

社交媒体效应、投资者认知和股票横截面收益

——“学海拾珠”系列之五十一

主要观点：

分析师：严佳炜

执业证书号：S0010520070001

邮箱：yanjw@hazq.com

分析师：朱定豪

执业证书号：S0010520060003

邮箱：zhudh@hazq.com

联系人：吴正宇

执业证书号：S0010120080052

邮箱：wuzy@hazq.com

本篇是“学海拾珠”系列第五十一篇，本期推荐的海外文献研究了社交媒体效应、投资者认知和股票横截面收益之间的关系。研究发现在著名的 Fama-French 因子模型中加入社交媒体因子显著提高了其对股票横截面收益的解释力。当控制大众媒体效应和流动性效应时，其结果是稳健的。此外，以社交媒体效应构建的投资组合能够产生显著的收益，投资组合经理和从业者应该在他们的投资组合和风险管理中考虑社交媒体溢价。

回到 A 股市场，在目前 alpha 日益稀缺的背景下，如何结合另类数据获得额外的信息增量，对传统的多因子选股模型进行补充是非常有价值的。我们可以通过爬虫获取股吧的数据，并基于股票分论坛的热度构建因子或个股择时信号构建相应的策略。

● 股吧发帖数量是衡量投资者认知度的良好的代理变量

股吧发帖数量衡量指标与传统的投资者认知度衡量指标，如股东数量和增长、大众媒体报道和广告成本等，存在正相关关系。而且股吧发帖数量具有明显的优势：第一，媒体报道和广告投资是单向信息传播，而股吧发帖数量则是直接衡量指标，因为投资者在股吧发帖前肯定对一只股票有一定的认知度。其次，股票的股东数忽略了认知该股票并可能在未来将其添加到其投资组合中的潜在投资者，而股吧发帖数量反映了当前和未来投资者的股票认知度。第三，股吧发帖数量的历史数据和实时数据都可以轻松获取，不受财报的季度披露限制。

● 社交媒体因子能够带来显著的风险调整后回报

作者通过构建基于社交媒体溢价的社交媒体因子进一步识别与投资者认知相关的风险，发现社交媒体因子提供了著名的 Fama-French 因子与 Fang 和 Peress(2009)的大众媒体因子未涵盖的新信息。由社交媒体因子增强的 Fama-French 因素模型能够产生显著的风险调整后回报。风险定价测试发现，对于每单位风险，社交媒体效应可以带来市场投资组合收益 2% 的变化。GRS 统计数据还表明，社交媒体因素增强了现有因子模型的有效性。

● 风险提示

作者结论基于历史数据与海外文献进行总结；不构成任何投资建议。

相关报告

1. 《投资者评价基金时会考虑哪些因素？——“学海拾珠”之五十》
2. 《公司盈利季节性和股票收益——“学海拾珠”之四十九》
3. 《信息消化与资产定价——“学海拾珠”系列之四十八》
4. 《日历更替：研究盈余公告发布时点影响的新视角——“学海拾珠”系列之四十七》
5. 《收益的季节性是由于风险还是错误定价？——“学海拾珠”系列之四十六》
6. 《公司复杂性对盈余惯性的影响——“学海拾珠”系列之四十五》
7. 《如何衡量基金经理把握股票基本面的能力？——“学海拾珠”系列之四十四》
8. 《企业预期管理与股票收益——“学海拾珠”系列之四十三》

正文目录

1 简介	4
2 实证研究：社交媒体效应和股票收益	5
2.1 数据	5
2.2 实证结果	6
3 社交媒体溢价的经济解释	10
4 社交媒体效应溢价	14
5 结论	17
风险提示:	18

