

Market Participant Sentiment and Market Pricing Efficiency: Empirical Research Based on Large-Cap Stocks and Small-Cap Stocks

Renze

School of Economics and Management, University of Electronic Science and Technology of China, Chengdu, Sichuan 610054, China

Email: 18501309995@163.com

Keywords: *sentiment of market participants, stock price synchronization, blue-chips and small-cap stocks*

ABSTRACT. This article takes the CSI 300 Index stocks and the CSI 500 index stocks in China's A-share market as research objects, and uses crawler technology to crawl the number and content of stock bar postings and quantify the results. As a variable to measure investor attention and sentiment, the results empirically analyze the impact of investor attention on the synchronization of stock prices and the difference between optimism and pessimism. We Regard the CSI 300 index as a representative of blue-chip large-cap stocks, and the CSI 500 constituents as a representative of small and medium-cap stocks, and investigate the impact of investor attention on stock price synchronization in the blue chip sector and in the small and medium sector. The conclusion proves that stocks with higher investor attention have lower stock price synchronization, and optimism is more effective in reducing stock price synchronization than pessimism, and this phenomenon is more pronounced in small and medium-cap stocks.

市场参与者情绪与市场定价效率 基于大盘蓝筹股及中小盘个股的实证研究

任泽

电子科技大学经济管理学院，中国四川省成都市，**610054**

Email: 18501309995@163.com

关键词：市场参与者情绪；股价同步性；蓝筹股与小盘股

中文摘要. 本文将我国A股市场上沪深300指数成分股以及中证500指数成分股作为研究对象，运用爬虫技术对股吧发帖数量和内容进行爬取并对所得结果进行量化处理，将其作为衡量投资者关注度和情绪的变量，实证分析投资者的关注度对个股股价同步性产生的影响以及乐观与悲观情绪的影响差别，并用沪深300指数成分股作为蓝筹大盘股的代表，用中证500成分股代表中小盘个股，研究了投资者关注度对股价同步性的影响在蓝筹板块和在中小板块中的差别。结论证明投资者关注度越高的个股其股价同步性就越低，并且乐观的情绪相较于悲观的情绪更能有效降低股价同步性，且这一现象在中小盘股中表现的更为明显。

1. 引言

行为金融学是专家学者们近年来研究的热点领域之一，它的出现对金融理论界和实证界产生巨大影响。关于对股价同步性机理，行为金融学术界主要观点为“信息效率观”，其认为，

R^2 的高低在一定程度上体现了公司不为外界所知的特质信息在该公司股价中被反映的程度： R^2 越高，则代表上市公司不为外界所获悉的特质信息或私有信息在股价中反映度就越低，所以在这种情况下，上市公司的股价同步性就越高，反之就越低[1]。近年来利用网络自媒体评论内容来衡量投资者情绪的方法得到了越来越多研究者的认可。本文基于新闻网站与社交媒体等网络平台，利用爬虫技术进行文本挖掘和情感分析技术，提取并构建投资者情绪指标。以往的研究只表明投资者关注度可整体上降低股价同步性[2]，本文在其基础之上收集了上证 300 板块和中证 500 板块中 800 支个股从 2013 年到 2018 年的发帖进行热帖数量统计以及使用网络爬虫技术对多空关键词进行提取，再使用情感分析技术确定发帖的情绪统计出看多贴和看空贴的数量进行回归分析，做出以下贡献与创新：

首先，验证这种效应在蓝筹股与中小板块之间的差别。蓝筹股在中国市场带有明显的价值性投资属性，受到机构投资者的青睐，总体来看股价运行比较稳健，而中小板个股则具有中国市场特有的“炒概念”的属性[3]，因此本文针对这两个不同板块中普通投资者关注度对股价同步性影响的差别展开了研究，并且证明在中小板块中投资者交流降低股价同步性的作用比在蓝筹股中的作用明显。

其次，进一步分析了看空和看多情绪分别的影响，表明当看多的积极情绪占比较高时对降低股价同步性的作用更明显，这可能是因为积极的情绪氛围更能提高投资者群体对个股特质信息和基本面的关注与了解，而悲观的情绪氛围会降低投资者对研究个股基本面的兴趣[4]。

