

百度关注度指数与股票价格关系研究^①

张悟移¹, 李 杰²

昆明理工大学 管理与经济学院, 昆明 650093

摘要: 针对投资者有限关注与股票市场的联动作用进行了分析, 回顾了关于传统金融学理论和现代金融学理论的国内外文献, 并且比较了两者的区别. 介绍了有限关注以及目前的几种理论模型, 根据模型得出个体注意力是有限的, 投资者风险认知不足会直接影响关注行为. 基于传统的投资者关注度代理指标包括换手率、交易量等, 利用互联网提供的代理变量可以更直接地反映投资者的关注行为, 因此研究选取百度指数作为投资者关注行为的代理变量, 当投资者在搜索该股票的信息时, 无疑表示投资者对这支股票产生了关注. 百度指数提供的关键词被检索的次数, 更加直接地反映了投资者对某只股票的关注度, 百度指数提供的数据也会驱动投资者的交易行为, 从而引起股票价格的波动. 通过建立 VAR 模型、格兰杰因果检验和脉冲响应函数分析对投资者关注行为和股票收盘价之间的关系进行实证分析, 结果显示, 投资者的关注行为与股票收盘价之间存在一定的影响.

关键词: 行为金融学; 有限关注; 百度指数; 股票市场

中图分类号: F503

文献标志码: A

文章编号: 1000-5471(2019)02-0075-09

伴随着我国经济的发展, 金融领域出现的有些现象无法用以前的理论来进行解释, 如处置效应, 即投资者倾向于持有亏损的股票而过早得卖掉盈利的股票; 公告效应, 国家的新政策与公司的相关规定会引起股价的起伏^[1]. 1965 年美国著名经济学家 Eugene Fama 首次提出了“有效市场理论”^[2]: 在一个有效的资本市场中, 资产价格能够完全地反映所有可供获得的信息, 市场中大部分的投资者都想追求利益最大化且都是理性的. 根据这一假设, 市场参与者具有完美理性和无限信息处理能力, 在买卖股票时会迅速有效地利用可能的信息, 并且能够及时充分地将所有公开信息反应在股价中, 即投资者是不会出现系统的非理性行为. 股票价格的变化与该股票的信息是息息相关的, 信息的不可测性决定了资产的未来价格会像化学分子运动中的“布朗运动”一样随机游走.

如今, 互联网时代的来临改变了整个金融市场的格局. 面对过量的信息, 投资主体特别是缺乏经验的经济个体在决策时容易受到这些信息的影响, 使自身的情绪产生波动, 如紧张、反应过度或反应不足等, 从而无法集中注意力, 导致投资决策产生偏差. 放眼中国证券市场, 个体投资者的羊群效应以及非理性特征非常突出, 投资者对市场的认识可能会发生系统性的启发性认知偏差, 股价也发生非理性的偏离^[3]. 投资者的非理性行为渐渐成为了中国金融市场异象层出不穷的原因之一.

1 国内外文献综述

关于关注行为影响股票市场的相关研究, 国内外都有不少的实证研究. 近年来, 许多学者先后选取不

① 收稿日期: 2018-03-08

基金项目: 国家自然科学基金项目(71562023); 云南省省院省校教育合作人文社科项目(SYSX201504).

作者简介: 张悟移(1965-), 男, 教授, 博士生导师, 管理学博士, 主要从事供应链金融与管理的研究.

通信作者: 李 杰, 硕士研究生.

同的关注行为代理指标,如:利用涨停效应作为投资关注的代理变量。Seasholes 和 Wu 认为涨停报道会影响投资者的注意力配置^[4]。他们为了吸引投资者的关注,仅仅将股票涨停现象作为研究对象,研究结果显示,涨停现象先是导致投资者对该股票的超买行为,随后引发股价持续上涨,但随后反转;财经新闻也纳入了投资关注的代理变量。以 Tetlock 为代表的研究发现,新闻的文本内容所包含的与股票价格有关的“软信息”对投资者的情绪有引导作用^[5]。新闻事件背后所指向的政治、经济、社会等各个方面的变动能够影响投资者的情绪,导致股票价格发生变动。诸如此类的间接指标还有广告费用、股票成交量、换手率等等。随着互联网的蓬勃发展,人们对于信息的及时性和有效性的要求越来越高,主动获取信息的欲望也越来越强,互联网已经成为了投资者了解市场的重要平台。在国外,谷歌(Google)甚至为了专门挖掘互联网数据的价值设立了首席经济学家的职位。

在国内,互联网搜索数据的价值逐渐开始被许多学者发现和研究。宋双杰等利用谷歌趋势的搜索量数据,从行为金融学的角度出发,构建异常搜索量指标,研究了中国股市的 IPO 异象^[6]。姜翔程等利用 GARCH 族模型对我国股市的波动性进行了研究^[7]。结果表明,IPO 前个股的网络搜索量对于该股票市场热销程度、首日超额收益和长期低迷表现这三大 IPO 市场异象有良好的、统一的解释^[8]。过去的搜索引擎只是起着查询信息的作用,而现在它还会记录人们的主观搜索行为,百度搜索就提供了一个将投资者有限关注的心理活动转化为一种可视化、可量化的平台^[9]。