图表目录

图表 1 2012.1-2017.6 社交媒体和大众媒体报道汇总统计	6
图表 2 (3 × 3) VW 投资组合的每月超额收益	7
图表 3 风险因素的汇总统计	8
图表 4 SMF 的因子冗余测试	8
图表 5 FAMA-MACBETH 回归的因子风险溢价	9
图表 6 因子模型的 GRS 检验	10
图表 7 投资者认知和股吧发帖数量	11
图表 8 按股东数和股吧发帖数量分组的股东数回归	12
图表 9 25 个大众投资组合在不同情绪组中的每月超额收益	13
图表 10 针对情感因素的 SMF 的因素冗余测试	13
图表 11 股吧发帖数量投资组合的月度回报	14
图表 12 按公司特征分组的组合回报	15
图表 13 按流动性和情绪特征分组的交易利润	16
图表 14 策略的累积回报	16
图表 15 社交媒体效应策略年度表现	17

1 简介

媒体报道是股票横截面收益的重要决定因素。在大众传媒时代，报纸、电视等媒体传播的信息影响着投资者的交易活动和股票价格。Fang 和 Peress(2009)记录了大众媒体报道对股票横截面收益的影响，并提出了“大众媒体效应”：**股票被大众媒体报道率越低，收益率越高**。这种效应已在多个国家得到论证(Ferguson et al., 2015; Aman et al., 2018; Griffin et al., 2011; Turner et al., 2018; Zou, Cao and Wang, 2018)，Fang 和 Peress(2009)等通过默顿(1987)的“投资者认知假说”来解释大众传媒效应。根据这一论点，**大众媒体报道扩大了投资者的认知程度。大众媒体报道较少的股票需要提供更高的回报，以补偿持有不完全分散化投资组合的投资者。**

随着互联网的出现，社交媒体已成为投资者发布、收集和交换信息的主流平台(Allcott and Gentzkow, 2017; Lazer et al., 2018; Lee et al., 2016)。因此，研究社交媒体对股票横截面收益的影响非常重要。与大众媒体相比，社交媒体可以更好地代表投资者对股票的认知度。首先，**大众媒体只能单向传播信息**，很难知道投资者是否收到了信息，有多少投资者了解这些证券，以及认知度如何。但是，当投资者在社交媒体平台上发布有关一只股票的消息时，通常对这只股票已经有所了解，这更容易量化他们对这只股票的认知度。其次，**社交媒体促进了投资者之间的互动**，从而减少了信息不对称性(Miller and Skinner, 2015)。通过投资者互动和对在线消息的回应，那些最初对股票了解有限的人也会产生一定的认知。最后，**社交媒体平台可以记录投资者的互动和对在线消息的回应**。这可以作为投资者认知度的直接衡量标准，而大众媒体并不容易获得这些信息。

作者利用中国最大、最活跃的金融社交媒体平台东方财富股吧（以下简称股吧）发布的在线消息数据，研究了社交媒体对中国股票横截面收益的影响。股吧平台兼容股吧交易 APP，被中国投资者，尤其是个人投资者广泛使用。股吧的影响力广泛，流量占中国所有金融网站的 5% 左右，每天影响着 1200 万至 2200 万投资者。作者收集了近 1 亿投资者在平台上发布的涵盖 3000 多只股票的数据。对于每一只股票，在股吧都有一个特定的子论坛。因此，可以计算出**每只股票论坛里的每月发帖数，作为投资者对该股票认知度的直接衡量标准**。通过将股票分类为不同社交媒体覆盖程度的投资组合，可以研究社交媒体溢价。如果社交媒体向更广泛的受众传播信息并扩大投资者的认知度，预计会看到显著的社交媒体溢价。然后作者研究了社交媒体溢价对股票横截面收益率的影响。通过测试社交媒体溢价的信息量，作者确定了在解释股票横截面收益率时与投资者认知相关的风险，并直接检验了默顿(1987)投资者认知假说。