2. 文献综述

近年来国内外学者从很多方面进行了研究，如张继德等（2014）认为普通投资者进行市场交易是一种注意力驱动交易行为，即对某一股信息的聚焦会使其实现在股票市场上进行频繁买卖，对于投资者关注的度量主要有传统的度量方式，包括换手率、交易量、广告支出、媒体报道等（Grullon 等，2004；Peng 等，2007；权小锋和吴世农，2010；Seasholes 和 Wu，2007；斯子文和朱叶，2016），但这些指标或是代表金融资产交易本身特性，或是基于外部信息传播会提高投资者关注的假设，并不能直接度量投资者关注程度（俞庆进和张兵，2012）；随着互联网的发展，有研究开始利用网络搜索量度量投资者关注度，这一指标能够更为直接反映投资者关注程度[5]。Da 等（2011）、宋双杰等（2011）利用谷歌搜索趋势来度量投资者关注。Dimpfl 和 Jank（2016）使用谷歌趋势中股票代码的搜索量指数来度量个体投资者对股票的关注程度。张继德等（2014）以百度指数用户关注度衡量普通投资者关注度，发现普通投资者高关注度将伴随高市场流动性，使用搜索量进行研究虽然能够有效度量投资者关注，但研究对象仅是数量上的指标。金雪军等（2013）利用股吧发帖进行文本分析，构建了看涨指数和意见趋同指数。他们发现发帖中包含的投资者意见与股票价格和交易量有显著关系。

3. 研究设计

3.1. 变量定义

3.1.1. 被解释变量：股价同步性

参考以往的研究采集利用个股、市场以及行业的周收益率数据，使用模型（1）回归得到的 R^2 作为股价同步性的衡量指标。由于 R^2 的取值范围在 0~1 之间，为了使其符合最小二乘法的回归要求，故对其进行 logit 转换，记为 $PrcSyn = \ln(\frac{R^2}{1-R^2})$ 。

$$r_{it} = \alpha_i + \beta_{i1}r_{Mt} + \beta_{i2}r_{It} + \varepsilon_{it} \quad (1)$$

式（1）中， r_{it} 表示个股 i 在第 t 周的收益率； r_{Mt} 表示整个市场第 t 周的收益率； r_{It} 表示个股 i 所属行业在第 t 周的收益率（行业内股票收益率的流通市值加权，不含个股 i ）。行业的分类标准依照证监会 2012 发布的行业分类标准。

3.1.2. 解释变量：投资者情绪

使用东方财富网股吧中个股的发帖数量和内容衡量该股的投资者情绪。具体而言，使用数据爬虫技术，提取股吧中沪深 300 指数和中证 500 指数成分股的热帖（阅读量大于等于 1000）数据，再使用情感分析技术识别发帖内容，将帖子标记为看多、看空和中立。这样统计每只股票年度热帖数量，以及其中的看多热帖数量、看空热帖数量和中性热帖数量。参考孙鲲鹏和肖星（2018）的做法，利用下式（2）、（3）、（4）所示的方法构建个股的情绪衡量指标，分为总体情绪（TotalSent）、看涨情绪（GoodSent）和看跌情绪（BadSent）。

$$TotalSent = \ln(Good_num + Bad_num + Neutral_num) \quad (2)$$

$$GoodSent = \frac{Good_num}{Good_num + Bad_num + Neutral_num} \quad (3)$$

$$BadSent = \frac{Bad_num}{Good_num + Bad_num + Neutral_num} \quad (4)$$

式（2）、（3）、（4）中， $Good_num$ 为股吧中个股的年度看多热帖数量； Bad_num 为股吧中个股的年度看空热帖数量； $Neutral_num$ 表示股吧中个股的年度中立热帖数量。

