

# 互联网异质性财经新闻对股市的影响

——来自中国互联网数据与上市公司的证据

刘海飞,许金涛

(南京大学 工程管理学院,江苏 南京 210093)

**摘要:** 结合文本挖掘技术和事件研究方法,分析互联网异质性财经新闻对中国股票短期异常收益的影响。研究发现:五种互联网异质性新闻均会引起股票短期的异常收益,其中,政策扶持类、兼收并购类、再融资类和盈利能力类新闻能够对公司股票产生正的异常收益,而违规处罚类新闻能够对公司股票产生负的异常收益。研究结论为监管机构、财经新闻媒体工作人员、上市公司、各类投资者在信息辨析与传播、交易、监管等方面提供了启示与证据。

**关键词:** 互联网新闻;异质性;文本挖掘;事件研究;异常收益率;媒体关注

中图分类号:F830 文献标识码:A 文章编号:1671-9301(2017)01-0076-13

DOI:10.13269/j.cnki.ier.2017.01.007

## 一、引言

计算机科学和信息通信技术推动了互联网行业的迅猛发展,由此建立的“互联网+”新型行业模式开始如火如荼地发展起来。与此同时,在利率市场化、汇率市场化和金融管制放松等一系列金融改革政策的驱动下,互联网金融孕育而生。互联网金融信息对金融市场的影响已经越发显现。刘海二<sup>[1]</sup>研究发现,对市场主体来说,互联网金融可以降低信息不对称程度。2015年3月,两会的政府工作报告中首次提出并正式确立“互联网+”计划的国家战略地位。随后股市的首个交易日,电商股、创投概念股集体飙升。截止到当天收盘,创业板指数达到2 142.67点,涨幅达到3.56%,其中共有28只创业板股票涨停。2015年12月,在第二届世界互联网大会即将开幕之际,概念股闻风而动,互联网产业掀起新一轮投资热潮。随着我国金融市场改革的不断深入,以及投资类型和投资风格的不断变动,准确判断不同异质性的财经新闻对上市公司的影响,对监管者、上市公司和投资者来说都具有重要的理论和现实意义。

然而,现有的研究体系中缺乏这方面的研究,主要原因是难以对文本信息进行科学的定量分析,无法从实证的角度展开互联网财经新闻对中国股市的影响研究。目前,随着科学技术的发展,将文本挖掘技术应用于金融市场的研究已成为一种趋势,因此可以考虑使用文本挖掘方法对互联网财经新闻进行采集、清洗、分类和研究。这有助于充分挖掘股票的价值和公司的潜力,帮助研究者和投资者更好地利用金融信息,深入系统地分析互联网异质性新闻文本信息对中国股市的影响。

收稿日期:2016-03-24;修回日期:2016-11-28

作者简介:刘海飞(1980—),男,安徽阜阳人,管理学博士,南京大学工程管理学院副教授,硕士生导师,研究方向为金融工程、行为金融、计算实验金融、金融大数据,以及交叉学科理论在资本市场与商业银行中的应用;许金涛(1991—),男,安徽蚌埠人,南京大学工程管理学院硕士研究生,研究方向为金融工程、数据挖掘及复杂网络。

基金项目:国家自然科学基金研究项目(71101068);江苏省自然科学基金面上项目(BK20161398);江苏省金融工程重点实验室项目(NSK2015-09)

本文将从政策扶持类、兼收并购类、再融资类、盈利能力类和违规处罚类五个方面区分互联网财经新闻,研究互联网异质性财经新闻对我国股市的影响。与现有研究相比,本文的贡献在于通过文本挖掘技术区分互联网财经新闻的异质性,在互联网财经新闻异质性下探究上市公司的股价运行机理,展开理论与实证相结合的分析。这有助于全面揭示中国金融市场异象,为监管者更好地制定监管政策和信息披露机制提供理论依据,为上市公司维护股价稳定和公司平稳发展提供新的启示,同时为投资者进行科学的投资决策提供指导性建议。

## 二、国内外文献回顾与评述

随着证券市场的不断发展和互联网信息平台的快速崛起,新闻信息对股市的影响成为近几年金融领域学者们关注和研究的热点,其相关研究的理论基础是市场有效性理论<sup>[2]</sup>,将资本资产定价模型中市场层面无法解释的公司层面和行为层面的信息视作异质性信息。学术界已有研究表明,互联网热点新闻对股价波动是有显著影响的<sup>[3-5]</sup>。依据研究的递进关系,下面将从财经新闻有效性、媒体关注有效性、异质性新闻有效性和文本挖掘的有效性等方面进行回顾。

**财经新闻有效性研究** Cutler *et al.*<sup>[6]</sup>将宏观经济变量进行向量自回归所得到的残差作为新闻因子,使用统计回归方法研究新闻是否会影响股票价格,研究结果表明股票市场会对新闻报道做出反应。Klibanoff *et al.*<sup>[7]</sup>研究新闻标题是否会影响到封闭式国家基金的波动。Chan<sup>[8]</sup>使用 Fama-French 三因子模型,研究发现上市公司发生负面新闻时对股票价格的影响比较大。Mitchell and Mulherin<sup>[9]</sup>使用道琼斯的新闻和公告研究新闻及公告的数量是否会影响到股票收益。杨继东<sup>[10]</sup>发现新闻媒体会通过影响投资者情绪及对收益预期的方式间接地影响到股票的价格。Li *et al.*<sup>[11]</sup>运用媒体意识量化交易策略,捕捉投资者情绪对股价的影响,同时运用实证分析,验证了新闻情绪对股价的影响。朱波和孙鹏阁等<sup>[12]</sup>研究表明,投资者情绪对股票收益有着显著效应,难以估值的股票对投资者情绪较为敏感,财务新闻质量的提高能减弱这类股票的投资者情绪效应,促进投资者理性投资。

**媒体关注有效性研究** Wysocki<sup>[13]</sup>对超过 90 万条的相关股评进行分析,发现当天股评的数量会对第二天的股票交易量及收益率产生影响。Huberman<sup>[14]</sup>发现如果上市公司利好消息被媒体重复报道会使公司股票存在很大程度的上涨并在一段时间内保持高位。Takeda and Yamazaki<sup>[15]</sup>采用了事件研究法研究发现所有的纪录片都对股价产生了不同程度的影响,涉及经营现状和未来发展时,这种影响更为明显。Zimbra *et al.*<sup>[16]</sup>发现相关股评可以反映出投资者的情绪和决策,可以以此对相关个股的股价进行预测。赵伟和梁循<sup>[17]</sup>发现当新闻数量明显增长时,相关股票的收益率也会产生明显的波动,并且其他会导致收益率波动的随机影响因素会随着新闻数量增长而逐渐消失。饶育蕾和王攀<sup>[18]</sup>研究发现新股的百度指数越高,则其发行价就越高。饶育蕾和彭叠峰等<sup>[19]</sup>研究媒体关注度对股票月收益的影响,发现媒体关注度对股票月收益率有负向的影响。Sprenger *et al.*<sup>[20]</sup>通过对微博论坛上与股票相关的信息运用计算语言的方法,论证了微博情绪与股票收益、信息量和交易量、争论和波动率之间存在关联,同时还分析了导致微博论坛信息有效聚合的机制。张正勇<sup>[21]</sup>发现媒体关注的压力与企业自愿披露 CSR 报告的程度具有显著的正相关关系,促使投资者能获取更多公司信息。