2 关注行为对股票价格的研究

采用格兰杰因果检验和 VAR 方法,研究关注行为对股票收盘价的影响,进一步考察关注行为对股市的影响。

2.1 研究假设

每一个个体投资者所做的决策都是经过一定的思考而形成的,很大部分是由他们的内心支配着行为。事实上,人们的注意力是有限的,投资者的决策行为会受到新闻媒体、政策变动、证券分析师推荐等的影响。而关注行为之所以会受到影响,是因为投资者文化水平、情绪、心理承受能力、闲暇时间的多少、信息掌握程度、风险认知水平存在差异,从而导致关注行为个性化、差异化,使同一市场下存在层次不齐的投资风险。目前,这种行为偏差逐渐成为了一种不确定性因素,对整个证券市场的股票收盘价产生影响。

基于此,我们提出假设:股票收盘价会受到关注行为的影响。

2.2 样本的选取和变量定义

本文百度指数的数据主要来自股票公司的全称的百度关注度指数,采用创业板和主板市场多只股票指标,涉及医药板块、机械板块、地产板块等多领域。主要选取近几年(2011 年 01 月 01 日至 2015 年 12 月 31 日)的数据,股票名称搜索量来自百度指数,用以描述关注度指数,股票收盘价数据来自网易财经。

我国主板市场成立较早,主要是在发展蓝筹股。我国创业板市场实行“风险自负”的投资原则,即监管当局不判断公司的前景、质量等情况,仅是考虑其是否满足上市条件。本文选取创业板市场与主板市场的股票进行研究更具有意义。

在本部分实证的分析中,本文采用了一些股票市场中比较常用的概念。为了能够清晰地了解各个概念的含义,现将其定义描述如下:

1) 收盘价:指当日股票市场交易结束时的成交价格。

2) 百度指数:以百度海量网民行为数据为基础,能够告诉用户某个关键词在百度的搜索规模。

2.3 股票价格与百度关注度指数

选取从 2011 年 1 月 4 日至 2015 年 12 月 31 日的相关数据,将百度关注度指数与股票收盘价画出变动走势图,以其中一支股票为例,结果如图 1,可发现无论是收盘价还是百度指数,两个图形中存在相似的整体变化趋势。2014 年到 2015 年的波动较大,其中最高值均出现在 2015 年初,在标记区间 A 段内,关注度

与股票收盘价有明显相似的变化趋势. 而对比来看, 2011 年到 2013 年波动趋于缓和.

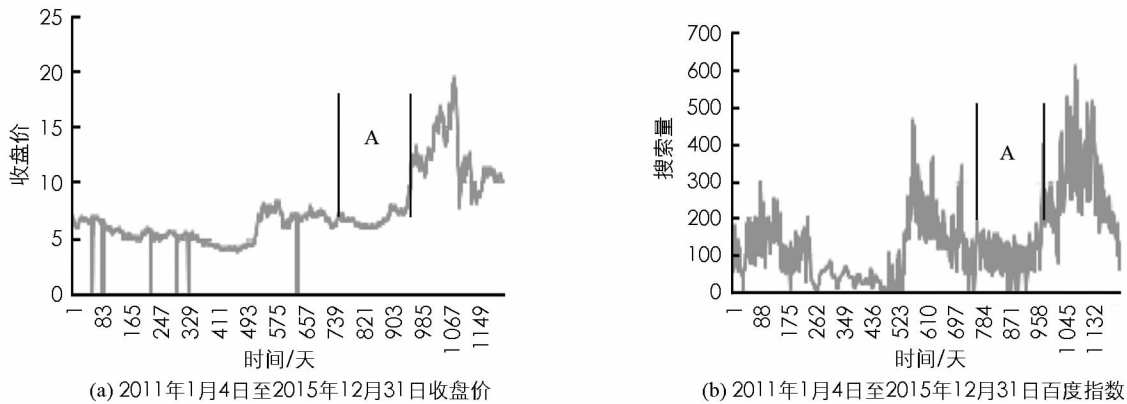


图 1 某只股票的收盘价与百度搜索指数

2.4 关注行为对股票价格的影响——格兰杰因果检验

随着对行为金融学的逐渐了解, 我们不得不去深思, 投资者的行为因素是否会影响股票收盘价, 会产生怎样的影响. 为了更好地判断时间序列之间的相关性, 我们需要运用格兰杰因果检验来确定序列之间的引导关系, 以此结果判断一个变量的变化是否会引起另一个变量的变化, 从而检验投资者关注行为与股票收盘价之间的引导关系.

格兰杰因果关系检验模型如下所示:

$$y_t = \sum_{i=1}^q \alpha_i X_{t-i} + \sum_{j=1}^q \beta_j y_{t-j} + u_{1t} \tag{1}$$

$$X_t = \sum_{i=1}^s \lambda_i X_{t-i} + \sum_{j=1}^s \delta_j y_{t-j} + u_{2t} \tag{2}$$

其中白噪音 u_{1t} 和 u_{2t} 假设是不相关的.
式(1) 假设当前的 y 变量与 y 自身以及 x 的过去值有关, (2) 式对 x 假设了相似的行为.
对式(1) 而言, 其零假设 $H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_q = 0$
对式(2) 而言, 其零假设 $H_0 : \delta_1 = \delta_2 = \cdots = \delta_s = 0$
具体的检验步骤如下:

首先, 将当前的 y 对全部的滞后项 y (即 y 对 y 的滞后项 $y_{t-1}, y_{t-2}, \cdots, y_{t-q}$) 做回归, 其中不包括滞后项 x , 然后从这个受约束回归中得到残差平方和 RSS_R .
然后, 做一个含有滞后项 x 的无约束的回归, 在第一步的回归式中加入滞后项 x , 然后从这个回归中获得无约束的残差平方和 RSS_U .

