

# 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

俞庆进 张 兵

(南京大学工程管理学院, 江苏南京 210093)

**摘 要:**搜索引擎是当今人们获取信息的入口,百度指数反映了关键词被搜索的次数,直接衡量着投资者的有限关注。本文实证检验了百度指数和创业板股票市场表现的相关性,在此基础上,验证了投资者有限关注能影响股票的市场交易活动。投资者的有限关注能给股票带来正向的价格压力,而这种压力会很快发生反转。本文还利用百度指数逐日可测的优点,研究发现非交易日的投资者关注将显著影响下一交易日股票集合竞价时的价格跳跃。

**关键词:**有限关注;百度指数;股票收益

**JEL 分类号:**G14,G12,G10 **文献标识码:**A **文章编号:**1002-7246(2012)08-0152-14

## 一、引 言

据日本《信息流通调查报告》估计,人类标准供给信息量每 10 年约增加 4 倍,而个人消费量几乎没有大的变化。这种海量信息供给与个体有限的消费能力之间的矛盾,正如 Herbert Simon 所提出的“信息的丰富导致了注意力的稀缺”,在这样的时代环境下,对投资者有限关注的研究越发显得重要。

传统的经济理论认为股票市场是有效的,股票价格的短期波动是对市场信息的反应,而长期的价格偏离是对投资者承担风险的补偿(Baker 和 Nofsinger, 2010)。在投资者理性与市场有效性基础之上,Markowitz(1952)提出了经典的资产组合理论。然而实证却发现很多经典理论无法解释的市场“异象”,如“过度自信”、“盈余公告效应”、“媒体效应”等。Kahneman(1973)建立了行为经济学理论,对上述市场“异象”进行解释和研究。

有限关注是行为金融的一个核心概念,该理论认为由于时间和精力限制,投资者往

收稿日期:2012-02-16

作者简介:俞庆进,南京大学工程管理学院研究生,Email:qingjinyu@gmail.com.

张 兵,经济学博士,南京大学工程管理学院教授。

\* 作者感谢匿名评审人的建设性建议,当然文责自负。本文感谢国家自然科学基金项目(批准号:70932003 和 71171108)的资助。

往不能及时地获取,并充分地理解市场上的信息。个体只会对吸引其关注的信息加以分析和判断,调整自己的投资行为,从而导致了暂时性定价偏差的发生。只有引起投资者关注的信息才能对市场产生影响,故而是一个合适的投资者关注度代理变量,不仅能反映资本市场运行的基本状况,在某种程度上,还可以预测未来市场的发展和走向。

传统投资者关注度的代理变量主要包括超额收益、交易量、换手率、媒体报道和广告支出等,最近有学者开始利用互联网工具研究投资者关注,搜索引擎作为互联网上主要的信息获取渠道,被广泛地应用在日常生活信息获取当中。2008年CNNIC调查显示,经常使用搜索引擎的网民占全体网民的72.4%,而且学历越高,搜索引擎使用率越高。证券投资者利用搜索引擎挖掘相关公司的股票信息已经是一种普遍的行为。当投资者搜索一只股票的信息时,毫无疑问,该投资者对这只股票产生了关注。搜索量是某个关键词在搜索引擎中被检索的次数,这种数据更为直接和精确地衡量了投资者对具体某只股票关注的程度。

和以往研究相比,本文的独特价值在于,第一,首次利用百度指数来衡量国内个体投资者的关注度,解决了前人研究中投资者关注不能直接衡量的缺陷,并且验证了百度指数作为投资者关注度代理变量的可行性。第二,实证发现中国投资者的关注对股票市场有正向的价格压力,但这种压力会迅速反转,本文对这个动态过程的短期效应进行了检验,发现了投资者有限关注指标具有区别于其他市场指标的独特性。第三,研究了非交易日的投资者关注对下一交易日市场开盘价格跳跃的影响,拓展了百度指数这个投资者关注代理变量潜在的应用价值。

全文的框架如下:第二部分是在文献回顾基础上,提出相应的研究假设;第三部分是样本选择和变量定义,第四部分是实证结果与分析;第五部分是全文的结论。

## 二、文献回顾和研究假设

### (一)投资者关注度代理变量的研究进展

投资者有限关注理论的研究属于新兴领域,难点之一在于“关注”较难进行衡量(贾春新等,2010)。传统的投资者关注度代理变量主要包括超额收益、交易量、媒体报道、换手率和广告支出等。Seasholes和Wu(2007)认为涨停报道会吸引投资者关注,他们利用涨停板事件作为投资者关注度的代理变量,实证发现股票交易涨停事件会影响投资者注意力配置。Peng和Xiong(2007)利用成交量研究投资者关注,发现市场信息与个股信息对注意力的竞争。Lou(2010)利用美国股票市场1974~2006年间上市公司广告支出数据进行实证检验,发现广告支出会增加投资者的买入,提高股票的回报。Chemmanur和Yan(2010)基于投资者有限关注理论,分别研究了广告支出对股票的短期回报和长期回报的影响。张雅慧等(2011)研究发现股票收入的“媒体效应”可以用过度关注弱势理论来解释。贾春新等(2010)利用谷歌历史资讯作为投资者有限关注的度量指标,研究了限售股解禁报道对股票收益的影响。张永杰等(2011)研究了互联网开源信息对资产定价

的影响,发现互联网信息已经蕴含对中国股市日异常收益率有显著解释能力的信息内容。李小晗和朱红军(2011)利用不同周历阶段作为投资者关注度的表征变量,研究了投资者有限关注对信息解读效率的影响。

然而 Da, Engelberg 和 Gao(2011)指出,利用上述代理变量对投资者关注的研究,都存在一个关键的假设,即这些代理变量反映的现象必须是投资者的关注所造成的。超额收益、成交量、换手率、涨停事件等都是金融资产本身的交易特性和价格行为,并不能直接反映公众注意力的分布和强弱,这些变量的波动有可能是由与投资者关注无关的因素所导致。而媒体报道、广告支出、互联网资讯等和投资者关注并不直接对应,它们只有在被投资者收听或阅读时,才能真正转化为投资者的关注。Da, Engelberg 和 Gao(2011)采用谷歌趋势的周度数据来度量投资者关注,宋双杰等(2011)也利用谷歌趋势的数据作为投资者关注度的代理变量研究了中国股票市场的 IPO 异象,这为投资者有限关注的研究提供了新的思路。