**异质性新闻有效性研究** Niederhoffer<sup>[22]</sup>利用计量经济学模型,将新闻变量作为哑变量,收集了 1950 年至 1966 年发表在《纽约时报》上的 432 篇重大新闻,他还根据新闻的标题内容,将新闻人工分为了 20 类,将每篇新闻按对股市的影响程度进行评级,使用统计的方法研究新闻发布后市场指数的涨跌幅情况,从而研究股票市场受哪些重大新闻事件的影响以及不同类别的新闻在发布后第一天对市场的影响情况。他发现在重大新闻发布的当天,股票市场的反应最强烈,并且第二天市场指数的变化趋势往往依赖第一天市场指数的变化情况。Keown and Pinkerton<sup>[23]</sup>从并购类的新闻事件入手,研究新闻公布前后股票收益率的变化。发现该类新闻确实会对公司股价产生影响,并且指出该影响共持续 12 天,最早在新闻公布 11 天前就产生影响,并且占了整个定义事件窗内累积异常收益的

50%。Li *et al.*<sup>[24]</sup>通过定量媒体意识交易策略来调查媒体对股市的影响,发现媒体对公司的影响根据公司特点和新闻内容的不同而各不相同。

文本挖掘技术的有效性研究 Wüthrich *et al.*<sup>[25]</sup>通过关键字提取技术对股票的走势进行预测。Tan *et al.*<sup>[26]</sup>提出了两阶段模型,将文本挖掘分为文本求精和知识提取。Lavrenko *et al.*<sup>[27]</sup>通过语言模型来判断新闻对股票价格走势的影响。Schumaker and Chen<sup>[28]</sup>利用支持向量机建立回归模型分析新闻信息与公布 20 分钟后股票价格的关系。Antweiler and Frank<sup>[3]</sup>对华尔街日报的 25 万多条新闻进行分类,并且使用事件法研究每一类新闻对股市的影响,得出了与“有效市场假说”相悖的结论。Lægreid and Sandal<sup>[29]</sup>研究出了能够从论坛中提取投资者观点并进行相关的股票价格预测的算法。Tetlock<sup>[30]</sup>分析新闻信息对交易量的影响,发现新闻报道悲观程度越高或者越低,都会对股票的交易量带来很大的影响。Tetlock *et al.*<sup>[31]</sup>又利用 Harvard-IV-4 词库对新闻信息进行分类,分为正面信息和负面信息。陈华和梁循<sup>[32]</sup>利用文本挖掘技术进行股票新闻板块的归类,并利用简单词频统计方法研究该篇个股新闻对其他个股是否会产生交互的影响。Smailović *et al.*<sup>[33]</sup>运用支持向量机的情绪分类器对推特新闻进行文本挖掘研究公众情绪与股票价格之间的关系。

综上,关于财经新闻对股市影响的研究,国内外学者重点关注小样本范围内某一类财经新闻的有效性研究,很少涉及新闻内容的深度挖掘。因此,有必要将文本挖掘技术引入互联网财经新闻的研究中,解决样本量不足和分类不全面等问题。本文将在现代金融系统理论和已有研究成果的背景下,通过文本挖掘技术和金融学研究方法系统地互联网财经新闻进行异质性划分,研究异质性财经新闻对中国股市的影响程度和持续期,提供中国市场的证据,这也是本文的贡献所在。

### 三、研究假设的提出

结合上市公司的内部因素和外部因素、主观需求和客观响应,同时兼顾事件的有效性和数据的可获取性,本文主要研究政策扶持类、兼收并购类、再融资类、盈利能力类和违规处罚类五种新闻对上市公司股价的影响。其中,政策扶持类新闻是指媒体报道政府发布与上市公司所属行业直接或间接相关的政策;兼收并购类新闻是指媒体报道上市公司发布与其他公司兼并重组类的公告;再融资类新闻是指媒体报道上市公司将进行再融资扩大规模的消息;盈利能力类新闻是指媒体报道上市公司目前具有较强的盈利能力的相关新闻;违规处罚类新闻是指媒体报道上市公司面临证监会或国家相关部门处罚的新闻。下面将结合相关文献和理论分析提出相关研究假设。

当今社会,国家的宏观政策对经济发展具有重要的引导作用。当政府决定扶持某一行业时,往往会通过降低税收、提供补贴、贷款优惠及开放出口等政策使该行业各类成本降低,进而增强了竞争实力,提高了可持续发展水平,因此股价自然会受影响上涨;政府出于对环境保护和宏观调控的要求,有时也会对某些行业制定各类限制政策,政府的限制规定会对企业业绩产生一定的负向影响。因此,投资者为了进行科学的投资决策,会非常关注国家政策的变动。Bonaparte and Kumar<sup>[34]</sup>发现,政治主义积极性越强信息收集成本越低并且市场参与率较高。Apergis<sup>[35]</sup>针对美国的股市进行了实证研究,运用 TVP-SVM 模型证明了在金融危机后劳务税收、国家财政支出和货币政策对股票收益有很明显的影响。因此,本文提出第一个假设:

假设 1: 政策扶持类的新闻会对公司股票产生异常收益率。

假设 1.1: 政策扶持类的新闻会对公司股票产生正的异常收益率。

兼收并购是指两家或多家独立的公司合并组成一家公司,通常由一家占优势的企业通过运用产权交易的方式获得对另一家企业的控制权或者另一家企业的资产所有权。兼收并购行为通过彻底改变公司的经营模式使公司的内在价值发生变化,进而对上市公司的股价产生影响。Keown and Pinkerton<sup>[23]</sup>从并购类事件出发,利用事件研究法研究了公司并购消息发布前后股票收益的变化,发现在信息正式发布之前 11 天就产生了异常收益,并且在公布之前所获得的异常收益总和大约占了

整个事件窗口期所获得的异常收益的 50%。因此,本文提出第二个假设:

假设 2: 兼收并购类的新闻会对公司股票产生异常收益率。

假设 2.1: 兼收并购类的新闻会提前对公司股票产生异常收益率。

企业发展过程中需要持续的资金支持,上市之后的后续发展在很大程度上取决于其融资能力。再融资作为上市公司的重大战略之一,它的制定和实施都会对公司股票的内在价值产生明显的影响。Asquith and Mullins<sup>[36]</sup>研究了美国股票市场上 1963 到 1981 年间的再融资样本,结果显示再融资消息公布当天,工业类股票的 CAAR 出现了 3% 的下跌,公用事业类股票的 CAAR 出现 0.9% 的下跌。同样的研究方法下,Masulis and Korwar<sup>[37]</sup>发现工业类股票的 CAAR 出现了 3.25% 的下跌,公用事业类股票的 CAAR 出现 0.68% 的下跌。Mann and Slicherman<sup>[38]</sup>以 1982 到 1984 年间 241 起再融资事件为样本,也同样发现相关的 CAAR 为 -2.64%。因此,本文提出第三个假设:

假设 3: 再融资类的新闻会对公司股票产生异常收益率。

假设 3.1: 再融资类的新闻会对公司股票产生负的异常收益率。

盈利能力反映了一定时期内上市公司的生产经营状况、资源配置效率及可持续发展水平。盈利能力在很大程度上决定了投资者投资收益的高低,因此投资者都会密切关注上市公司盈利能力。Chan<sup>[8]</sup>发现投资者对公开的公司信息会过于不理睬,而更在意关于公司的更加私人的信息,这些信息会影响投资者投资行为。因此,本文提出第四个假设:

假设 4: 盈利能力类的新闻会对公司股票产生异常收益率。

上市公司的负面新闻对股价的影响也越来越明显。从 2008 年三鹿牌婴幼儿奶粉事件导致乳制品行业股价全线大跌到近几年央视 3.15 晚会曝光的问题企业股价受到明显影响,说明此类违规处罚对股价的影响是显而易见的。除了产品质量问题,证券市场上的违规操作受到证监会或国家相关部门的处罚也会使股价有明显的异常波动,例如 2013 年的光大证券乌龙指事件,光大证券由于违规操作扰乱了证券市场受到了证监会的处罚,处罚前后光大证券的股价暴跌。因此,有必要将上市公司各类违规处罚的负面新闻分出一类并讨论其对上市公司股价的影响。因此,本文提出第五个假设:

假设 5: 违规处罚类的新闻会对公司股票产生异常收益率。

假设 5.1: 违规处罚类的新闻会对公司股票产生负的异常收益率。

#### 四、研究设计

##### (一) 数据采集与处理

本研究所需数据主要包括上市公司股票数据和互联网财经新闻数据。在上市公司股票数据方面,本文从 CCER 数据库中采集了 2014 年 1 月 1 日至 12 月 31 日的全 A 指数及其成份股的前复权日收盘价数据,共得到全 A 指数和 2 927 只股票的 245 个交易日数据。同时,在互联网财经新闻数据方面,为保证数据的准确性、有效性和可操作性,本文通过编写网络爬虫程序从和讯网的财经专栏上采集所有股票在 2014 年 1 月 1 日至 12 月 31 日全年的财经新闻数据,抓取的主要信息包括新闻发布时间、新闻标题、新闻正文、股票名称、股票代码等内容。特别地,为了使搜集到的公司新闻与该公司的股票交易数据更加匹配,正文中应至少有两次提到该公司的简称,因此共得到 2 927 家公司的 128 816 篇财经新闻。

然而,由于股票存在长期停牌和 ST 等现象,财经新闻也存在大量低质量数据,无法直接使用这些数据进行研究,因此有必要对数据进行清洗和筛选,主要步骤如下:(1) 因创业板中大部分公司成立时间短、规模较小且投资风险大,所以剔除创业板所有股票及其相关新闻数据;(2) 因 ST 和 \*ST 类股票的涨跌幅有严格的限制而无法及时反映事件对股价的影响,所以剔除 ST、\*ST、暂停上市和退市的股票及其新闻;(3) 剔除各上市公司简报公告类相关的新闻数据;(4) 删除重复报道的新闻,保留最早发布的新闻;(5) 为了消除不同新闻对上市公司股价的重叠影响,只保留每个月内只有一条新闻

的上市公司作为研究对象,这一步骤剔除了绝大部分样本;(6)为保证事件研究窗口的有效性,剔除连续停牌超过一个月的股票及其新闻。经过以上数据清洗后得到 599 家上市公司的 1 165 篇财经新闻。

(二) 财经新闻异质性识别

为了研究互联网异质性财经新闻对股价的影响,本文将基于文本分类技术对财经新闻进行分类。分类的主要步骤如下:

1. 分词处理 本文采用北京大学计算语言研究所提供的分词方法,使用 NLPPIR/ICTCLAS 2015 分词软件进行文本分词。

2. 文本清洗 去除分词中的停用词、独立词和高频词等,通过语义分析和同义项合并等方法依据分词间的相似度进行同类项合并,确定文本特征项。

3. 文本表示 通过向量空间模型将文本特征项表示为特征向量,并基于 TF-IDF 函数计算特征词的权重。TF-IDF 一般采用如下公式计算权重:

$$w_i(t, d) = \frac{tf(t, d) \times \log(N/n_t + 0.01)}{\sqrt{\sum_{t \in d} [tf(t, d) \times \log(N/n_t + 0.01)]^2}}$$

其中  $w_i(t, d)$  表示文档  $d$  中第  $i$  个特征项  $t$  的权重,  $N$  表示文档集中的文档数目,  $n_t$  表示文档集中出现  $t$  的文档数目。

4. 特征提取 为降低文本向量的维数,本文采用信息增益(IG) 文本特征选择法对本文特征项进行提取。信息增益的计算公式如下:

$$\begin{aligned} IG(t) &= H(C) - H(C|t) \\ &= - \sum_{c \in C} p(c) \log(p(c)) + p(t) \sum_{c \in C} p(c|t) \log(p(c|t)) + p(\bar{t}) \sum_{c \in C} p(c|\bar{t}) \log(p(c|\bar{t})) \\ &= \sum_{c \in C} \left[ p(c|t) \log\left(\frac{p(c|t)}{p(c)p(t)}\right) + p(c|\bar{t}) \log\left(\frac{p(c|\bar{t})}{p(c)p(\bar{t})}\right) \right] \end{aligned}$$

其中  $IG(t)$  表示特征词  $t$  的信息增益值,  $c$  表示文本类变量,  $C$  表示文本类的集合且  $C = (c_1, c_2, \dots, c_n)$   $H(C)$  表示的是在没有特征词  $t$  之前一个随机文本对分类结果的不确定程度。 $H(C|t)$  表示为获得特征词  $t$  后随机文本对分类结果的不确定性程度。

5. 文本分类 根据预定义的类别,手动选择一部分已知类别的样本利用支持向量机(SVM<sup>①</sup>) 训练分类模型,然后运用训练好的分类器模型对剩余样本进行分类。分类结果如表 1 所示。同时,经随机挑选部分新闻样本进行人工验证发现识别的效果较好,可以满足本文的研究需求。

表 1 互联网异质性新闻的分类识别结果统计表

新闻类别	新闻数量	比例
政策扶持类	205	18%
兼收并购类	132	11%
再融资类	255	22%
盈利能力类	347	30%
违规处罚类	226	19%
汇总	1 165	100%

(三) 异质性新闻有效性模型

本节将构建基于事件研究法的互联网异质性财经新闻对上市公司影响的研究模型。事件研究法(Event Study Method) 依据研究目的选取某一特定事件来研究其发生前后股价是否会产生异常波动,从而解释该事件对样本股票的价格和收益率变化产生的影响<sup>[39]</sup> 其理论基础是有效市场假说<sup>[2]</sup>。国内外大量研究发现中国资本市场存在弱式有效性<sup>[40-41]</sup> 这为本文开展互联网异质性新闻对上市公司股价影响的有效性检验提供了理论支撑。

事件研究法需要观察事件发生前后股票是否产生异常收益率并检验异常收益率的显著性,从而说明事件对股票价格变动是否产生了影响,因此需要对事件发生的前后情况进行定义,此处引入估计窗和事件窗。同时,研究中通常将事件发生的日期(时刻) 包括在事件窗内,并设定估计窗比事件窗要长<sup>[42]</sup>。本文将新闻报道日定为第 0 日,将估计窗设为  $[-100, -15]$  共 86 天,将事件窗设为