3.1.3 控制变量

本文选取了三类控制变量：第一类为关于公司其他分析信息，包括研报数量（ReptNum）和分析师数量（AnalyNum）；第二类为公司的关键财务指标，包括每股收益（EPS）、资产收益率（ROA）、资产规模（SIZE）、资产负债率（LEV）、资产周转率（CapTov）、毛利率（GrossProfit）；第三类为衡量公司治理的指标，如审计师是否来自于四大会计师事务所（Big4）、流通股比例（LiqShr）、第一大股东持股比例（FirstHoldShr）、机构持股比例（InsHoldShr）、管理层持股比例（AdmHoldShr）。关于控制变量的定义与说明详见表 1。

表 1 控制变量定义与说明

变量名	含义	定义与说明
ReptNum	研报的数量	一年内对公司进行过跟踪分析的研报数量的自然对数
AnalyNum	分析师的数量	一年内对公司进行过跟踪分析的分析师（团队）数量的自然对数
Big4	审计师来源	虚拟变量，审计师来自境内四大会计师事务所取 1，否则取 0
EPS	每股收益	净利润本期值/最新股本
SIZE	资产规模	流通市值的自然对数值
ROA	资产收益率	净利润/平均资产总额
LEV	资产负债率	负责合计/资产总计
CapTov	资产周转率	营业收入/平均资产总额
GrossProfit	毛利率	(营业收入-营业成本)/营业收入
LiqShr	流通股比例	流通股数占总股本的比例
FirstHoldShr	第一大股东持股比例	第一大股东持股数占其总股本的比例
InsHoldShr	机构持股比例	机构投资者持有的无限售流通 A 股所占的比例
AdmHoldShr	管理层持股比例	管理层持股数占总股本的比例

3.2. 回归模型设定

用式（5）所示的模型分析投资者关注度对股价同步性的影响，使用面板数据分析中的个体固定效应方法进行回归分析，若式（5）中的系数 β_1 回归结果显著，则表明投资者的情绪关注对个股股价的同步性存在显著影响。进一步，建立如式（6）所示模型，回归分析不同情绪对个股股价同步性的影响，若式（6）中的系数 β_1 或 β_2 显著，则表明投资者的乐观情绪或悲观情绪对个股股价同步性存在显著影响。

$$PrcSyn_{it} = \beta_0 + \beta_1 TotalSent_{it} + \beta_2 Controls_{it-1} + \varepsilon_{it} \quad (5)$$

$$PrcSyn_{it} = \beta_0 + \beta_1 GoodSent_{it} + \beta_2 BadSent_{it} + \beta_3 Controls_{it-1} + \varepsilon_{it} \quad (6)$$

式(5)、(6)中, PrcSynit 表示个股 i 第 t 年的股价同步性指标值; TotalSentit、GoodSentit、BadSentit 分别表示个股 i 第 t 年的总情绪、乐观情绪和悲观情绪指标值; Controlit 表示个股 i 第 t-1 年的控制变量; ε_{it} 为服从正态分布的随机误差项。

3.3 样本选取与数据来源

本文选取的样本为沪深 300 指数和中证 500 指数的成分个股, 样本期间为 2013 年至 2018 年。市场与个股的周收益率数据, 控制变量的数据来源于国泰安 (CSMAR) 经济金融数据库, 个股的机构持股比例来源于锐思 (RESSET) 金融研究数据库。经筛选最终获得 3697 个股票-年份观测值数据, 其中沪深 300 指数成分股的观测值为 1638 个, 中证 500 指数成分股的观测值 2049 个。表 2 给出了变量的描述性统计结果。

表 2 变量描述性统计

变量	均值	标准差	最小值	中位数	最大值	观测值
PrcSyn	-0.343	1.072	-7.341	-0.223	2.192	3687
TotalSent	8.505	0.906	0.000	8.526	12.237	3687
GoodSent	0.345	0.061	0.182	0.341	1.000	3687
BadSent	0.380	0.052	0.000	0.382	0.541	3687
ReptNum	29.910	29.913	0.000	21.000	219.000	3687
AnalyNum	13.332	11.095	0.000	11.000	66.000	3687
Big4	0.137	0.347	0.000	0.000	1.000	3687
EPS	0.620	0.864	-5.573	0.473	23.091	3687
SIZE	23.843	1.145	20.321	23.777	28.544	3687
ROA	0.060	0.060	-0.275	0.049	0.669	3687
LEV	0.472	0.204	0.019	0.481	1.280	3687
CapTov	0.706	0.517	0.016	0.586	5.115	3687
GrossProfit	0.301	0.189	-0.300	0.264	1.154	3687
LiqShr	0.759	0.260	0.037	0.840	1.000	3687
FirstHoldShr	0.390	0.165	0.036	0.383	0.900	3687
InsHoldShr	0.278	0.240	0.000	0.190	0.985	3687
AdmHoldShr	0.088	0.169	0.000	0.001	0.849	3687