投资者关注一只股票,并通过搜索希望深入了解。相关信息搜索量的高低反映了投资者的关注程度,而通过搜索得到的信息的质量和数量也必然会影响到投资者的投资期望和投资行为,进而影响资产价格的波动,最终表现出来就是搜索指数和市场指标有很强的相关性。另外,市场指标的突出表现也会引发投资者的关注和搜索,从另一个方向加强了这种相关性。这种相关性的存在,是我们研究投资者有限关注和股票收益二者关系的基础。中国投资者构成类型、组织结构和国外不同,目前国内中小投资者所占市场比例远高于国外市场,相较于机构投资者而言,中小投资者更多的采用搜索引擎收集信息。百度在国内搜索市场占据绝对的强势地位,在国内个体投资者中普及率非常高,因而百度指数可以有效刻画国内个体投资者对具体某只股票的关注程度。故本文提出如下假设:

H1: 百度指数代表的投资者关注度和股票市场表现指标具有相关性。

## (二) 投资者的有限关注对股票市场的影响

金融理论研究中,关注可以分为关注和有限关注,前者是指投资者因为特定事件而对相应股票产生偏离基本面的过度反应。后者是指由于时间和精力有限性,限制了投资者所能分析的信息量,使得投资者不可能考虑所有的股票,从而导致了暂时性定价偏差的发生(Aboody 等,2010)。Kahneman(1973)最先提出“有限关注”的概念,他认为个体对特定事物的关注是以减少对其他事物关注为代价的,因此关注是一种稀缺资源。在经济和公司基本面等其他状况没有发生根本变化的情况下,有限关注本身会对投资决策产生重大的影响(王春和徐龙炳,2009)。信息对股票市场的影响是通过投资者的交易行为实现的,只有被投资者关注到的信息才能通过交易反映到股票价格当中,因而投资者的关注是市场反应的前提条件(权小锋和吴世农,2010)。Della Vigna(2009)发现,有限关注可以用来解释许多被传统金融学理论视为“异象”的问题。例如信息内容上等价而仅是形式上不同也会对投资者行为产生不同的影响(Hirshleifer 等,2008);投资者由于时间精力的限制,不能有效处理所有可得信息,以致对股票基本面的相关信息反应不足(Engelherg 等,2010)。根据以上文献,我们提出如下假设:

H2:有限关注下,投资者关注本身会影响股票的市场表现。

Madhavan(1992)研究指出交易不活跃的股票,其信息的不对称程度越大,越是倾向于在集合竞价市场中交易。周五收盘之后,经过周六和周日两天,市场基本面情况可能发生了较大的变动,这种市场信息的变动,会使得投资者的关注有较大的改变。因为没有交易产生,这种改变无法反映在传统市场指标(收益率、成交量等)当中。但投资者可以通过搜索引擎继续挖掘自己关注的股票的信息,做出相关投资决策,这种关注度带来的价格压力会表现在下一交易日股票开盘时的集合竞价上。个体投资者的关注和搜索,被百度指数所捕捉,从而使得这种关注度可以衡量,研究非交易日的投资者关注对股票市场开盘的影响也就有了可能。因此我们提出如下假设:

H3:非交易日的投资者关注会反映在下一交易日股票市场开盘的价格跳跃中。

在投资者有限关注下,那些能够吸引投资者注意力的股票能否获得超过市场平均收益的溢价,一直是学术界关心的热点。Barber和Odean(2008)认为没有足够的精力对所有可供选择的股票进行分析和比较,因此倾向于购买吸引其注意力的股票,从而使该股票价格存在暂时的上升压力。这一假设在Da等(2011)对美国股票市场的研究中获得了实证。国外基于投资者有限关注的研究记录了完整的股票价格变动趋势(Seasholes和Wu,2007),但是国内相关研究大多止步于投资者关注带来的被关注股票的价格上升,而较少对后续的反转进行讨论(张雅慧等,2011)。饶育蕾等(2009)发现媒体对上市公司的关注度越高,在接下来的一个月中,其股票的平均收益率越低,并且认为这一现象与基于有限注意的过度关注弱势假设一致。杨晓兰(2010)的研究也发现和讯网络关注度对股票的当日收益率产生了正向影响,而次日这种影响就变成了负向。

当前我国股票市场信息披露机制尚不完善,而且没有卖空机制,大量个体投资者倾向依据内幕消息等这类吸引其注意力的信息进行买入操作。在个体投资者发生“基于关注度的购买行为”之后,持有高关注度股票的理性投资者和机构投资者可能会意识到其价格中包含的关注度溢价,出售股票获取溢价收益。因为前面的高关注度带来的溢价并没有基本面信息的支撑,个体投资者的关注转移后,不再有买家跟进,之前关注度越高,之后价格回落幅度越大。因此我们提出如下假设:

H4:有限关注下,投资者关注对股票当期收益有正向影响,但该影响会在近期反转。

### 三、样本选择和变量定义

#### (一)样本选择

本文样本选自于深圳证券交易所创业板。选择创业板作为研究对象,主要是从百度指数的代表性方面考虑。由于主板市场公司规模较大、业务较多,许多非证券投资者也可能使用百度搜索这些公司的信息,搜索指数刻画的投资者关注目前阶段代表性可能还不够。创业板上的公司多数比较年轻,获取其资讯的渠道较少。而且创业板公司一般属于“两高六新”的企业,容易成为某些题材、概念的炒作对象,个体投资者通过搜索寻找相关

信息的可能性更大,搜索指数能更好的刻画该市场投资者的关注度。

创业板上市的股票规模偏小,基金公司投资比例的限制以及风险控制的考虑,投资创业板的规模非常小。深交所金融创新实验室 2011 年 3 月发布的《2010 年度股票市场绩效报告》显示,2009 年创业板市场交易中,机构投资者成交金额占市场比例仅为 2.24%,2010 年来有所上升,但占比也仅为 8.45%。创业板市场交易主体是中小投资者,这提供了天然的试验场,极其适合研究中小投资者的关注。