$[-14, 15]$ 共30天,如图1所示。

异质性新闻有效性检验的关键是确定股票的预期正常收益率。预期正常收益率的估计通常采用两种模型,一是常均值收益模型,假定某支股票的平均收益在整个研究时段不变;另一个是市场模型,假定某支股票的收益率与市场收益率有稳定的线性关系。本文采用的是市场模型,将股票收益分解为来自市场层面的共同收益与来自公司层面的特有收益,通过去除与市场组合收益变化相关的收益部分,降低正常收益的方差,从而增强检测事件效应的有效性<sup>[43]</sup>。即:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t}$$

其中  $R_{i,t}$  表示  $t$  期第  $i$  只股票的收益率;  $R_{m,t}$  表示  $t$  期市场的收益率;  $\alpha_i$  表示第  $i$  只股票的超额收益率;  $\beta_i$  表示第  $i$  只股票的系统性风险;  $\varepsilon_{i,t}$  表示残差项且  $\varepsilon_{i,t} \sim N(0, \sigma_i^2)$ 。则预期的正常收益率为:

$$\hat{R}_{i,t} = \alpha_i + \beta_i R_{m,t}$$

当事件发生后,股票价格因受到事件影响而产生与预期正常收益偏离的波动,波动的差值定义为异常收益率。因此,异常收益率  $AR$  的计算公式为:

$$AR_{i,t} = R_{i,t} - \hat{R}_{i,t}$$

从而通过对异常收益率累加,可以得到累计异常收益率  $CAR$ ,即:

$$CAR_{i,t} = \sum_{k=t_0}^t AR_{i,k}$$

为进一步度量不同类型互联网财经新闻对股票价格的影响,此处引入平均异常收益率和平均累计异常收益率。其中,平均异常收益率为同期同类新闻异常收益的平均值,平均累计异常收益率为同期平均异常收益率的累计值。因此,平均异常收益率  $AAR$  和平均累计异常收益率  $CAAR$  分别可表示为:

$$AAR_t = \frac{1}{n} \sum_{i=1}^n AR_{i,t}$$

$$CAAR_{i,t} = \sum_{k=t_0}^t AAR_{i,k}$$

其中  $n$  表示第  $i$  只股票同期所属异质性新闻样本的数量。同时,为了验证上市变化是否由股价随机波动引起的,需要对指标进行显著性检验,即检验  $AAR_t$  和  $CAAR_t$  与 0 是否存在显著性差异。因此,构造的  $T$  统计量为:

$$t_{AAR_t} = \frac{AAR_t}{\sigma(AAR_t) / \sqrt{n}}$$

$$t_{CAAR_t} = \frac{CAAR_t}{\sigma(CAAR_t) / \sqrt{n}}$$

其中  $\sigma(x)$  表示  $x$  的标准差,且  $t_{AAR_t}$  和  $t_{CAAR_t}$  服从自由度为  $n-1$  的  $t$  分布。在给定的置信水平  $\theta$  (通常取 0.05) 下,当统计结果无法拒绝原假设时,有  $t_{AAR_t}$  (或  $t_{CAAR_t}$ ) 与 0 无异,表明该类异质性新闻对股票价格没有显著性影响;当统计结果拒绝原假设时,有  $t_{AAR_t}$  (或  $t_{CAAR_t}$ ) 与 0 存在显著性差异,若  $CAAR_t > 0$  表明该类异质性新闻对股价有显著正影响,若  $CAAR_t < 0$  表明该类异质性新闻对股价有显著负影响。

## 五、实证分析与显著性检验

本节将开展互联网异质性新闻对股票价格影响的有效性研究。通过前文中对所采集的财经新闻异质性识别可知,互联网异质性新闻可分为 5 类,下面将分别针对每种异质性财经新闻研究其对股票价

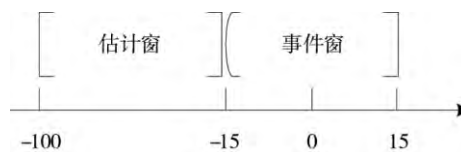


图1 事件研究法的研究窗划分示意图

格的影响及其显著性。本文选取全 A 指数作为市场指数,对每只股票在事件发生前的估计窗  $[-100, -15]$  内计算超额收益率  $\alpha_i$  和系统性风险  $\beta_i$ , 并以此计算每类新闻从估计窗到事件窗  $[-14, 15]$  的平均异常收益率 AAR 和平均累计异常收益率 CAAR, 同时对异常收益率的显著性进行检验。

(一) 政策扶持类新闻

本小节将对政策扶持类新闻对股票价格的影响进行研究,共包含 205 个样本,基于表 1 的统计数据,对政策扶持类新闻事件进行异常收益率的计算及其显著性检验,得到表 2 所示的统计结果。可以发现,平均异常收益率 AAR 在事件发生前 2 日内都是显著为正的,事件发生后 6 日内依然明显异于 0,从而使得平均累计异常收益率 CAAR 在事件发生前 1 日到发生后 12 日均是显著为正的。

同时,可以绘制政策扶持类新闻事件的异常收益率走势图,如图 2 所示。可以发现,平均累计异常收益率曲线呈现先快速上升后逐步下降的趋势,显著的平均异常收益率主要集中在  $[-2, 0]$  和  $[3, 6]$  之间,其中,  $[-2, 0]$  是正的异常收益率的主要贡献者,  $[3, 6]$  间的异常收益率有正有负,没有贡献。

因此,通过对上述政策扶持类新闻事件有效性的分析,可以得到以下结论:

- (1) 政策扶持类新闻对股票价格确实存在显著的影响,证明假设 1 成立;
- (2) 政策扶持类新闻对股票价格的影响主要集中在新闻发布的前 3 日至后 6 日;
- (3) 政策扶持类新闻对股票价格的影响是积极的,使得股票的收益率明显增加,可以看作是利好的信号;
- (4) 若投资者能在政策扶持类新闻发布前提前获知信息,并最晚在发布前 3 日买入股票,在新闻发布日后两日内卖出便可获得全程的超额收益。当然,也可以在回调后买入持有 1 ~ 2 日也可获得一定的超额收益。

(二) 兼收并购类新闻

本小节将对兼收并购类新闻对股票价格的影响进行研究,共包含 132 个样本,基于表 1 的统计数据,对兼收并购类新闻事件进行异常收益率的计算及其显著性检验,得到表 3 所示的统计结果。可以发现,平均异常收益率 AAR 在事件发生前 2 日内都是明显异于 0 的,事件发生当日存在 9% 以上的显著正收益,其后异常收益主要呈现负的情况,从而使得平均累计异常收益率 CAAR 在事件发生当