中国股市仍在发展中，但中国在将社交媒体应用于金融市场方面取得了长足的进步。根据麦肯锡的一项调查，91% 的中国受访者最近访问了社交媒体网站。83.6% 的个人投资者将基于互联网的社交媒体作为首选的投资信息来源。广泛使用股吧等社交媒体平台提供了庞大的在线帖子数据集，这是衡量投资者认知度的理想选择。而且中国股市以个人投资者为主(Han and Li, 2017)，个人交易量占总交易量的 85% 以上。社交媒体主要影响个人投资者而非机构投资者的认知，这有助于直接检验默顿的投资者认知假说。

作者报告了中国股市的显著的社交媒体溢价。基于此溢价，作者通过构建社交媒体因子进一步识别与投资者认知相关的风险。研究发现在著名的 Fama-French 因子模型中加入社交媒体因子显著提高了其对股票横截面收益率的解释力。当控制大众媒体效应和流动性效应时，其结果是稳健的。而且通过使用社交媒体数据来衡

量投资者认知度，提供了支持默顿投资者认知假说的直接证据。最后发现基于社交媒体因子的投资策略能够在风险调整的基础上产生可观的收益。作者说明了使用社交媒体数据来了解股票横截面收益率的重要性，这为未来包括金融其他领域在内的研究指明了方向。

虽然作者将在线发帖数视为投资者认知度的衡量标准，但它也常常与投资者情绪有关。Bollen et al. (2011) 将社交媒体发帖解释为衡量投资者情绪的指标。他们从 Twitter 中提取每日公众情绪数据，发现公众情绪与市场回报具有显著相关性。在将社交媒体效应与股票横截面收益率联系起来的研究中，Kim and Kim (2014) 基于 Yahoo 上 2005 年至 2010 年留言板上的 3200 万条消息测试了提取的投资者情绪对股票回报的预测能力。Leung and Ton (2015) 使用澳大利亚金融社交媒体平台 HotCopper 2003 年至 2008 年的 250 万个观察样本数据，发现在业绩不佳和小盘股样本中，股票收益仅与社交媒体上表达的情绪呈正相关。以上的研究主要关注在线帖子的内容而不是在线帖子的数量。为了进一步控制可能的情绪效应，作者采用了常见的情绪衡量指标，例如 MAX 效应 (Baliet al., 2011; Fong and Toh, 2014)、成交率和交易量 (Hong and Stein, 2007)。研究发现在控制情绪效应后，主要的结果保持不变。这支持作者将在线帖子数量解释为投资者认知度的衡量标准，并进一步支持作者的论点。

2 实证研究：社交媒体效应和股票收益

2.1 数据

作者从中国最大、最全面的金融社交媒体平台股吧上收集社交媒体数据，该平台的网络流量是中国金融社交媒体中最高的。股吧成立于 2006 年，旨在为投资者提供一个分享信息和观点的金融信息平台。目前股吧拥有数百万注册用户，每天发布超过 100000 个帖子，规模比其最接近的竞争对手大 10 倍左右。股吧包括以当前交易或已在市场交易所有股票的代码或名称命名的子论坛。在股吧上发帖，投资者至少需要知道这只股票的名称和代码，这保证了投资者在发帖前对这只股票有一定的认知度。作者从所有股吧子论坛检索可见内容，并使用大约 1.5 亿个帖子。需要注意的是，因为用户是主动进入子论坛的，作者认为子论坛的所有帖子都与该股票相关联。作者设定每只股票的关键变量**股吧发帖数量**是其论坛中所有帖子的**月度数量**，将作为投资者认知的衡量标准。

作者的样本包括 2012 年 1 月 1 日至 2017 年 6 月 30 日中国 A 股市场的所有上市公司。在样本期间，社交媒体成为主流，为作者的研究提供了足够的数据。剔除上市不足一年、账面价值为负、长期停牌或退市以及特殊处理的股票，以确保股票收益正确反映公司价值的变化。鉴于这些异常情况下的股票很少，我们样本中的股票平均数量约为 2500 只，涵盖了市场上大部分上市的股票，占各时期总市值的 90% 以上。因此，作者样本中的公司具有代表性。上市公司的会计和交易数据来自 Wind 数据库。