在时间区间的选择上,本文选用从 2011 年 4 月 1 日到 2012 年 3 月 31 日这一个月度的时间区间。我国股票创业板市场开板较晚,为了获得足够宽的市场截面数据,将时间起点选择为 2011 年 4 月 1 日。数据频率上,目前国内外的研究都是基于谷歌趋势提供的周数据,而本文利用的是百度指数的日度数据<sup>①</sup>,因而可以在较短的时间区间内得到足够多的时间序列观测值。

2011 年 4 月 1 日前,创业板共有 196 家公司上市,这 196 支股票样本上市时间在百度指数服务推出之后,可以得到完整的历史搜索数据。同时我们剔除了部分简称具有明显二义性的 6 支股票<sup>②</sup>,以及搜索数据缺失的 21 支股票,共得到创业板上 169 支不同批次上市的股票样本。这样我们考察的样本实际已经涵盖了大部分创业板股票,可以保证结果的可靠性。

## (二) 投资者关注度指标——百度指数

本文选择的是用户在百度中对某一个关键词搜索量的加权和,即百度指数的用户关注度,作为投资者关注度的代理变量。百度公司在国内搜索引擎市场处于垄断地位,据艾瑞咨询发布的《2010-2011 年中国搜索引擎年度监测报告》显示,2010 年第四季度,百度的网页搜索请求量市场份额达到 83.6%。相较于采用谷歌搜索的数据对国内股票市场的研究,采用百度指数的数据作为投资者关注度的代理变量具有更好的代表性<sup>③</sup>。

个体投资者利用百度检索投资对象的相关信息时,他可能使用上市公司的证券简称、证券代码或公司名称进行搜索,不同关键词的搜索量是不一样的。公司名称搜索数据的噪音比较大,因为用户在搜索公司名称时,可能是查询公司地址、公司招聘等其他信息。而搜索证券简称和证券代码的用户,更可能是该公司的潜在投资者,这种关注也正是我们所要分析的个体投资者的有限关注。Da 等(2011)的文章中用股票代码(ticker)的搜索量作为投资者关注度,宋双杰等(2011)利用公司名称构建投资者关注度的代理变量。考虑到国内股票市场基本上是代码和简称并重<sup>④</sup>,本文基于证券简称和证券代码搜索量之和构建了投资者关注度的指标:

① 百度公司没有提供可以直接下载的数据,作者利用 matlab 程序和手工整理完成了原始数据的收集,有意者可以发信索取原始数据和程序代码。

② 被剔除的六支股票分别为:机器人,同花顺,东方财富,乐视网,向日葵,汤臣倍健。这些简称或者是有约定俗成含义,或者是有相应互联网产品,搜索量和其他股票简称相比,基本上高一个数量级以上。

③ 事实上,用于用户搜索量少,谷歌趋势无法提供大部分创业板股票的搜索量指数。

④ 表 1 中数据描述性统计中可以发现,股票简称和股票代码的搜索量基本上在同一数量级水平。

$$AT_{i,t} = \log(\text{NameIndex}_{i,t} + \text{NumberIndex}_{i,t})$$

其中  $\text{NameIndex}_{i,t}$ ,  $\text{NumberIndex}_{i,t}$  分别代表第  $t$  期第  $i$  支股票的股票简称和股票代码的百度指数,  $AT_{i,t}$  为相应的投资者的关注度变量, 其值为前二者搜索指数之和的自然对数。

通过百度指数提供的相似搜索词功能可以发现, 股票简称和股票代码往往不是相似关键词, 这说明投资者使用搜索引擎检索股票信息的习惯比较固定, 少有投资者会混用股票名称和股票代码进行搜索。因此股票简称和股票代码分别得到的各自搜索量虽有少量重合的部分, 但其整体才构成投资者关注的完整指标。

(三) 市场交易指标——收益率、换手率和成交量

对创业板股票市场表现的考察, 我们主要选择了四组市场指标, 分别是第  $i$  支股票第  $t$  期的:  $A\text{Ret}_{i,t}$  (个股日异常收益率)、 $AbsA\text{Ret}_{i,t}$  (个股绝对值日异常收益率)、 $T\text{vol}_{i,t}$  (个股成交量)、 $D\text{trd}_{i,t}$  (流通股日换手率)。数据来自于 wind 数据库。本文超额收益率采用市场调整模型<sup>①</sup>进行计算:

$$A\text{ret}_{i,t} = \text{Ret}_{i,t} - \text{MRE}_t$$

$$AbsA\text{ret}_{i,t} = | \text{Ret}_{i,t} - \text{MRE}_t |$$

$\text{MRE}_{i,t}$  为创业板指数日收益率,  $\text{Ret}_{i,t}$  为个股日收益率。考虑到引发投资者关注的信息本身有利好和利空两种属性, 都可能对应其关注程度的提高。而这两类消息使得投资者对某只股票存在着看涨和看跌两种预期, 故而引入收益率的绝对值指标, 捕捉这两种方向的预期对市场的影响。由于百度指数提供的是逐日可测的观测数据, 与市场指标有很好的对应, 所以并没有对这些指标做更多的变换。主要变量的描述性统计见表 1。从表中可以发现, 关注度指标的标准差较小, 说明从样本总体而言, 各只关注度水平差异不大。

表 1 主要变量描述性统计

|                    | 均值       | 中位数       | 标准差     | 最小值         | 最大值      |
|--------------------|----------|-----------|---------|-------------|----------|
| <i>NameIndex</i>   | 232. 44  | 180. 00   | 191. 25 | 8. 88       | 3807. 69 |
| <i>NumberIndex</i> | 203. 94  | 168. 92   | 124. 04 | 9. 23       | 1766. 15 |
| 投资者关注              | 5. 9467  | 5. 8684   | 0. 4711 | 4. 3220     | 8. 3397  |
| 绝对值日异常收益率          | 0. 0133  | 0. 0095   | 0. 0137 | 9. 43E - 07 | 0. 1362  |
| 异常收益率              | -0. 0001 | -0. 0012  | 0. 0191 | -0. 1192    | 0. 1362  |
| 成交量(股)             | 1387418  | 721824. 5 | 2109949 | 14983       | 41549619 |
| 换手率(%)             | 2. 7181  | 1. 7102   | 3. 6702 | 0. 0592     | 55. 8504 |