表 2 政策扶持类新闻事件的异常收益率检验统计表

事件窗	AAR	CAAR	事件窗	AAR	CAAR	事件窗	AAR	CAAR
-14	0.15 (-0.976)	0.15 (0.976)	-4	0.131 (0.533)	0.732 (0.995)	6	0.418** (2.464)	2.580*** (3.152)
-13	0.086 (0.498)	0.237 (0.977)	-3	-0.079 (-0.328)	0.654 (0.884)	7	-0.247 (-1.555)	2.333*** (2.867)
-12	0.11 (0.504)	0.346 (1.01)	-2	0.628** (2.546)	1.282 (1.469)	8	-0.247 (-1.555)	2.086** (2.477)
-11	0.028 (0.135)	0.374 (1.001)	-1	0.694*** (3.283)	1.976** (2.359)	9	-0.027 (-0.121)	2.059** (2.459)
-10	0.303 (1.285)	0.677 (1.374)	0	0.818*** (3.989)	2.794*** (3.712)	10	-0.47 (-2.076)	1.589** (2.066)
-9	-0.127 (-0.553)	0.551 (1.066)	1	-0.008 (-0.035)	2.786*** (3.531)	11	-0.182 (-0.823)	1.406* (1.905)
-8	-0.256 (-1.242)	0.294 (0.482)	2	0.038 (0.17)	2.824*** (3.445)	12	0.113 (0.474)	1.519* (1.886)
-7	0.286* (1.738)	0.581 (1.049)	3	-0.389* (-1.698)	2.435*** (3.081)	13	-0.391* (-1.895)	1.128 (1.432)
-6	-0.258 (-1.387)	0.323 (0.533)	4	-0.420* (-1.758)	2.015*** (2.733)	14	-0.118 (-0.7)	1.01 (1.271)
-5	0.279 (1.349)	0.601 (0.905)	5	0.147 (0.693)	2.161*** (2.774)	15	-0.097 (-0.542)	0.914 (1.173)

注: \* 表示  $p < 0.1$ , \*\* 表示  $p < 0.05$ , \*\*\* 表示  $p < 0.01$ , 括号内为 T 检验的统计值。

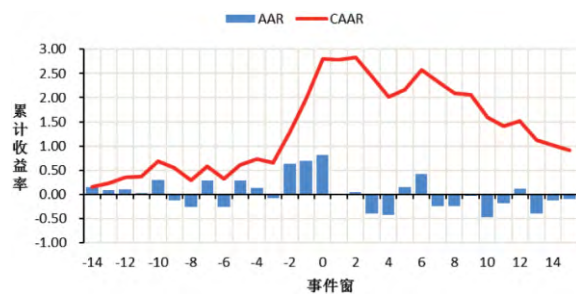


图 2 政策扶持类新闻事件的异常收益率走势图

日至后 2 日是显著为正的。

同时,可以绘制兼收并购类新闻事件的异常收益率走势图,如图 3 所示。可以发现,平均累计异常收益率曲线呈现先快速上升后快速下降的趋势,显著的平均异常收益率主要集中在 [-2, 2] 之间。同时,正的异常收益率主要位于事件发生当日及以前,而负的异常收益率主要位于事件发生之后。

表 3 兼收并购类新闻事件的异常收益率检验统计表

事件窗	AAR	CAAR	事件窗	AAR	CAAR	事件窗	AAR	CAAR
-14	0.32 (1.303)	0.32 (1.303)	-4	0.399 (1.033)	0.445 (0.617)	6	0.051 (0.258)	0.765 (1.215)
-13	0.094 (0.376)	0.414 (1.274)	-3	0.145 (0.456)	0.591 (0.681)	7	-0.354* (-1.891)	0.412 (0.804)
-12	-0.051 (-0.188)	0.363 (1.067)	-2	0.460* (1.705)	1.052 (1.08)	8	-0.354* (-1.891)	0.057 (0.401)
-11	-0.21 (-0.894)	0.154 (0.719)	-1	0.424* (1.666)	1.476 (1.528)	9	-0.23 (-0.941)	-0.174 (0.212)
-10	0.217 (0.711)	0.371 (1.003)	0	0.909*** (3.728)	2.385*** (3.079)	10	-0.357 (-1.143)	-0.532 (-0.023)
-9	-0.066 (-0.238)	0.304 (0.825)	1	-0.269 (-0.989)	2.116** (2.555)	11	0.301 (1.076)	-0.23 (0.147)
-8	-0.247 (-1.208)	0.057 (0.282)	2	-0.741*** (-3.209)	1.375* (1.746)	12	0.073 (0.357)	-0.158 (0.192)
-7	-0.138 (-0.682)	-0.083 (-0.002)	3	-0.307 (-1.17)	1.068 (1.482)	13	-0.084 (-0.529)	-0.241 (0.111)
-6	0.267 (1.224)	0.186 (0.448)	4	-0.482 (-1.668)	0.586 (1.097)	14	-0.242 (-1.574)	-0.484 (-0.136)
-5	-0.138 (-0.7)	0.048 (0.228)	5	0.129 (0.465)	0.716 (1.17)	15	-0.137 (-0.721)	-0.62 (-0.259)

注: \* 表示  $p < 0.1$ , \*\* 表示  $p < 0.05$ , \*\*\* 表示  $p < 0.01$  括号内为  $T$  检验的统计值。

因此,通过对上述兼收并购类新闻事件有效性的分析,可以得到以下结论:

(1) 兼收并购类新闻对股票价格确实存在显著的影响,证明假设 2 成立; (2) 兼收并购类新闻对股票价格的影响主要集中在新闻发布的前 6 日至后 4 日,证明假设 2.1 成立; (3) 兼收并购类新闻对股票价格的影响是积极的,可以产生较高的平均异常收益率和累计平均异常收益率,除有消息表明并购重组失败,一般可以看作是利好的信号; (4) 若投资者能在兼收并购类新闻发布前提前获知信息,并最晚在发布前 5 日买入股票,在新闻发布日当天卖出便可获得全程的超额收益。但若投资者在新闻发布后还追进去,则是不明智的选择。

(三) 再融资类新闻

本小节将对再融资类新闻对股票价格的影响进行研究,共包含 255 个样本,基于表 1 的统计数据,对再融

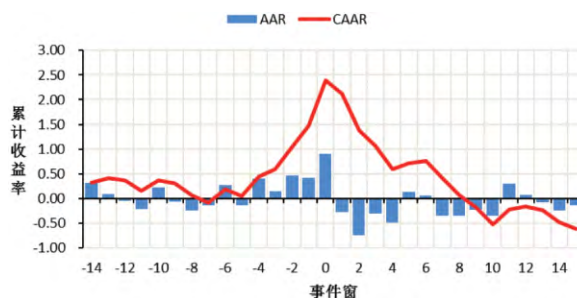


图 3 兼收并购类新闻事件的异常收益率走势图

表 4 再融资类新闻事件的异常收益率检验统计表

事件窗	AAR	CAAR	事件窗	AAR	CAAR	事件窗	AAR	CAAR
-14	-0.14 (-0.852)	-0.14 (-0.852)	-4	-0.091 (-0.42)	1.470** (2.185)	6	0.136 (0.907)	2.182*** (2.732)
-13	0.495** (2.536)	0.354 (1.057)	-3	-0.101 (-0.372)	1.369** (2.047)	7	-0.07 (-0.435)	2.113*** (2.651)
-12	-0.249 (-1.447)	0.105 (0.323)	-2	-0.268 (-1.109)	1.100* (1.79)	8	-0.07 (-0.435)	2.043** (2.461)
-11	0.001 (0.003)	0.106 (0.262)	-1	0.545*** (2.184)	1.645** (2.272)	9	-0.014 (-0.07)	2.029** (2.355)
-10	0.32 (1.297)	0.426 (0.736)	0	1.205*** (4.926)	2.849*** (4.485)	10	0.063 (0.367)	2.092** (2.365)
-9	0.389* (1.963)	0.815 (1.403)	1	-0.313 (-1.397)	2.536*** (3.785)	11	-0.223 (-1.249)	1.869** (2.175)
-8	0.359* (1.714)	1.174* (1.993)	2	-0.292 (-1.133)	2.244*** (3.241)	12	-0.314* (-1.703)	1.555* (1.91)
-7	0.376* (1.902)	1.550*** (2.645)	3	0.018 (0.079)	2.262*** (3.085)	13	0.067 (0.386)	1.622* (1.954)
-6	0.077 (0.31)	1.628** (2.493)	4	-0.151 (-0.569)	2.111*** (2.751)	14	-0.208 (-1.383)	1.414 (1.652)
-5	-0.066 (-0.312)	1.561*** (2.379)	5	-0.065 (-0.304)	2.046** (2.592)	15	-0.056 (-0.34)	1.359 (1.581)