零假设是 $H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_q = 0$, 即滞后项 x 不属于此回归.
接着, 用 F 检验此假设, 即:

$$F = \frac{\frac{RSS_R - RSS_U}{q}}{\frac{RSS_U}{n - k}} \tag{3}$$

遵循自由度为 q 和 $(n - k)$ 的 F 分布. 其中样本容量为 n 、滞后项 x 的数量是 q , 即有约束回归方程中待估参数的个数, 无约束回归中待估参数的个数为 k .
如果计算的 F 值大于给定显著性水平 α 下 F 分布的相应的临界值, 则拒绝原假设, 认为 x 是 y 的格兰杰原因.

最后, 根据 F 值判定是否存在由 x 到 y 的单向因果关系.
分 4 种情形讨论:

- 1) y 随着 x 的变化而变化,即存在由 x 到 y 的单向因果关系.
- 2) x 随着 y 的变化而变化,即存在由 y 到 x 的单向因果关系.
- 3) x 与 y 互相影响,即存在由 x 到 y 的单向因果关系,同时也存在由 y 到 x 的单向因果关系.
- 4) x 与 y 是独立的,或 x 与 y 间不存在因果关系.

我们运用 Matlab 软件建立格兰杰因果关系,研究 x 与 y 的因果关系,统计出所选每只股票收盘价与百度关注度指数之间的关联程度(图 2—图 8),具体分布情况统计如表 1 所示:

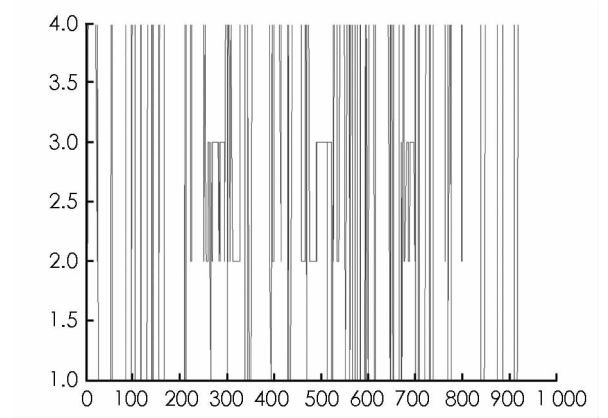


图 2 首创股份(股票代码: 600 008)
格兰杰检验结果分布图

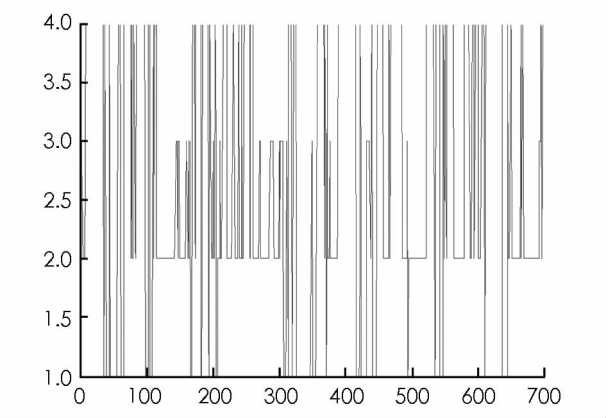


图 3 浙江广厦(股票代码: 600 052)
格兰杰检验结果分布图

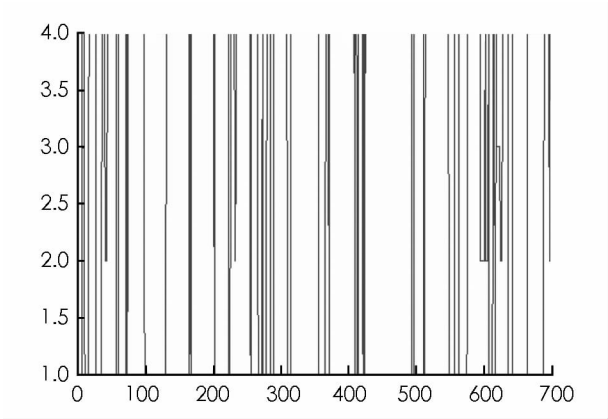


图 4 华业资本(股票代码: 600 240)
格兰杰检验结果分布图

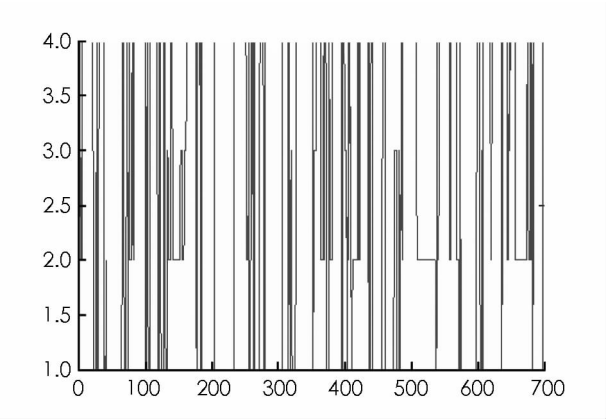


图 5 中集集团(股票代码: 000 039)
格兰杰检验结果分布图

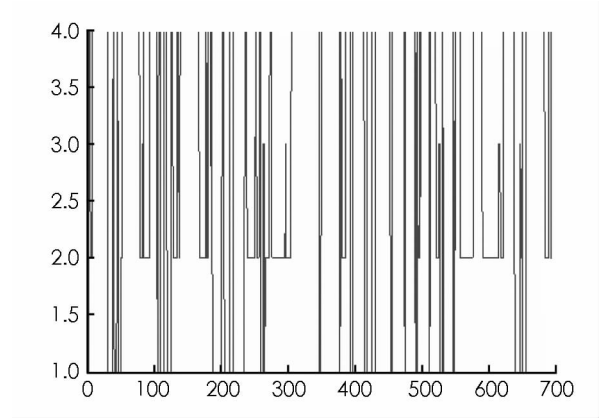


图 6 北大医药(股票代码: 000 788)
格兰杰检验结果分布图

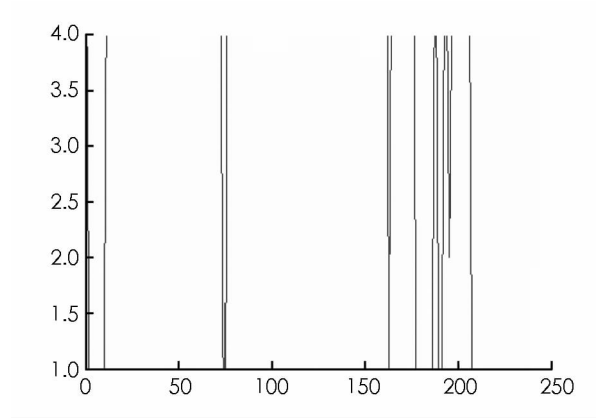


图 7 浙江金科(股票代码: 300 459)
格兰杰检验结果分布图

表 1 格兰杰检验分布情况

股票代码	第 1 类 关系	第 2 类 关系	第 3 类 关系	第 4 类 关系
600 008	256	85	77	23
600 052	116	251	43	88
600 240	167	20	6	5
000 039	212	116	25	45
000 788	133	162	9	94
300 459	27	1	0	80
300 005	118	66	24	87

我们将统计结果分为 4 种因果关系, 1 代表 x 是引起 y 变化的原因, 2 代表 y 是引起 x 变化的原因, 3 代表 x 和 y 互为因果关系, 4 代表 x 与 y 间不存在因果关系。其中, x 代表百度关注度指数, y 代表日股票收盘价(剔除节假日)。

综合以上几只股票的分析结果, 主板市场和创业板市场中, 关注度与股票收盘价的因果关系基本是相同的, 但是在分布情况上有些差异。主板市场股票明显比创业板股票更多地表现出关注度与股票收盘价是存在相互关系的。这主要是因为主板市场主要是在发展蓝筹股, 用于服务各个行业的龙头、对整个经济平稳健康发展发挥决定性作用的企业, 创业板市场专注于服务自主创新能力高、有巨大发展潜能的中小型企业。

从结果分布情况可以看出, 无论是蓝筹股还是创业板股票, 每只股票分布在第一种情况的最多, 即说明 x 构成了 y 变化的原因。在研究 x 与 y 之间存在因果关系时, 比较数据分布情况, 我们得出结论就是, 百度关注度指数的变化构成了影响股票收盘价的原因, 投资者的关注行为因素对股票收盘价的影响具有明显的显著性^[10]。

2.5 关注行为与收盘价的 VAR 方法研究——脉冲响应方法

为了进一步确定股票关注度与股票收盘价之间的相关性, 接下来拟建立自回归(VAR)模型来进行精确的研究。

滞后阶数为 p 的 VAR 模型表达式为:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + Bx_t + \epsilon_t$$

(4)

其中: y_t 为 k 维内生变量向量, x_t 为 d 维外生变量向量, p 为滞后阶数, 样本个数为 T 。 $k \times k$ 维矩阵 $A_1, \cdots, A_p$ 和 $k \times d$ 维矩阵 B 为被估计的系数矩阵^[11]。 ϵ_t 是 k 维扰动向量, 它们相互之间可以同期相关, 但不与自己的滞后值相关及不与等式右边的变量相关^[12]。

由图 9(a)可以看出, 对于关注度指数的冲击, 股票收盘价在初期上升较快, 之后一直平稳上升, 到第 10 期达到最大值, 总体保持在正向水平。关注度对股票收盘价的反应为正向, 初期上升较明显, 后面趋于平稳。如图 9(b), 因此, 关注度指数与股票收盘价之间互相都表现出明显的正向影响, 而且这种影响是长期的。