4. 实证结果

表 3 给出了模型(5) 和模型(6) 基于全样本数据的回归结果。

表 3 全样本回归结果

变量	被解释变量: PrcSyn				
	(1)	(2)	(3)	(4)	(5)
TotalSent	-0.085** (-2.49)				-0.273*** (-7.848)
GoodSent		-5.168*** (-11.28)		0.501 (0.74)	-0.435 (-0.64)
BadSent			7.896*** (16.01)	8.309*** (11.14)	8.768*** (11.84)
ReptNum	-0.002 (-0.87)	-0.004** (-2.11)	-0.004** (-2.18)	-0.004** (-2.14)	-0.002 (-1.28)
AnalyNum	0.011** (2.18)	0.021*** (4.15)	0.021*** (4.22)	0.021*** (4.14)	0.015*** (2.90)
Big4	-0.174	-0.143	-0.096	-0.094	-0.111

	(-1.19)	(-1.00)	(-0.68)	(-0.67)	(-0.80)
EPS	0.055	0.064	0.053	0.052	0.073*
	(1.33)	(1.59)	(1.34)	(1.31)	(1.85)
SIZE	0.501***	0.359***	0.186***	0.182***	0.275***
	(12.94)	(10.07)	(4.93)	(4.73)	(6.91)
ROA	-0.548	-0.329	0.277	0.302	0.187
	(-0.84)	(-0.53)	(0.44)	(0.48)	(0.29)
LEV	-0.507**	-0.491**	-0.259	-0.246	-0.295
	(-2.14)	(-2.12)	(-1.14)	(-1.08)	(-1.30)
CapTov	-0.125	-0.079	-0.087	-0.089	-0.117
	(-1.34)	(-0.87)	(-0.98)	(-1.00)	(-1.32)
GrossProfit	0.250	0.272	0.211	0.206	0.282
	(0.85)	(0.95)	(0.75)	(0.73)	(1.01)
LiqShr	0.208**	0.131	0.112	0.114	0.130
	(1.96)	(1.26)	(1.10)	(1.12)	(1.29)
FirstHoldShr	-0.642**	-0.392	-0.210	-0.208	-0.374
	(-2.09)	(-1.300)	(-0.71)	(-0.71)	(-1.28)
InsHoldShr	-0.732***	-0.486***	-0.582***	-0.596***	-0.636***
	(-8.34)	(-5.56)	(-6.94)	(-6.91)	(-7.43)
AdmHoldShr	0.863**	0.723**	0.867***	0.879***	0.933***
	(2.53)	(2.16)	(2.64)	(2.68)	(2.87)
R-Squared	0.083	0.120	0.155	0.155	0.173
N	3687	3687	3687	3687	3687