注: \* 表示  $p < 0.1$ , \*\* 表示  $p < 0.05$ , \*\*\* 表示  $p < 0.01$  括号内为  $T$  检验的统计值。



资类新闻事件进行异常收益率的计算及其显著性检验,得到表4所示的统计结果。可以发现,平均异常收益率AAR在事件发生前10日开始便整体维持着较为显著的正收益,而从事件发生的第二天开始便呈现反向走势。

同时,可以绘制再融资类新闻事件的异常收益率走势图,如图4所示。可以发现,平均累计异常收益率曲线呈现阶梯式上升然后逐步下降的趋势,显著的平均异常收益率主要集中在[-13],[-9,-7]和[-1,0]之间,其中,[-1,0]是正的异常收益率的主要贡献者。

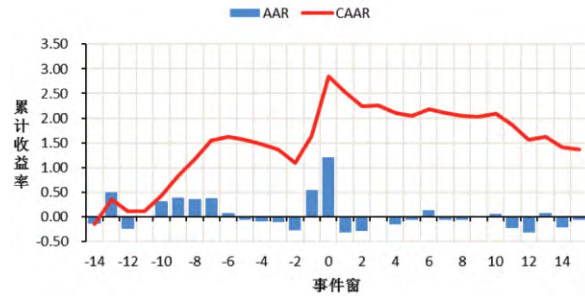


图4 再融资类新闻事件的异常收益率走势图

因此,通过对上述再融资类新闻事件有效性的分析,可以得到以下结论:

(1) 再融资类新闻对股票价格存在显著的影响,证明假设3成立;(2) 再融资类新闻对股票价格的影响范围很广,事件窗的各阶段均存在显著影响,且影响持续性较高。显然,再融资类新闻要比其他类型的新闻更早被投资者获取,通常正式新闻公布前十多天股价就开始发生异动。(3) 再融资类新闻对股票价格的影响是积极的,使得股票的收益率明显增加,可以产生较高的平均异常收益率和平均累计异常收益率,这说明大多数投资者相信再融资可以为企业创造价值,并获得超额收益,证明假设3.1不成立。当然,也可能是由于本文选择的财经新闻的时间为2014年,此时股市正处于回暖的阶段,因此可以看作是利好的信号。若选用熊市时期的新闻进行研究,再融资类新闻对股价的影响可能会有差异。(4) 若投资者能在再融资类新闻发布前提前获知信息,并最晚在发布前10日买入股票,在新闻发布日当天卖出便可获得全程的超额收益。

(四) 盈利能力类新闻

表5 盈利能力类新闻事件的异常收益率检验统计表

事件窗	AAR	CAAR	事件窗	AAR	CAAR	事件窗	AAR	CAAR
-14	-0.024 (-0.225)	-0.024 (-0.225)	-4	-0.016 (0.679)	0.095 (1.35)	6	-0.048 (-0.411)	1.096* (0.594)
-13	0 (-0.004)	-0.025 (-0.183)	-3	-0.053 (-0.39)	0.042 (1.243)	7	0.066 (0.477)	1.162 (0.685)
-12	0.036 (0.274)	0.012 (-0.01)	-2	0.024 (1.618)	0.066 (1.538)	8	0.066 (0.477)	1.228 (0.74)
-11	-0.046 (-0.339)	-0.034 (-0.16)	-1	0.007 (0.052)	0.073 (1.474)	9	-0.073 (-0.479)	1.156 (0.634)
-10	-0.001 (-0.009)	-0.035 (-0.147)	0	0.102 (1.558)	0.175* (2.009)	10	-0.176 (-1.31)	0.98 (0.446)
-9	-0.144 (-1.075)	-0.18 (-0.552)	1	0.124 (0.158)	0.299* (1.985)	11	0.095 (0.726)	1.075 (0.531)
-8	0.012 (0.098)	-0.167 (-0.448)	2	0.272* (-1.894)	0.571** (1.567)	12	-0.062 (-0.423)	1.013 (0.462)
-7	0.109 (0.912)	-0.059 (-0.014)	3	0.312* (-1.852)	0.882* (1.102)	13	-0.148 (-1.147)	0.864 (0.251)
-6	0.101 (2.517)	0.043 (0.824)	4	-0.079 (0.472)	0.804 (1.142)	14	0.02 (0.149)	0.884 (0.269)
-5	0.069 (1.603)	0.111 (1.242)	5	0.341** (-2.28)	1.144 (0.698)	15	0.026 (0.186)	0.91 (0.302)

注:\*表示 $p < 0.1$ ,\*\*表示 $p < 0.05$ ,\*\*\*表示 $p < 0.01$  括号内为T检验的统计值。

同时,可以绘制盈利能力类新闻事件的异常收益率走势图,如图5所示。可以发现,平均累计异常收益率曲线在接近事件公告日时快速上升,之后平缓下降,平均异常收益率主要在[2,5]间较显著。

因此,通过对上述盈利能力类新闻事件有效性的分析,可以得到以下结论:

(1) 盈利能力类新闻对股票价格的影响确实存在,证明假设4成立;(2) 分析得出盈利能力类新

闻对股票价格的影响是积极的,因为盈利能力展现了上市公司良好的经营能力,易受到投资者的关注;(3) 盈利能力类新闻对股票价格的影响力是持续的,能够使高异常收益率维持较长时间而不会迅速回归到正常收益率的水平;(4) 盈利能力类新闻在发布前不易获知,所以投资者可以在新闻发布当天买入股票,并持有 8 个交易日左右卖出即可获得全程的超额收益率。



图 5 盈利能力类新闻事件的异常收益率走势图

(五) 违规处罚类新闻

本小节将对违规处罚类新闻对股票价格的影响进行研究,共包含 226 个样本,基于表 1 的统计数据,对违规处罚类新闻事件进行异常收益率的计算及其显著性检验,得到表 6 所示的统计结果。可以发现,在事件发生前后存在显著为负的异常收益率,并持续影响至时间窗的结束。

表 6 违规处罚类新闻事件的异常收益率检验统计表

同时,可以绘制违规处罚类新闻事件的异常收益率走势图,如图 6 所示。可以发现,平均累计异常收益率曲线在接近公告日前 2 日开始迅速下降,并维持了较长一段时间,其中平均异常收益主要在  $[-1, 1]$  间较为显著。

