪对股票市场流动性有显著的正向影响,微博情绪越高,市场交易活动越活跃,股票市场流动性越强;滞后一期的短期和中期微博情绪均对股票市场收益率有显著的预测能力,而且中期微博情绪的预测能力会在第三期发生反转;交易日微博情绪和非交易日微博情绪对股票市场收益率的预测能力不同。神经网络预测也证实了,纳入微博情绪的模型预测结果误差更小,进一步说明微博情绪对股票市场存在正向影响。

本研究的主要贡献在于:首次利用微博数据中心提供的“微指数”平台挖掘相关信息,并以此来衡量投资者情绪,并且验证了微博情绪作为投资者情绪代理变量的合理性;分析了微博情绪影响股票市场的渠道;从向量自回归(线性)和神经网络(非线性)两个角度同时验证了微博情绪对股票市场的影响。

文本挖掘及语义分析技术虽已日臻成熟,但在微博舆情研究方面,研究数据完全依赖于微博提供的部分 API 接口,且对用户设有访问权限,需要繁琐的步骤设置。在这种情况下,微指数的优势就凸显出来,其可以提供更大的数据量,时效性更强,易于获取,可以获得更加全面、真实的投资者情绪信息,对于学术研究来说,有重大的应用价值。而在实践方面,机构投资者则可以通过挖掘微博数据构建相关指标以预测散户交易行为,并纳入实际的投资组合中,辅助其进行更高效的投资决策。对于证监会等监管部门而言,若能捕捉微博数据中的股票市场信息,建立相应的微博舆情监控系统,则能够有效改善、优化市场监管,促进股票市场平稳健康的运行。

本研究采用“牛市”和“熊市”两个关键词在微博上出现的频率进行分析,数据存在一定的噪音。在后续研究中,将通过挖掘更纯净的数据和更高效的方法来构建投资者情绪,以期进一步提高股价预测精度。

参考文献

- [1] Karabulut Y. Can facebook predict stock market activity? [J]. Ssrn Electronic Journal, 2011.
- [2] Bollen J, Mao H, Zeng X. Twitter mood predicts the stock market[J]. Computer Science, 2010, 2(1):1-8.
- [3] Da Z, Engelberg J, Gao P. In search of attention[J]. The Journal of Finance, 2011, 66(5):1461-1499.
- [4] 宋双杰, 曹晖, 杨坤. 投资者关注与 IPO 异象——来自网络搜索量的经验证据[J]. 经济研究, 2011(S1):145-155.
- [5] 俞庆进, 张兵. 投资者有限关注与股票收益——以百度指数作为关注度的一项实证研究[J]. 金融研究, 2012(8):152-165.
- [6] 周胜臣, 施询之, 瞿文婷, 等. 基于微博搜索和 SVM 的股市时间序列预测研究[J]. 计算机与现代化, 2013(4):22-26.
- [7] 程琬芸, 林杰. 社交媒体的投资者涨跌情绪与证券市场指数[J]. 管理科学, 2013(5):111-119.
- [8] 赖凯声, 陈浩, 钱卫宁, 等. 微博情绪与中国股市:基于协整分析[J]. 系统科学与数学, 2014, 34(5):565-575.
- [9] 黄润鹏, 左文明, 毕凌燕. 基于微博情绪信息的股票市场预测[J]. 管理工程学报, 2015, 29(1):47-52.
- [10] 苏志, 张骐. 基于网络大数据的社会经济预测综述[J]. 情报杂志, 2015(7):18-21.
- [11] Baker M, Wurgler J. Investor sentiment and the cross-section of stock returns[J]. Journal of Finance, 2006, 61(4):1645-1680.
- [12] Thaler R H. Advances in behavioral finance[M]. Advances in behavioral finance. Russell Sage Foundation, 1993.
- [13] Engelberg J, Sasseville C, Williams J. Market madness? The case of mad money[J]. Social Science Electronic Publishing, 2010, 58(2):351-364.
- [14] 饶育蕾, 彭叠峰, 成大超. 媒体注意力会引起股票的异常收益吗?——来自中国股票市场的经验证据[J]. 系统工程理论与实践, 2010, 30(2):287-297.
- [15] 王勇, 杨庆运. 我国网络关注度对股票收益的影响——基于和讯关注度的实证研究[J]. 投资研究, 2014(2):143-149.
- [16] Louis H, Sun A X. Abnormal accruals and managerial intent: Evidence from the timing of merger announcements and completions[J]. Ssrn Electronic Journal, 2015, 33(3):1101-1135.
- [17] 谭伟强. 我国股市盈余公告的“周历效应”与“集中公告效应”研究[J]. 金融研究, 2008(2):152-167.
- [18] 权小锋, 吴世农. 投资者关注、盈余公告效应与管理层公告择机[J]. 金融研究, 2010(11):90-107.
- [19] 董颖红, 陈浩, 赖凯声, 等. 博客基本社会情绪的测量及效度检验[J]. 心理科学, 2015(5):1141-1146.
- [20] Mao H, Counts S, Bollen J. Quantifying the effects of online bullishness on international financial markets[J]. Statistics Paper, 2015.
- [21] Zhou S, Shi X, Sun Y, et al. Stock market prediction using heat of related keywords on micro blog[J]. Journal of Software Engineering & Applications, 2013, 6(3):37-41.
- [22] Kamstra M J, Levi M D. Winter blues: A SAD stock market cycle[J]. American Economic Review, 2002, 93(1):324-343.

(责编:贺小利;校对:王平军)