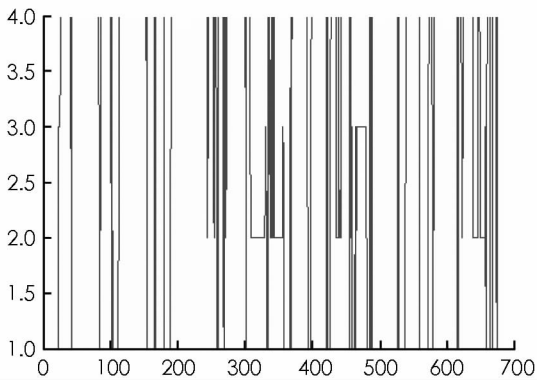


图 8 探路者(股票代码: 300 005)
格兰杰检验结果分布图

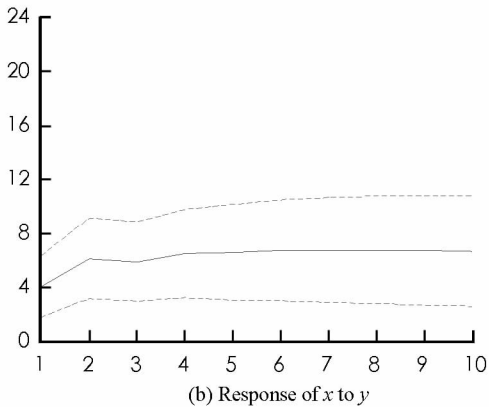
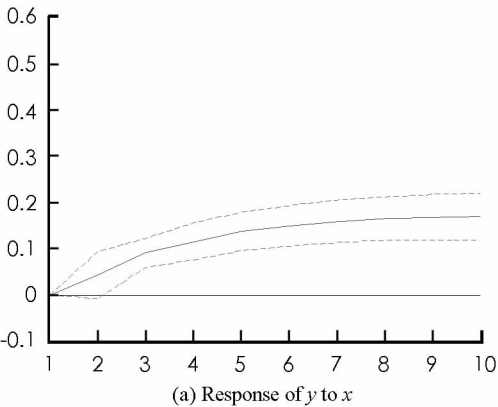


图 9 股票代码: 600 008 脉冲响应函数结果

由图 10(a)可以看出,对于关注度指数的冲击,股票收盘价在前面几期上升迅速,到第 5 期达到最大值,之后有所下降,但总体均表现为正向响应.关注度对股票收盘价的响应初期为正向且接近零,先增加到第 2 期,然后开始衰减,总体的响应为正向,如图 10(b).因此,关注度指数与股票收盘价之间互相都表现出正向影响,而且这种影响是长期的.

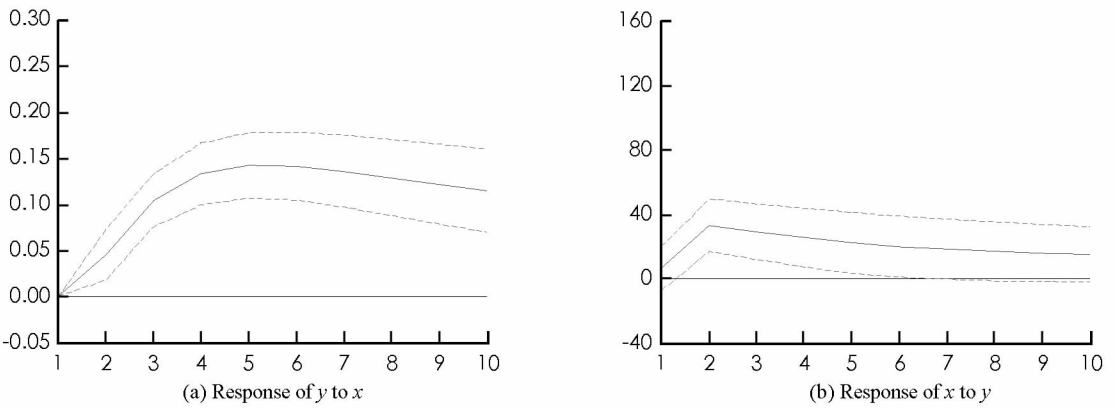


图 10 股票代码: 600 052 脉冲响应函数结果

由图 11(a)可以看出,对于关注度指数的冲击,股票收盘价在初期基本无响应,从零开始衰减到第 10 期,整体表现为一定的负向响应.而关注度对股票收盘价的响应先正而后负向,总体上表现为负向响应,正向反应只表现了两期,从第 2 期就开始出现衰减,第 4 期接近于零,一直到第 7 期达到最低点后就稳定在负向水平,如图 11(b).虽然该股票收盘价对关注度表现出负向的响应,同时,关注度指数与股票收盘价之间互相都表现出负向影响,而且这种影响是长期的,但是两者总体的响应程度较小,说明负面影响程度不高.