事件窗	AAR	CAAR	事件窗	AAR	CAAR	事件窗	AAR	CAAR
-14	0.087 (0.438)	0.087 (0.438)	-4	0.026 (0.128)	0.413 (0.817)	6	0.167 (0.923)	-0.832 (-0.462)
-13	-0.017 (-0.086)	0.069 (0.2)	-3	-0.011 (-0.05)	0.401 (0.8)	7	-0.055 (-0.296)	-0.888 (-0.513)
-12	-0.193 (-0.803)	-0.124 (-0.193)	-2	0.236 (1.039)	0.637 (1.016)	8	-0.055 (-0.296)	-0.943 (-0.549)
-11	-0.139 (-0.626)	-0.263 (-0.469)	-1	-0.401** (-2.006)	0.236 (0.385)	9	-0.107 (-0.552)	-1.050 (-0.637)
-10	-0.334 (-1.229)	-0.597 (-0.973)	0	-0.646*** (-3.003)	-0.410 (-0.491)	10	-0.17 (-0.857)	-1.220 (-0.75)
-9	-0.088 (-0.384)	-0.686 (-0.981)	1	-0.247* (-1.42)	-0.657 (-0.707)	11	0.128 (0.755)	-1.093 (-0.661)
-8	0.382 (1.985)	-0.303 (-0.113)	2	-0.176 (0.936)	-0.832 (-0.541)	12	-0.1 (-0.434)	-1.193 (-0.728)
-7	0.212 (1.214)	-0.091 (0.276)	3	-0.228 (-1.046)	-1.060 (-0.725)	13	0.009 (0.055)	-1.183 (-0.703)
-6	0.161 (0.955)	0.070 (0.483)	4	0.116 (0.562)	-0.944 (-0.625)	14	-0.528*** (-3.099)	-1.711 (-1.238)
-5	0.317 (1.327)	0.387 (0.814)	5	-0.056 (-0.305)	-1.000 (-0.656)	15	-0.165 (-0.99)	-1.876 (-1.383)

注: \* 表示  $p < 0.1$ , \*\* 表示  $p < 0.05$ , \*\*\* 表示  $p < 0.01$  括号内为  $T$  检验的统计值。

因此,通过对上述违规处罚类新闻事件有效性的分析,可以得到以下结论:

(1) 违规处罚类新

闻对股票价格的影响确实存在,证明假设 5 成立;(2) 违规处罚类新闻是五类样本中唯一的负面新闻类型,其对公司股价会产生明显的负异常收益率,证明假设 5.1 成立;(3) 违规处罚类新闻对股价的影响在新闻公布当天的负向作用最为显著,并使得影响持续很长一段时间;(4) 违规处罚类新闻从新闻公布日前 1 日开始对股价产生明显的影响,说明部分投资者已初步知晓上市公司的违规行为,并直观体现在了股价变化上。由于企业的违规违法行为,使得投资者在较长一段时间内不信任该公司,否定企业价值,因此异常收益率一路走低,并且会随着时间的推移开始第二波股价的下跌。

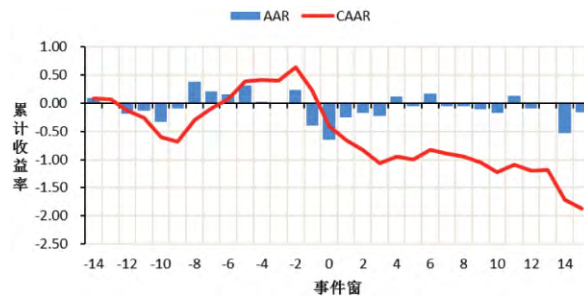


图 6 违规处罚类新闻事件的异常收益率走势图

## 六、结论与建议

本文通过对已有文献的梳理,结合金融理论和文本挖掘技术,基于事件研究法构建了异质性财经新闻有效性模型,开展各类互联网财经新闻对股价的影响研究。研究发现,政策扶持类、兼收并购类、再融资类和盈利能力类新闻对股价具有积极的推动作用,其中,政策扶持类在新闻发布前后都有影响,加持都能带来超额收益;兼收并购类在新闻发布前加持收益最大,但是发布后一两天内加持也能有部分超额收益;再融资类新闻发布前十多天具有较大影响,加持有超额收益,但是后期影响很小;盈利能力类新闻则在前期无影响,后期能保持一段时间的影响,建议发布后加持获得收益。而违规处罚类新闻具有负方向的作用,而且影响会持续很长一段时间。总之,随着互联网媒体的不断发展和投资者对上市公司的持续关注,异质性财经新闻对股价的影响已有规律可寻。

根据以上结论,本文有如下建议:首先,从政府监管机构的层面来看,一方面,应该加强监管以杜绝内幕交易行为,制定严格的信息披露机制,尤其是在上市公司处于政策扶持、兼收并购和再融资等重大决策情况时;另一方面,金融市场监管机构应与新闻媒体监管机构形成长效合作机制,不仅要确保财经类新闻报道客观、公正、全面,而且要充分发挥媒体报道对整个金融市场的监督作用。其次,本文的研究也对财经类新闻媒体及其工作人员的专业性和职业素养提出了更高的要求,新闻媒体及其工作人员不仅要具有较高的财经领域专业理解能力,而且应该客观、公正、全面地对财经类信息进行阐述和报道,杜绝偏私和造传谣言。再次,上市公司应自觉遵守法律法规,依法及时披露公司信息,严杜信息泄露和欺骗。最后,作为投资者,应该理性看待股票市场和财经信息,掌握客观判断和甄别的能力,执行科学的投资决策,切忌盲目跟风炒作热点。

当然,本文的研究还存在一定的局限性。投资者如今不止获取门户网站的财经新闻专栏消息,随着新兴媒体的崛起,微博、论坛等新媒体平台上的信息也受到广泛关注。如果想要获得更精准的分析,应该设法获取涵盖所有对市场有影响的消息源,并研究它们之间的交互效应,这样才会使股市异常波动的分析更加精准,更加具有实用价值。

## 注释:

①此处使用的是 LIBSVM 软件包。LIBSVM 是台湾大学林智仁博士等开发设计的一个操作简单、易于使用、快速有效的通用 SVM 软件包,可以解决分类问题、回归问题以及分布估计等问题,提供了线性、多项式、径向基和 S 形函数四种常用的核函数供选择,可以有效地解决多类问题、交叉验证选择参数、对不平衡样本加权、多类问题的概率估计等。

## 参考文献:

- [1] 刘海二. 互联网金融的风险与监管 [J]. 经济与管理研究, 2015(10): 54-61.
- [2] FAMA E F. Efficient capital markets: a review of theory and empirical work [J]. Journal of finance, 1970, 25(2): 383-417.
- [3] BEAUDRY P, PORTIER F. Stock prices, news, and economic fluctuations [J]. American economic review, 2006, 96(4): 1293-1307.
- [4] ANTWEILER W, FRANK M Z. Do US stock markets typically overreact to corporate news stories? [Z/OL]. SSRN, 2006, [2016-11-05]. <https://ssrn.com/abstract=878091>.
- [5] FUNG G P C, YU J X, LAM W. News sensitive stock trend prediction [C]//CHEN M S, YU P S, LIU B. Advances in knowledge discovery and data mining: 6th Pacific-Asia conference. Berlin: Springer, 2002: 481-493.
- [6] CUTLER D M, POTERBA J M, SUMMERS L H. What moves stock prices? [J]. Journal of portfolio management, 1989, 15(3): 4-12.
- [7] KLIBANOFF P, LAMONT O, WIZMANT T A. Investor reaction to salient news in closed-end country funds [J]. Journal of finance, 1998, 53(2): 673-699.