5. 稳健性检验

分别利用沪深 300 指数和中证 500 指数成分股作为研究样本对实证结果进行稳健性检验，实证结果如表 4 和表 5 所示。

表 4 沪深 300 指数成分股回归结果

变量	被解释变量: PrcSyn					
	(1)	(2)	(3)	(4)	(5)	(6)
TotalSent	-0.087*	-0.298***				
	(-1.65)	(-4.96)				
TotalSent*D2015		0.065***				
		(6.97)				
GoodSent			-4.503***	-3.801***		
			(-6.11)	(-5.03)		
GoodSent*D2015				0.903***		
				(3.91)		
BadSent					7.456***	6.678***
					(9.27)	(7.87)
BadSent*D2015						0.548***
						(2.73)
ReptNum	-0.005*	-0.007**	-0.007**	-0.008***	-0.008***	-0.009***
	(-1.76)	(-2.49)	(-2.54)	(-3.15)	(-3.03)	(-3.41)
AnalyNum	0.023***	0.029***	0.031***	0.035***	0.033***	0.039***
	(3.29)	(4.04)	(4.57)	(5.18)	(4.97)	(5.39)
Big4	-0.103	-0.046	-0.117	-0.086	-0.024	-0.010
	(-0.49)	(-0.23)	(-0.57)	(-0.39)	(-0.11)	(-0.05)
EPS	0.013	0.011	0.022	0.016	0.018	0.015
	(0.23)	(0.24)	(0.47)	(0.34)	(0.41)	(0.35)
SIZE	0.448***	0.541***	0.335***	0.347***	0.163***	0.189***

	(7.28)	(8.79)	(5.96)	(6.19)	(2.74)	(3.23)
ROA	-1.447	-1.523	-1.301	-1.286	-0.600	-0.651
	(-1.40)	(-1.51)	(-1.28)	(-1.28)	(-0.60)	(-0.66)
LEV	-0.773*	-0.886**	-0.781*	-0.824*	-0.565	-0.625
	(-1.73)	(-2.02)	(-1.77)	(-1.88)	(-1.31)	(-1.45)
CapTov	0.098	0.092	0.147	0.143	0.131	0.133
	(0.56)	(0.54)	(0.87)	(0.86)	(0.78)	(0.79)
GrossProfit	0.389	0.324	0.369	0.318	0.337	0.295
	(0.81)	(0.69)	(0.79)	(0.68)	(0.71)	(0.64)
LiqShr	0.524***	0.507***	0.452***	0.441***	0.469***	0.453***
	(2.99)	(2.99)	(2.79)	(2.59)	(2.69)	(2.81)
FirstHoldShr	0.111	-0.063	0.475	0.424	0.428	0.400
	(0.21)	(-0.12)	(0.91)	(0.82)	(0.85)	(0.79)
InsHoldShr	-0.850***	-0.653***	-0.652***	-0.547***	-0.752***	-0.648***
	(-6.19)	(-4.76)	(-4.75)	(-3.86)	(-5.58)	(-4.81)
AdmHoldShr	1.207*	1.278*	0.925	0.881	0.957	0.904
	(1.70)	(1.85)	(1.33)	(1.27)	(1.41)	(1.33)
R-Squared	0.093	0.135	0.123	0.135	0.163	0.169
N	1244	1244	1244	1244	1244	1244

表 5 中证 500 指数成分股回归结果

变量	被解释变量: PrcSyn					
	(1)	(2)	(3)	(4)	(5)	(6)
TotalSent	-0.097**	-0.371***				
	(-2.16)	(-7.57)				
TotalSent*D2015		0.085***				
		(11.81)				
GoodSent			-5.533***	-4.585***		
			(-9.50)	(-7.81)		
GoodSent*D2015				1.299***		
				(7.53)		
BadSent					8.094***	6.905***
					(12.98)	(10.67)
BadSent*D2015						0.838***
						(6.05)
ReptNum	0.001	0.001	-0.002	-0.003	-0.000	-0.002
	(0.48)	(0.18)	(-0.56)	(-0.99)	(-0.12)	(-0.59)
AnalyNum	-0.001	0.001	0.011	0.015**	0.008	0.012*
	(-0.09)	(0.20)	(1.47)	(2.01)	(1.11)	(1.67)
Big4	-0.204	-0.247	-0.143	-0.155	-0.149	-0.159
	(-1.01)	(-1.26)	(-0.72)	(-0.79)	(-0.76)	(-0.82)
EPS	0.183**	0.152*	0.180**	0.145*	0.146*	0.120
	(2.27)	(1.95)	(2.29)	(1.86)	(1.89)	(1.57)
SIZE	0.559***	0.684***	0.386***	0.405***	0.218***	0.259***
	(11.06)	(13.69)	(8.29)	(8.80)	(4.42)	(5.24)
ROA	-0.635	-0.878	-0.354	-0.413	0.314	0.179
	(-0.70)	(-1.00)	(-0.40)	(-0.47)	(0.36)	(0.21)
LEV	-0.391	-0.448	-0.348	-0.336	-0.141	-0.167
	(-1.38)	(-1.64)	(-1.26)	(-1.23)	(-0.52)	(-0.62)
CapTov	-0.182	-0.239**	-0.143	-0.155	-0.143	-0.155
	(-1.64)	(-2.23)	(-1.32)	(-1.46)	(-1.35)	(-1.48)
GrossProfit	0.125	0.220	0.163	0.153	0.046	0.072
	(0.33)	(0.60)	(0.44)	(0.42)	(0.13)	(0.20)
LiqShr	0.017	-0.042	-0.060	-0.104	-0.097	-0.118
	(0.13)	(-0.32)	(-0.45)	(-0.80)	(-0.75)	(-0.92)
FirstHoldShr	-0.969**	-1.235***	-0.784**	-0.820**	-0.472	-0.562
	(-2.54)	(-3.35)	(-2.11)	(-2.24)	(-1.29)	(-1.55)
InsHoldShr	-0.636***	-0.367***	-0.357***	-0.198*	-0.463***	-0.324***
	(-5.55)	(-3.24)	(-3.12)	(-1.73)	(-4.21)	(-2.92)