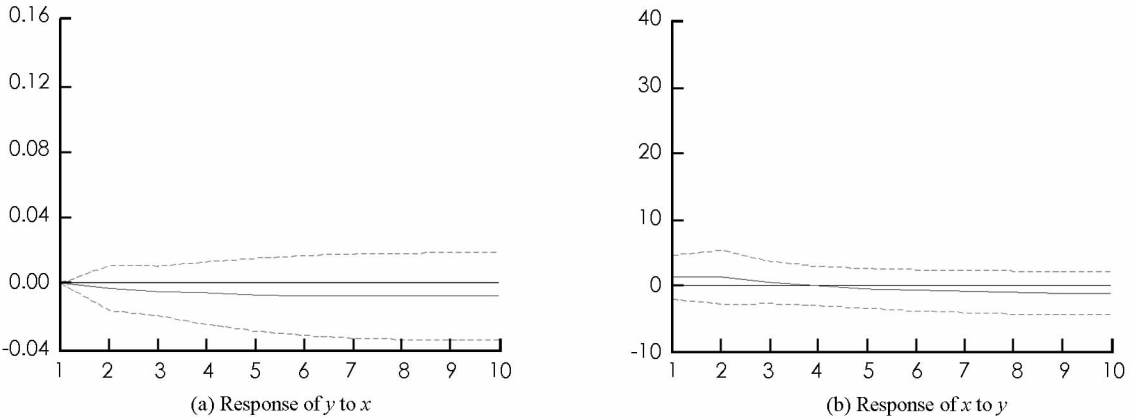


图 11 股票代码: 600 024 脉冲响应函数结果

从图 12(a)中可以看出,股票收盘价对关注度的反应整体上接近于零.从初期到第 3 期表现出微弱的正向响应,第 4 期开始基本为零,然后在第 6 期开始衰减,表现为负向反应,到第 10 期达到最低值.而关注度指数对股票收盘价的反应是正向应,第 2 期达到最高值后开始衰减,到第 10 期达到最低值,且与初期值相近,如图 12(b).由此说明,关注度指数对股票收盘价的影响接近零,股票收盘价对关注度指数具有正向影响,两者之间的这种相关性是长期的.

从图 13(a)中可以看出,股票收盘价对关注度的反应整体上都保持在正向水平,但响应程度较低.初期基本无响应,但从第 2 期开始表现出一定的正向响应,第 3 期达到最高值,之后虽有所下降,但变化很小.而关注度指数对股票收盘价的反应是正向应,上升到第 3 期时达到最高值,然后一直处于衰减状态,到第 10 期达到正向的最低值,如图 13(b).由此说明,关注度与股票收盘价之间存在正向影响,这种影响是长期的.