① 在计算超额收益时, 除文中采用的直接减去指数收益外, 我们同时采用了 CAPM 模型、Fama - French 三因素模型计算的基准收益进行了稳健性的分析, 发现显著性和系数符号基本没有变化。

#### (四) 非交易日市场影响的衡量指标

周五收盘之后,至下周一市场开盘集合竞价时,股票价格较上周收盘会发生较大变化。本文从价格和收益两个角度,定义了四个市场跳跃指标:

$$\begin{aligned} Dprice_{i,t} &= Oppmon_{i,t} - Clpfri_{i,t-1} \\ AbsDprice_{i,t} &= |Oppmon_{i,t} - Clpfri_{i,t-1}| \\ Gapret_{i,t} &= \frac{Oppmon_{i,t} - Clpfri_{i,t-1}}{Clpfri_{i,t-1}} \\ AbsGapret_{i,t} &= \left| \frac{Oppmon_{i,t} - Clpfri_{i,t-1}}{Clpfri_{i,t-1}} \right| \end{aligned}$$

其中,  $Oppmon_{i,t}$  是股票  $i$  在第  $t$  期周一的开盘价,  $Clpfri_{i,t-1}$  是股票  $i$  在第  $t-1$  期周五的收盘价,  $Dprice_{i,t}$  是价格跳跃幅度,  $Gapret_{i,t}$  是收益率跳跃的程度,分别引入其绝对值指标  $AbsDprice_{i,t}$ ,  $AbsGapret_{i,t}$  是为了捕捉两个方向的价格跳跃。

### 四、实证模型与结果分析

#### (一) 平稳性检验

本文重点关注的六个市场指标分别是有限关注程度(AT)、收益率(Ret)、超额收益率(ARet)、绝对值超额收益率(ARet)、换手率(Dtrd)、成交量(Tvol)。为了避免出现伪回归问题,我们对这 169 只股票组成的面板数据的每个指标都进行了单位根检验。所有指标的 LLC(Levin-Lin-Chu)检验和 ADF 检验的结果都在 1% 水平下拒绝存在单位根为零假设,说明我们选取的时间序列是平稳的。

#### (二) 市场指标与投资者关注的相关性分析

本文首先对市场指标和投资者关注做单变量回归,检验百度指数所代表的投资者的关注是否与股票的市场表现显著相关。其次,根据张永杰等(2011)研究,互联网开源信息内容蕴含了市场指标所未包含的信息,因而我们还希望检验搜索指数指标是否同样蕴含着其他市场指标所未能反映的市场信息。

根据以上两个检验目标,我们设计了如下两个模型,见公式(1)和(2)。依据 F 检验和 Hausman 检验的结果,采用固定效应模型进行回归分析。模型(1)检验了同期的市场指标和百度指数的相关性,模型(2)是在利用历史信息解释当期收益的普通 AR 模型基础上,增加滞后一期的百度指数项,考察滞后的投资者关注度对市场指标的解释效果。其中 AR 项的滞后阶数通过 AIC 和 SBIC 的信息准则确定。

$$MarInd_{i,t} = m_1 + \alpha_{i,1} + \beta_1 AT_{i,t} + \varepsilon_{i,t} \quad (1)$$

$$MarInd_{i,t} = m_2 + \alpha_{i,2} + \sum_{\lambda=1}^{\lambda=n} \gamma_{i,\lambda} MarInd_{i,t-\lambda} + \beta_2 AT_{i,t-1} + \varepsilon_{i,t} \quad (2)$$

其中:  $MarInd_{i,t}$  是第  $i$  支个股在第  $t$  期的市场指标,分别代表超额收益率、绝对值超额收益率、换手率、成交量,其中为了避免异方差的出现,对成交量数据取对数处理;  $AT_{i,t}$  是第

$i$  支个股在第  $t$  期的百度指数,  $m$  代表总体均值截距项,  $\alpha_i$  表示截面对总体均值偏离的截面截距项,  $\varepsilon_{i,t}$  是回归残差。回归结果见表 2。

表 2 模型 1,2 回归结果  $\beta$  系数统计

|    | 因变量     | $ARet$      | $AbsAret$  | $Dtrd$     | $Tvol$     |
|----|---------|-------------|------------|------------|------------|
| 模型 | $\beta$ | 0.0073 ***  | 0.0185 *** | 7.8746 *** | 1.8579 *** |
| 1  | $t$ 统计量 | 21.51       | 83.13      | 186.25     | 169.51     |
| 模型 | $\beta$ | -0.0026 *** | 0.0064 *** | 0.9165 *** | 0.2292 *** |
| 2  | $t$ 统计量 | -7.76       | 25.18      | 19.76      | 20.66      |

注:  $\beta$  值即为模型(1)和(2)中投资者有限关注度的系数。\*、\*\*、\*\*\*分别代表在 10%、5%、1% 的水平下显著,下文同此注

模型(1)的回归中,将市场指标作为因变量,投资者关注度作为自变量,我们发现所有股票的市场指标和百度指数在 1% 显著性水平下相关,尤其是绝对值异常收益率、换手率和成交量指标,  $t$  值都超过 50, 显示出其与百度指数非常强的相关性。这说明了投资者关注程度越高,股票的市场表现就越活跃。这一结果验证了本文的假设 1, 百度指数代表的投资者关注度和股票市场指标具有显著的相关性。这同时也说明了百度指数作为个体投资者关注度指标一定程度上的合理性,为我们下面的研究打下坚实的基础。