AdmHoldShr	0.680*	0.473**	0.577	0.485	0.767**	0.662*
	(1.70)	(1.22)	(1.47)	(1.25)	(1.99)	(1.74)
R-Squared	0.094	0.154	0.132	0.156	0.164	0.179
N	2443	2443	2443	2443	2443	2443

通过表4和表5，可以清晰地看到模型中的指标除少数控制变量之外，均满足稳定性检验，尤其是其中最重要的几项解释变量如 TotalSent、TotalSent*D2015、GoodSent、GoodSent*D2015、BadSent、BadSent*D2015的稳定性检验结果表明投资者自媒体交流会对其他投资者情绪产生影响，这种股吧中的情绪会对二级市场产生显著的影响。

6.结论

本文采用最小二乘法，构建回归模型实证分析了在沪深 300 板块的蓝筹个股和中证 500 板块的中小板个股中，（1）投资者关注度提升均可以显著降低个股的股价同步性；（2）投资者的乐观情绪对降低股价同步性的作用更明显；（3）投资者关注度降低股价同步性的作用在中小板个股中更显著，这可能是由于中小板中非理性的散户投资者比例更多的原因。

中国股票市场的股价同步性现象在世界主要经济体中是名列前茅的，同涨同跌现象严重，难以体现出个股上市公司的基本面的差异，这反映了我国资本市场的定价效率偏低的现状以及投资者习惯于盲目跟风炒作的特性。投资者通过各种自媒体平台的交流、研究可以大大增强对投资标的基本面的了解，提高对其理性投资、价值投资的认知，因此可以降低个股的同步性并提高资本市场的定价效率，对提升我国资本市场的定价效率并引导真正的价值性投资所有助益。

References

- [1] Aboody D, Lehavy R, Trueman B, Limited Attention and the Earnings Announcement Returns of Past Stock Market Winners, Social Science Electronic Publishing, 2010, pp. 317-344.
- [2] Dimpfl T, Jank S, Can Internet Search Queries Help to Predict Stock Market Volatility?, European Financial Management, 2016, pp.171-192.
- [3] Wei-Qi L, Xin-Xin L, Individual /institutional investor sentiment and stock returns: Study based on Shanghai A-share market, Journal of Management Sciences in China, 2014, pp.30-36.
- [4] Qiang Z, Shu-E Y, Noise Trading, Investor Sentiment Volatility, and Stock Returns, Systems Engineering - Theory and Practice, 2009, pp. 40-47.
- [5] Berger A N, Hasan I, Zhou M. Bank Ownership and Efficiency in China: What Will Happen in the World's Largest Nation?. Journal of Banking & Finance, 2007, pp. 113-130.
- [6] Allen F, Santomero A M, What Do Financial Intermediaries Do?, Center for Financial Institutions Working Papers, 1999, pp. 271-294.