- [8] CHAN W S. Stock price reaction to news and no-news: drift and reversal after headlines [J]. *Journal of financial economics*, 2003, 70(2): 223-260.
- [9] MITCHELL M L, MULHERIN J H. The impact of public information on the stock market [J]. *Journal of finance*, 1994, 49(3): 923-950.
- [10] 杨继东. 媒体影响了投资者行为吗? ——基于文献的一个思考[J]. *金融研究* 2007(11): 93-102.
- [11] LI Q, WANG T J, GONG Q X, et al. Media-aware quantitative trading based on public web information [J]. *Decision support systems*, 2014, 61: 93-105.
- [12] 朱波 孙鹏阁 龙云庚. 高质量的会计信息促进了投资者理性投资吗 [J]. *经济与管理研究* 2015(8): 130-138.
- [13] WYSOCKI P D. Cheap talk on the web: the determinants of postings on stock message boards [Z]. *University of Michigan working paper*, 1998, No. 98025.
- [14] HUBERMAN G. Contagious speculation and a cure for cancer: a nonevent that made stock prices soar [J]. *Journal of finance*, 2001, 56(1): 387-396.
- [15] TAKEDA F, YAMAZAKI H. Stock price reactions to public TV programs on listed Japanese companies [J]. *Economics bulletin*, 2006, 13(11): 1-7.
- [16] ZIMBRA D, FU T, LI X. Assessing public opinions through Web 2.0: a case study on Wal-Mart [Z]. *The 13th international conference on information systems meeting paper*, 2009.
- [17] 赵伟 梁循. 互联网金融信息量与收益率波动关联研究[J]. *计算机技术与发展* 2009(12): 1-4.
- [18] 饶育蕾 王攀. 媒体关注度对新股表现的影响——来自中国股票市场的证据[J]. *财务与金融* 2010(3): 1-7.
- [19] 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(2): 287-297.
- [20] SPRENGER T O, TUMASJAN A, SANDNER P G, et al. Tweets and trades: the information content of stock microblogs [J]. *European financial management*, 2014, 20(5): 926-957.
- [21] 张正勇. 媒体关注、产权性质与社会责任信息披露——来自社会责任报告自愿披露的经验证据[J]. *南京财经大学学报* 2015(3): 70-77.
- [22] NIEDERHOFFER V. The analysis of world events and stock prices [J]. *Journal of finance*, 1971, 44(2): 193-219.
- [23] KEOWN A J, PINKERTON J M. Merger announcements and insider trading activity: an empirical investigation [J]. *Journal of finance*, 1981, 36(4): 855-869.
- [24] LI Q, WANG T J, LI P, et al. The effect of news and public mood on stock movements [J]. *Information sciences*, 2014, 278(10): 826-840.
- [25] WÜTHRICH B, PERMUNETILLEKE D, LEUNG S. Daily prediction of major stock indices from textual WWW Data [J]. *HKIE Transactions*, 1998, 5(3): 151-156.
- [26] TAN A H, RIDGE K, LABS D, et al. Text mining: the state of the art and the challenges [C]. *Proceedings of the PAKDD workshop on knowledge discovery from advanced databases*, 2000: 65-70.
- [27] LAVRENKO V, SCHMILL M, LAWRIE D, et al. Language models for financial news recommendation [C]. *Proceedings of the ninth international conference on information and knowledge management (CIKM00)*, 2000: 389-396.
- [28] SCHUMAKER R, CHEN H. Textual analysis of stock market prediction using breaking financial news: the AZFin text system [J]. *ACM transactions on information systems*, 2009, 27(2): 1139-1141.
- [29] LÆGREID T, SANDAL P C. Financial news mining: extracting useful information from continuous stream of text [J]. *Institutt for datateknikk og informasjonsvitenskap*, 2006, 13: 222-228.
- [30] TETLOCK P C. Giving content to investor sentiment: the role of media in the stock market [J]. *Journal of finance*, 2007, 62(3): 1139-1168.
- [31] TETLOCK P C, SAAR-TSECHANSKY M, MACSKASSY S. More than words: quantifying language to measure firms' fundamentals [J]. *Journal of finance*, 2008, 63(3): 1437-1467.
- [32] 陈华 梁循. 互联网股票新闻归类和板块分析的方法[J]. *电脑开发与应用* 2006(11): 2-4.
- [33] SMAILOVIĆ J, GRČAR M, LAVRAČ N, et al. Stream-based active learning for sentiment analysis in the financial do-

- main [J]. *Information sciences*, 2014, 285: 181-203.
- [34] BONAPARTE Y, KUMAR A. Political activism, information costs, and stock market participation [J]. *Journal of financial economics*, 2013, 107(3): 760-786.
- [35] APERGIS N. Policy risks, technological risks and stock returns: new evidence from the US stock market [J]. *Economic modelling*, 2015, 51(C): 359-365.
- [36] ASQUITH P, MULLINS D W. Equity issues and offering dilution [J]. *Journal of financial economics*, 1986, 15(1-2): 61-89.
- [37] MASULIS R W, KORWAR A N. Seasoned equity offerings: an empirical investigation [J]. *Journal of financial economics*, 1986, 15(1-2): 91-118.
- [38] MANN S V, SICHERMAN N W. The agency costs of free cash flow: acquisition activity and equity issues [J]. *Journal of business*, 1991, 64(2): 213-227.
- [39] FAMA E F, FISHER L, JENSEN M C, et al. The adjustment of stock prices to new information [J]. *International economic review*, 1969, 10(1): 1-21.
- [40] 张兵, 李晓明. 中国股票市场的渐进有效性研究 [J]. *经济研究*, 2003(1): 54-61.
- [41] 吴振翔, 陈敏. 中国股票市场弱有效性的统计套利检验 [J]. *系统工程理论与实践*, 2007(2): 92-98.
- [42] BOEHMER E, MASUMECI J, POULSEN A B. Event-study methodology under conditions of event-induced variance [J]. *Journal of financial economics*, 1991, 30(2): 253-272.
- [43] 袁显平, 柯大钢. 事件研究方法及其在金融经济研究中的应用 [J]. *统计研究*, 2006(10): 31-35.

(责任编辑: 木子)

## The impact of internet heterogeneous financial and economic news on stock market: evidence from Chinese internet data and listed companies

LIU Haifei, XU Jintao

(School of Management and Engineering, Nanjing University, Nanjing 210093, China)

**Abstract:** This paper analyzes the impact of internet heterogeneous financial and economic news on short-term abnormal yield of Chinese stock by text mining technology and event research methodology. It is found that five kinds of internet heterogeneity news can cause short-term abnormal yield of stock. Among them, the policy support, mergers and acquisitions, refinancing and profitability news can produce positive abnormal yield to the company's stock, and the violation punishment news can produce negative abnormal yield. The conclusion of this paper will provide some important inspiration and evidence for regulators, financial news media staffs, listed companies and investors in the aspects of analysis, communication, transaction and regulation of information.

**Key words:** internet news; heterogeneity; text mining; event research; abnormal yield; media attention