模型(2)的回归中,使用了 AR 项来反映市场的历史信息,用来检验滞后一期的投资者历史关注是否蕴含了还未表现在其他市场指标中的信息。模型 2 的回归结果显示,尽管已经包含了历史信息项,滞后一阶的百度指数,即投资者的历史关注还是和市场指标显著相关。这一结果验证了假设 2 中,有限关注下,投资者关注本身就能影响股票市场的表现这一推断。历史的投资者关注可以一定程度预测股票未来的表现,这一结果凸显了百度指数作为投资者关注度的代理变量区别现有市场指标的独特价值。随着互联网应用的深入,我们可以期待这一投资者关注的代理变量会对股票收益有越来越强的解释能力。

### (三) 投资者关注度对市场交易活动的影响

从上节实证中,我们已发现投资者关注和股票的成交量、换手率存在着显著的相关性。由于市场可能存在着交易量放大,但是股票价格在盘中调整的情况,故而收益率指标并不能很好的衡量市场交易活动程度的强弱。为此,我们选择了成交量和换手率这两个指标,研究投资者关注对市场交易活动的影响。

根据 Fama 和 French(1992) 的三因子模型,股票的收益和公司规模、市场组合收益率和市净率三个因素有关。据此,我们分别构建了以换手率( $Dtrd$ )和成交量( $Tvol$ , 取对数后值)为被解释变量,以公司市值( $SCA$ , 取对数后值)、市场组合收益率( $MRE$ )、市净率( $PB$ )为控制变量,以投资者关注( $AT$ )为解释变量的面板回归模型。依据 Hausman 检验结果,我们采用固定效应模型进行面板回归,具体模型如下:

$$MarInd_{i,t} = \alpha_i + \beta_0 + \beta_1 \cdot AT_{i,t-k} + \beta_2 SCA_{i,t} + \beta_3 PB_{i,t} + \beta_4 MRE_t + \varepsilon_{i,t} \quad (3)$$

$MarInd_{i,t}$  为第  $i$  支股票第  $t$  期的市场交易指标,这里分别代表换手率和成交量,  $SCA_{i,t}$ ,



$PB_{i,t}$ , 分别为股票  $i$  在第  $t$  期的公司规模, 市净率,  $MRE_t$  采用的是创业板指数的日收益率作为市场组合收益率, 仅随时间  $t$  变化, 不受个股  $i$  影响。和前文类似,  $\beta_0$  代表总体均值截距项,  $\alpha_i$  表示截面对总体均值偏离的截面截距项,  $\varepsilon_{i,t}$  是回归残差。 $AT_{i,t-k}$  为个股  $i$  在  $t$  期时, 对应的投资者关注度变量滞后  $k$  期值。

面板回归结果整理如表 3, 我们分别将关注度变量滞后了 0 至 4 期 (分别用  $AT(-k)$  表示,  $k$  为滞后期数), 还考察了滞后一周和滞后的第二周投资者关注度对市场交易活动影响<sup>①</sup>。在控制了市场组合收益率、个股规模、市净率等影响因素后, 投资者关注度对近期市场交易指标存在显著的正向影响。不仅是当期的投资者关注, 最近两周的投资者关注都会显著提高当期股票的市场成交量表现。滞后的投资者关注对当期的成交量和换手率有显著的正向影响, 验证了假设 2: 投资者关注本身会影响股票的市场交易表现。投资者关注了某只股票, 通过百度搜索查询相关信息, 然后在市场买卖该股票, 从而引起了市场指标的变化, 而不只是股票的市场表现突出, 引发了投资者的搜索和关注。从表 3 结果还可以发现, 投资者的有限关注对个股交易活动的影响是随时间推移而衰减的。

表 3 投资者关注度与换手率、成交量面板回归系数统计表

| 关注度  | AT                      | AT(-1)                  | AT(-2)                  | AT(-3)                  | AT(-4)                 | $\frac{1}{n} \sum_{k=1}^{n=5} AT(-k)$ | $\frac{1}{n} \sum_{k=6}^{n=10} AT(-k)$ |
|------|-------------------------|-------------------------|-------------------------|-------------------------|------------------------|---------------------------------------|--|
| Dtrd | 7. 871 ***<br>(177. 06) | 6. 738 ***<br>(137. 46) | 5. 593 ***<br>(106. 33) | 4. 990 ***<br>(92. 13)  | 4. 504 ***<br>(81. 54) | 6. 885 ***<br>(115. 39)               | 4. 242<br>(64. 27)                     |
| Tvol | 1. 967 ***<br>(172. 14) | 1. 715 ***<br>(138. 17) | 1. 487 ***<br>(113. 11) | 1. 365 ***<br>(101. 31) | 1. 260 ***<br>(91. 88) | 1. 854 ***<br>(125. 65)               | 1. 236 ***<br>(75. 89)                 |

(四) 非交易日投资者关注对下周市场开盘影响

本节和前面所用的关注度数据有所区别, 投资者关注度变量用  $SAT_{i,t}$  代表, 其值为股票  $i$  在第  $t$  期周六周日两天股票简称和代码的搜索指数之和的对数<sup>②</sup>。具体模型见公式 (4), 非交易日没有其他市场指标, 这里只考虑投资者关注度的影响, 做的是单变量回归。与前文一致, 根据  $F$  检验和 Hausman 检验的结果, 选择固定效应模型进行回归。其中, 被解释变量  $JumpInd_{i,t}$  为价格跳跃指标, 分别代表  $Dprice_{i,t}$ ,  $Gapret_{i,t}$ ,  $AbsDprice_{i,t}$ ,  $AbsGapret_{i,t}$ 。变量都通过了单位根检验, 拒绝存在单位根的零假设。回归结果如表 4。

$$JumpInd_{i,t} = \alpha_i + \beta_0 + \beta_1 SAT_{i,t} + \varepsilon_{i,t} \tag{4}$$

从表 4 实证结果可以发现, 一个单位的关注度提高, 对应了 0. 0975 单位的价格跳跃, 也对应着 0. 3% 单位收益率的跳跃。这验证了假设 3: 非交易日的投资者关注, 会反映在交易日股票市场开盘的价格跳跃中。投资者在非交易日对股票进行信息的搜集和分析,