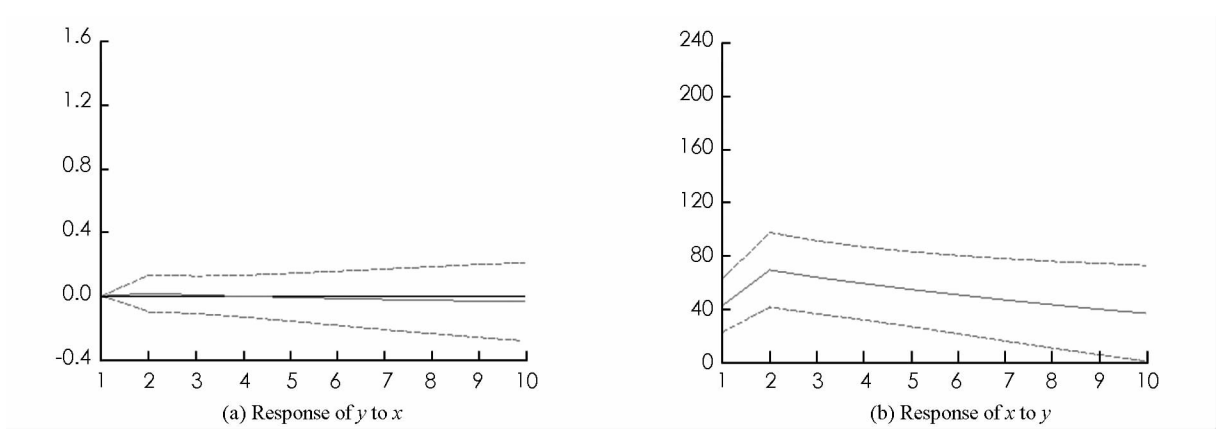


图 12 股票代码: 000 788 脉冲响应函数结果

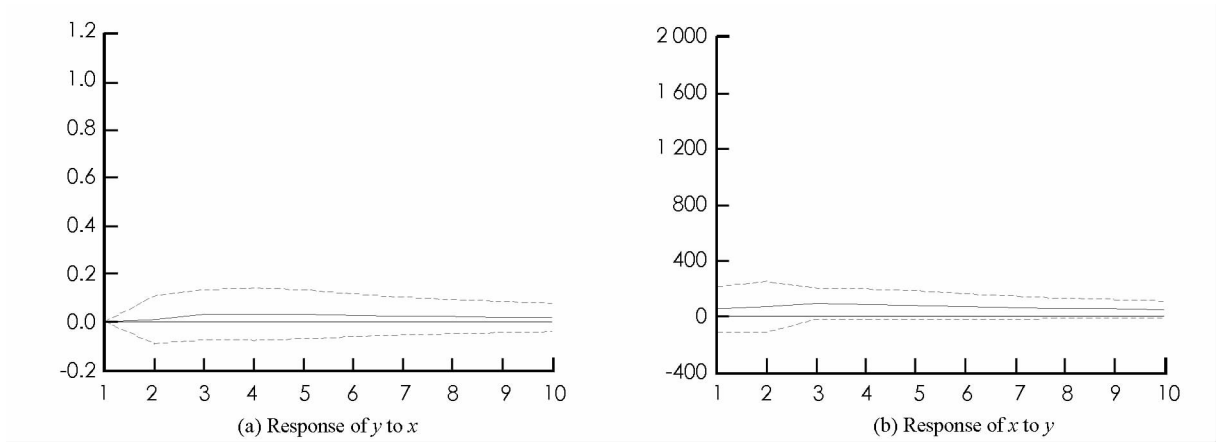


图 13 股票代码: 000 039 脉冲响应函数结果

由图 14(a)可以看出, 对于关注度指数的冲击, 股票收盘价在初期基本无响应, 到第 2 期时表现出一定的正向响应, 到第 3 期达到最高值, 直到第 8 期开始有所下降, 总体表现为正向响应. 而关注度对股票收盘价的反应也是正向的长期性的, 如图 14(b). 总体来看, 股票收盘价与关注度两者的反应都是正向的, 响应情况也比较平稳.

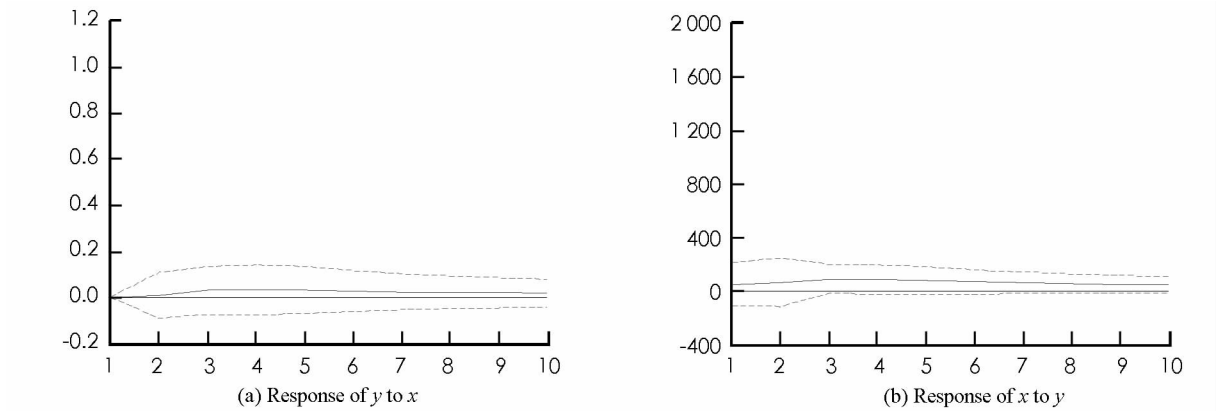


图 14 股票代码: 300 005 脉冲响应函数结果

由图 15(a)可以看出, 对于关注度指数的冲击, 股票收盘价在初期基本无响应, 第 2 期开始表现出一定的正向响应, 随后一直趋于上升阶段, 到第 10 期达到最高值, 总体表现为正向响应. 而关注度对股票收盘价的反应初期为正, 在第 6 期达到最高值后上升基本无变化, 如图 15(b). 总体来看, 股票收盘价与关注度两者的反应都是正向的, 响应情况也比较明显.

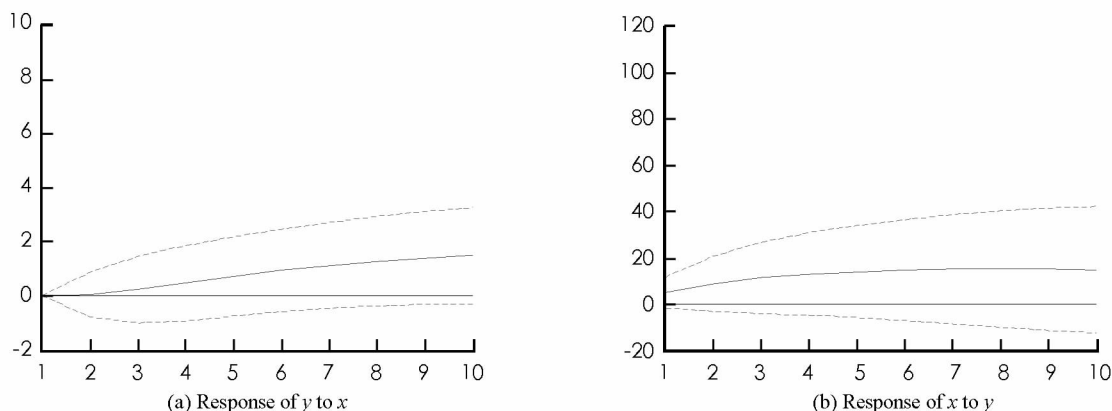


图 15 股票代码: 300 459 脉冲响应函数结果

综合上面的函数结果图分析, 股票收盘价与关注度指数之间存在长期的影响, 而不同性质的股票的响应情况有明显差异. 我们得出以下结论: 首先, 主板市场中的股票收盘价受关注度的影响程度上, 一线蓝筹股明显比二线蓝筹股强烈, 中集集团(股票代码: 000039)的股票收盘价受关注度的影响程度明显低于首创股份(股票代码: 600008). 其次, 创业板市场中股票收盘价与关注度双方表现出正向响应, 而且两者响应程度基本一致. 最后, 我们将主板市场与创业板市场的股票结合在一起分析比较, 发现关注度对股票收盘价有显著的正向影响, 那些市场形象良好、公司收益稳定的上市公司的股票收盘价对关注度的响应程度要低于经营风险高、处于发展阶段的公司.