① 这里对滞后一周和滞后两周的处理采用的是较为简单的  $\frac{1}{n} \sum_{k=1}^{n=5} AT(-k)$  和  $\frac{1}{n} \sum_{k=6}^{n=10} AT(-k)$  方式, 并非恰好就是该股票上一周或者上上周的关注, 但股票交易日排列并不规则, 这样处理比较直观方便。

② 事实上, 由于法定假期的存在, 周末序列和非交易日序列并不完全一致, 但二者相比差别不大。

带来的交易压力反映在次日开盘后的集合竞价上。不仅是利好消息,利空消息也会吸引投资者的关注。由于管理层倾向于在周五公布不利公告(权小锋和吴世农,2010),这增加了非交易日投资者对负面消息关注的可能,这一点也反映在我们的实证结果中,体现在以跳跃指标绝对值为被解释变量的结果更加显著。

表4 非交易日投资者关注对市场开盘影响的回归结果统计

| 被解释变量              | 价格指数                   |                         | 收益率指标                   |                      |
|--------------------|------------------------|-------------------------|-------------------------|----------------------|
|                    | <i>Dprice</i>          | <i>AbsDprice</i>        | <i>Gapret</i>           | <i>AbsGapret</i>     |
| Intercept          | -0.6588 **<br>( -4.74) | -0.5082 ***<br>( -4.64) | -0.0203 ***<br>( -6.11) | -0.0097<br>( -3.75)  |
| SAT                | 0.0975 ***<br>(4.31)   | 0.1237 ***<br>(6.94)    | 0.0030 ***<br>(5.60)    | 0.0027 ***<br>(6.39) |
| Adj R <sup>2</sup> | 0.0009                 | 0.127                   | 0.0047                  | 0.016                |
| F                  | 1.39 ***               | 7.07 ***                | 1.20 **                 | 1.71 ***             |

这一发现对于个体投资者来说,具有一定的应用价值,通过监测某一只股票前一日的关注度数据,就能一定程度上预测其次日价格跳跃的幅度。但这种观测还有缺陷,对跳跃幅度的解释能力还不够强,跳跃的方向也不能通过监测关注度直接得到。

我们前面的实证结果显示投资者关注对股票市场交易活动能产生影响,投资者的关注能导致近期股票交易量和换手率的放大。但对于当期的投资者关注和股票的市场表现来说,究竟是投资者关注引起市场表现的改变,还是市场表现引发了投资者的关注,并不能很好的区分开来。由于通过百度指数所获得的投资者关注数据不能分成股票交易期间的投资者关注和休市后的投资者关注,所以当期的百度指数可能反映的是休市后投资者针对市场表现突出的股票为下一期的投资做“功课”,因而当期投资者关注和股票收益的正相关关系无法说明投资者关注带来了当期股票收益率的提高<sup>①</sup>。

然而,非交易日的投资者关注和股票开盘价格跳跃关系的研究,为验证假设4中的投资者关注本身对股票当期收益有正向的影响这一假设,提供了绝佳的试验场。非交易日的投资者关注,不会受到下周市场开盘价格跳跃的影响。而从下周股票开盘到连续竞价开始,时间间隔很短,这段时间相应关键词的搜索次数相比周末两天的搜索可以忽略,因此从市场指标和投资者关注角度考虑,这种非交易日的关注可以视为不会受到当期市场表现影响的,价格跳跃发生当期的投资者关注度<sup>②</sup>。

我们实证结果显示非交易日的投资者关注对股票下周一的价格跳跃和收益率的跳跃

① 非常感谢审稿专家指出这一点。

② 事实上,周末的投资者关注可能受到上周五交易的影响,因而可能认为这代表的是滞后一期的投资者关注,但是结合下文的研究,我们发现滞后一期的关注对当期收益存在显著的负向影响,和当期关注的影响截然不同,因而视为当期的投资者关注更加合理。

都有显著的正向影响,由于这种正相关关系排除了收益率表现突出导致搜索量扩大这种可能,因而验证了假设 4 中的投资者关注本身对股票当期收益有正向的影响这一假设。

(五) 投资者有限关注对市场的价格压力及其反转效应

互联网上的信息更新非常迅速,一条消息被公布到网上,数小时内就会被百度抓取,并经过用户的搜索提供给用户。投资者获取信息之后,往往在当天或者数天之内就会做出交易决策。因此投资者关注对个股的价格压力在短期内就可以实现。Barber 和 Odean (2008)认为,个体投资者是受关注股票的净买入者 (net buyers of attention - grabbing stock),其关注会使股票产生暂时的超额收益。由于百度搜索指数逐日可测,这里可以很直接地观测到这种个体投资者关注带来的超额收益以及其后的收益回落。

我们构建了以日收益率 (Ret) 为被解释变量,和前文第三节类似的面板回归模型,由于需要检验反转效应,我们把关注度变量分别取了当期和滞后一期,构建了两组模型:

$$Ret_{i,t} = \alpha_i + \beta_0 + \beta_1 * AT_{i,t} + \beta_2 SCA_{i,t} + \beta_3 PB_{i,t} + \beta_4 MRE_t + \varepsilon_{i,t} \tag{5}$$

$$Ret_{i,t} = \alpha_i + \beta_0 + \beta_1 * AT_{i,t-1} + \beta_2 SCA_{i,t} + \beta_3 PB_{i,t} + \beta_4 MRE_t + \varepsilon_{i,t} \tag{6}$$

$Ret_{i,t}$  为第  $i$  支股票第  $t$  期的收益率指标,  $AT_{i,t}$ ,  $AT_{i,t-1}$  分别为当期和滞后一期的投资者关注度,其余指标和上节一致,不再赘述。

面板回归结果如表 5 所示,在控制了市场组合收益率、个股规模、盈利能力、市净率等影响因素后,投资者当期关注度与上期关注度对当期市场指标的影响存在显著差异。当期关注度对当期收益率有显著的正向影响,1 个单位的关注度能带来 0.0070 个单位的超额收益率。结合前文利用非交易日投资者关注度对股票市场影响的研究结果,我们认为这一结果进一步验证了假设 4 中当期投资者关注对股票价格有正向影响这一假设。而上期投资者关注度对当期收益率有显著的负向影响,关注度提高一个单位,当期收益率下降 0.0035 个单位,支持了假设 4 中的投资者关注度的正向影响会在近期反转这一假设。