总结上文的实证结果: 关注度变化对股票收盘价有显著的正向的长期的影响.

2.6 关注度的内生性与外生性

我们通过格兰杰因果检验证实了关注度是引致股票收盘价变化的原因, 然后利用脉冲响应进一步解释关注度冲击与股票收盘价之间的关系, 反映了模型受到关注度冲击时对股票收盘价的影响具有长期性. 主板市场和创业板市场中, 关注度冲击时对股票收盘价的影响与股票收盘价冲击时对关注度的影响都具备长期性. 创业板股票的股票收盘价与关注度的响应程度基本一致, 即关注度会影响收盘价, 而收盘价也会对关注度产生影响. 而主板市场中, 一线蓝筹股的关注度对收盘价的冲击会出现正向反应, 二线股票也会出现收盘价引起关注度变化, 同时关注度也会引起收盘价出现正向反应的现象, 三线蓝筹股的情况没有前面两者显著.

综合分析, 主板市场和创业板市场股票的关注度对股票价格表现出内生性, 对一线蓝筹股这类公司形象好、收益稳定的股票, 关注度对股票收盘价有外生性, 但影响程度不大. 对于那些经营风险高、公司规模小、仍处于发展阶段的公司, 关注度对股票收盘内生性程度低于外生化程度, 关注度总体表现为内生变量.

3 相关结论

通过建立相关模型进行实证分析, 我们发现了一些较有意义的结论, 主要总结为:

- 1) 本文通过建立格兰杰因果检验印证了关注行为对股票价格有显著的影响这一假设, 结果表明关注度是股票收盘价变化的原因.
- 2) 实证发现, 主板市场中, 股票收盘价受关注度的影响程度要高于创业板市场. 采用脉冲响应函数分析, 得到关注行为对股票价格会产生持续性长期的影响, 关注度能够在投资者的决策过程中产生影响.

参考文献:

- [1] 林 莉. 证券市场异象的行为金融学解释以及对我国的启示 [J]. 财经界(学术版), 2015(15): 7.
- [2] 陈奇斌. 尤金·法马与有效市场理论 [J]. 福建论坛·人文社会科学版, 2014(1): 38—41.
- [3] 刘敬亮. 中国股票市场个人投资者的非理性行为分析 [J]. 赤峰学院学报(自然科学版), 2016, 32(1): 97—99.
- [4] SEASHOLES M S, WU G. Predictable Behavior, Profits, and Attention [J]. Journal of Empirical Finance, 2007,

14(5): 590—610.

[5] TETLOCK P C. Giving Content to Investor Sentiment: The Role of Media in the Stock Market [J]. Social Science Electronic Publishing, 2007, 62(3): 1139—1168.

[6] 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, 46(S1): 145—155.

[7] 姜翔程,熊亚敏. 基于 GARCH 族模型的我国股市波动性研究 [J]. 西南师范大学学报(自然科学版), 2017, 42(2): 115—119.

[8] 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.

[9] 廖 微. 基于百度指数的投资者关注与股票市场量价研究 [D]. 广州: 广东商学院, 2013.

[10] 高红霞,梅 辉,谭显中,等. 股票日收盘价的协整与因果关系实证研究 [J]. 重庆理工大学学报(自然科学版), 2017, 31(4): 145—149.

[11] 肖丁丁,朱桂龙,王 静. 政府科技投入对企业 R&D 支出影响的再审视——基于分位数回归的实证研究 [J]. 研究与发展管理, 2013, 25(3): 25—32.

[12] 袁 航. 基于 VAR 模型对我国城镇居民消费和收入之间关系的研究 [J]. 金融经济, 2014(8): 85—86.

On Relationship between Baidu’s Attention Index and Stock Price

ZHANG Wu-yi, LI Jie

School of Management and Economics, Kunming University of Science and Technology, Kunming 650093, China

Abstract: In his paper, the contact between limited attention index of investors and the stock market has been analyzed, the literature reviewed, and the differences between traditional and modern financial theory compared. The limited attention and modern theoretical models have been introduced, which support them according to the model. The lack of risk perception of investors can directly influence their behaviors. Comparing with traditional proxy indicators of investor attention including exchange rate, volume, and so on, proxy indicators provided from the Internet can reflect the concerns of investors more directly. Baidu index is used as a proxy variable for the attention of investors. When investors are searching for information about the stock on the Internet, undoubtedly, it means that they concern with this stock. The number of retrieved keywords provided by Baidu index directly reflects the attention investors. Baidu index will also affect investors’ transactions and lead to the stock market volatility. In order to analyze the relationship between the attention of investors and the price of the stock, we establish the VAR model, Granger and Impulse Response Function. The result shows that there is a relationship between the attention of investors and the stock market.

Key words: behavioral finance; limited attention; Baidu index; stock market

责任编辑 汤振金