表 5 投资者有限关注与收益率指标的面板回归结果

| 被解释变量              | Ret                       |                            |
|--------------------|---------------------------|----------------------------|
|                    | 模型 5                      | 模型 6                       |
| Intercept          | -0.1189 ***<br>( - 8.26 ) | -0.1668 ***<br>( - 11.47 ) |
| AT <sub>t</sub>    | 0.0070 ***<br>( 19.50 )   |                            |
| AT <sub>t-1</sub>  |                           | -0.0035 ***<br>( - 9.57 )  |
| SCA                | 0.0035 ***<br>( 5.12 )    | 0.0086 ***<br>( 12.47 )    |
| PB                 | 0.0001 ***<br>( 2.45 )    | 1.64E - 05<br>( 4.96 )     |
| MRE                | 0.992 ***<br>( 201.25 )   | 0.9908 ***<br>( 199.78 )   |
| Adj R <sup>2</sup> | 0.5088                    | 0.505                      |
| F                  | 242.72 ***                | 238.47 ***                 |

国外实证研究认为,投资者关注带来的超额收益会在长期反转,在我们对中国创业板市场的研究中,发现这种反转在短期就能实现。我们又对当期市场指标与滞后 0 - 4 期投

投资者的关注度分别进行了面板回归,回归结果统计见表 6。表 6 投资者关注度系数的统计中,和滞后一期结果类似,滞后二、三、四期的投资者关注度的系数全部显著为负,和当期关注度系数的符号相反。滞后期搜索量的提高反而会使得当期的收益率降低,这也支持了假设 4 当期投资者关注带来的超额收益会在短期反转这一假设。

结合前文第四节和本节的研究,当期和滞后的投资者关注对股市的影响截然不同,当期的关注能带来股票暂时的超额收益,而滞后的关注将使得这种收益回落。这种现象也说明投资者关注是独立于其他市场基本面信息而对市场产生影响的,因为如果这种正向的价格压力是公司的基本面信息带来的,那么这种价格压力有着实际的需求支撑,就不应该有如此显著而快速的反转现象的出现。

表 6 投资者关注度与收益率面板回归系数统计表

| 关注度 | AT                    | AT(-1)                   | AT(-2)                  | AT(-3)                  | AT(-4)                  | $\frac{1}{n} \sum_{k=1}^{n=5} AT(-k)$ | $\frac{1}{n} \sum_{k=6}^{n=10} AT(-k)$ |
|-----|-----------------------|--------------------------|-------------------------|-------------------------|-------------------------|---------------------------------------|--|
| Ret | 0.0070 ***<br>(19.50) | -3.48E-03 ***<br>(-9.57) | -0.0040 ***<br>(-11.16) | -0.0039 ***<br>(-10.90) | -0.0039 ***<br>(-10.80) | -0.0050 ***<br>(-11.89)               | -0.0045 ***<br>(-10.65)                |

反转现象同样出现在以媒体报道作为投资者关注代理变量的研究中(张雅慧等,2011)。但不同的是,百度指数代表的关注度数据来自投资者的自主搜索,没有媒体报道多少的差异,只能用有限关注的“过度关注弱势”理论解释:投资者对吸引其注意力的股票出现了过度反应,价格短暂上升之后,被关注的股票收益下跌是股价回归基础价值的表现。在投资者有限关注下,高关注度本身在当期会带来投资者的购买力量,从而产生正向的价格压力,出现关注度溢价。因为这种关注度并没有基本面的信息支撑,投资者获取到其他相关信息后,理性投资者可能意识到价格中包含的关注度溢价,出售股票获取溢价收益。之前关注度越高,随后价格回落幅度越大。

前文表 3 中,我们发现以换手率和成交量为被解释变量的模型中,投资者有限关注对股票当期及近期的成交量、换手率都有正向的影响,这和表 6 收益率为被解释变量的回归结果其实是一致的,因为当期关注度会带来买入压力,使得成交量和换手率放大,而滞后的关注带来的是卖出压力,也能导致成交量和换手率的放大。

学者通过对美国股票市场的研究发现关注度产生的价格反转的长期效应,如 Da 等(2011)发现,谷歌趋势衡量的关注度与美国股票两周内收益率是正相关关系,第四周关注度的影响才显著为负。而本文研究发现国内股票市场这种反转在短期内就能实现,除了研究方法不同,数据频率有差别外,这更可能是和中国个体投资者投资习惯有关。国内个体投资者多数倾向短线操作,交易频率远超于美国,存在显著的过度交易倾向(李心丹等,2002)。另外由于国内信息披露等机制不完善,金融市场上“羊群效应”也较国外更为明显,这使得国内股票市场上股票的关注度溢价远大于国外市场,因而由投资有限关注导致的溢价消失的也更快,其价格的反转效应也在更短的时间内就能显现出来。

### (六) 稳健性分析

前文研究中,我们从创业板 2011 年 4 月 1 日前上市的 196 支样本股票中选择了其中的 169 支进行研究,实证了本文的假设。为避免样本筛选中的主观判断导致我们的结论出现偏差,这里将有搜索量数据的全部 175 支备选股票纳入研究样本中,进行稳健性检验。检验同样发现股票的市场指标和百度指数代表的投资者关注都显著相关。在针对收益率模型的面板回归中,发现投资者有限关注对股票当期的收益率有正向影响,但关注度带来的超额收益会很快发生反转。在非交易日关注度对下周市场开盘影响的研究中,也同样验证了前文的结论。

另外,我们分别利用股票简称和股票代码的关注度数据,对前文假设进行稳健性分析,虽然具体数值大小上略有差异,但是结论是一致的。

## 五、结 论

本文利用百度指数作为投资者关注度的代理变量进行研究,它能直接衡量投资者对具体某只股票的关注程度。结合创业板上市较早的 196 支股票样本,我们揭示出这一代理变量和市场指标有显著的相关性。投资者的有限关注能独立对股票市场产生影响,投资者关注在当期能够对市场造成正向的价格压力,而这种压力将很快发生反转。我们还发现非交易日的投资者关注与下一交易日集合竞价时股票的价格跳跃显著相关。

搜索引擎提供的搜索量指标数值精确、易于获取,对于社会科学领域的研究来说,具有重大的应用价值。本文提出的利用百度指数代表的投资者关注,能精确地衡量投资者的有限关注程度,揭示其与股票收益的联系,为搜索量指标在国内金融领域的研究和应用做了初步的探索。

本文的研究中,还存在不少缺陷和不足,一方面由于机构投资者较少利用搜索引擎搜集股票信息,该部分投资者的关注利用百度指数无法有效捕捉,另一方面搜索引擎的应用还有一定的门槛,其所代表的关注度还没有覆盖所有个体投资者。后续研究可以进行数据的深入挖掘,更准确地把握百度指数所能代表的投资者的有限关注。

## 参 考 文 献

- [1] 贾春新、赵宇、孙萌和汪博, 2010,《投资者有限关注与限售股解禁》,《金融研究》第 11 期 108 ~ 122 页。
- [2] 李小晗和朱红军, 2011,《投资者有限关注与信息解读》,《金融研究》第 8 期 128 ~ 142 页。
- [3] 李心丹、王冀宁和傅浩, 2002,《中国个体证券投资者交易行为的实证研究》,《经济研究》第 11 期 54 ~ 63 页。
- [4] 权小锋和吴世农, 2010,《投资者关注、盈余公告效应与管理层公告择机》,《金融研究》第 11 期 90 ~ 107 页。
- [5] 饶育蕾、彭叠峰和成大超, 2009,《公众注意力是否会引起股票的异常收益?——来自中国股票市场的经验证据》, CFRN 工作论文, <http://www.cfrn.com.cn/getPaper.do?id=1609>。
- [6] 饶育蕾、彭叠峰和周相宜, 2010,《基于有限注意的排行榜效应研究——来自上海股票市场的证据》, CFRN 工作论文, <http://www.cfrn.com.cn/getPaper.do?id=2340>。
- [7] 宋双杰、曹晖和杨坤, 2011,《投资者关注与 IPO 异象——来自网络搜索量的经验证据》,《经济研究》第 S1 期。

- 145 ~ 155 页。
- [8] 王春和徐龙炳, 2009, 《投资者关注研究新进展》, 《上海财经大学学报: 哲学社会科学版》第5期 91 ~ 97 页。
- [9] 杨晓兰, 2010, 《我国股票市场的网络关注度效应——一个基于和讯关注度的实证检验》, CFRN 工作论文, <http://www.cfrn.com.cn/getPaper.do?id=2485>.
- [10] 张雅慧、万迪防和付雷鸣, 2011, 《股票收益的媒体效应: 风险补偿还是过度关注弱势》, 《金融研究》第8期 143 ~ 156 页。
- [11] 张永杰、张维、金曦和熊熊, 2011, 《互联网知道的更多么? ——网络开源信息对资产定价的影响》, 《系统工程理论与实践》第04期 577 ~ 586 页。
- [12] Aboody D, Lehavy R, Trueman B. 2010, "Limited attention and the earnings announcement returns of past stock market winners". *Review of Accounting Studies*. 15(2). pp. 317 ~ 344.
- [13] Baker H K, Nofsinger J R. 2010, "Behavioral finance: investors, corporations, and markets", Wiley.
- [14] Barber B M, Odean T. 2008, "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors". *Review of Financial Studies*. 21(2). pp. 785 ~ 818.
- [15] Chemmanur T, Yan A. 2010: "Advertising, Investor Recognition, and Stock Returns". AFA 2011 Denver Meetings Paper.
- [16] Da Z, Engelberg J, Gao P. 2011, "In Search of Attention". *The Journal of Finance*. 66(5). pp. 1461 ~ 1499.
- [17] Della Vigna S. 2009, "Psychology and Economics: Evidence from the Field" (2009)". *J. Econ. Literature*. 47(2). pp. 315.
- [18] Engelberg J, Sasseville C, Williams J. 2010, "Market Madness? The Case of Mad Money". SSRN working paper, <http://ssrn.com/paper=870498>
- [19] Fama E F, French K R. 1992, "The cross - section of expected stock returns". *Journal of finance*. pp. 427 ~ 465.
- [20] Hirshleifer D A, Myers J N, Myers L A, et al. 2008, "Do individual investors cause post - earnings announcement drift? Direct evidence from personal trades". *The Accounting Review*. 83. pp. 1521 ~ 1550.
- [21] Kaheman D. 1973, "Attention and Effort", NJ: Prentice - Hall: Englewood Cliffs.
- [22] Lou D. 2010, "Maximizing Short - Term Stock Prices through Advertising". SSRN working paper, <http://ssrn.com/paper=1571947>
- [23] Madhavan A. 1992, "Trading mechanisms in securities markets". *Journal of Finance*. pp. 607 ~ 641.
- [24] Markowitz Harry M. 1952, "Portfolio selection". *Journal of Finance*. 7(1). pp. 77 ~ 91.
- [25] Peng L, Xiong W, Bollerslev T. 2007, "Investor Attention and Time - varying Comovements". *European Financial Management*. 13(3). pp. 394 ~ 422.
- [26] Seasholes M S, Wu G. 2007, "Predictable behavior, profits, and attention". *Journal of Empirical Finance*. 14(5). pp. 590 ~ 610.

**Abstract:** We propose a direct measure of investor attention using Baidu Index. In a sample of Growth Enterprise Market, we find the investor attention can affect the stock market performance with this proxy of investor attention. An increase in Baidu Index predicts higher stock price in that day and an eventual price reversal in the next few days. It also contributes to the jump of stock price in call auction.

**Keywords:** Limited attention, Baidu Index, Stock performance

(责任编辑: 李景农) (校对: LN)