为了比较社交媒体与大众媒体的影响，作者按照 Fang 和 Peress (2009) 的方法构建了一个媒体报道曝光变量。作者从 RESSET 数据库的“每日新闻”部分收集数。“每日新闻”栏目主要从上海证券报、证券时报、中国经济网、新浪财经、腾讯财经等知名媒体收集特定个股的公开新闻。

表 1 展示了**股吧发帖数量**和**媒体报道**的汇总统计数据。股吧每个股票分论坛平均每月发布 700 个帖子。2012 年以来，社交媒体和移动网络在中国变得越来越流

行，股吧发帖数量在此期间得到了快速发展。股吧发帖数量在 2015 年达到顶峰，当时中国股市的繁荣吸引了大量投资者，导致社交媒体帖子激增。与社交媒体帖子相比，每只股票的平均新闻报道数量很少（每月约 7.8 篇）。与 Fang 和 Peress(2009)类似，在给定的一个月內，大约 20%的股票没有被大众媒体报道。相比之下，在作者的样本中，每只股票每个月都有社交媒体报道。

图表 1 2012.1-2017.6 社交媒体和大众媒体报道汇总统计

Period	Mean	Std	Min	P5	P25	Median	P75	P95	Max
<i>Panel A: Gubapost</i>									
2012.1~2012.6	492.8	843.8	11	61	154	293	551	1459	22867
2012.7~2012.12	370.5	592.2	10	49	115	216	410	1138	17655
2013.1~2013.6	439.9	677.5	7	60	144	260	487	1276	14440
2013.7~2013.12	598.7	909.1	13	87	203	363	660	1825	34041
2014.1~2014.6	426.2	547.8	12	64	149	270	490	1296	10653
2014.7~2014.12	753.4	1018.2	1	96	249	464	876	2277	30154
2015.1~2015.6	1289.6	2403.2	1	129	343	694	1410	4156	135392
2015.7~2015.12	1002.0	1385.9	1	136	327	618	1150	3052	37582
2016.1~2016.6	738.9	859.8	1	124	289	494	875	2087	18907
2016.7~2016.12	737.9	1278.1	1	110	271	480	842	2094	89106
2017.1~2017.6	623.3	801.6	1	98	229	400	705	1844	14676
overall	700.7	1186.7	1	82	214	406	784	2175	135392
<i>Panel B: News</i>									
2012.1~2012.6	3.6	6.8	0	0	0	2	4	14	110
2012.7~2012.12	3.6	7.4	0	0	0	1	4	14	195
2013.1~2013.6	4.1	8.3	0	0	0	2	5	15	190
2013.7~2013.12	6.6	12.4	0	0	1	3	8	23	517
2014.1~2014.6	5.9	11.2	0	0	1	3	7	20	215
2014.7~2014.12	6.8	12.0	0	0	1	3	8	23	200
2015.1~2015.6	6.7	12.3	0	0	1	3	8	23	359
2015.7~2015.12	10.9	21.5	0	0	2	5	12	40	603
2016.1~2016.6	14.2	18.0	0	1	4	9	18	46	598
2016.7~2016.12	12.4	25.0	0	0	3	7	14	41	1395
2017.1~2017.6	4.1	8.4	0	0	1	2	5	15	193
overall	7.8	15.8	0	0	1	3	9	29	1395

资料来源：华安证券研究所整理

此外，作者还测试了股吧发帖数量与公司特征之间的相关性。结果表明，股吧发帖数量与新闻报道（系数为 46%）高度相关，说明股吧发帖数量与投资者认知度有关。规模大、波动性大、回报高的股票会吸引更多的股吧发帖数量可能是合理的，因为研究发现相关系数分别为 23%、22%和 11%。股吧发帖数量也与分析师覆盖率相关，系数为 9%。

2.2 实证结果

本节研究了社交媒体对股票横截面收益率的影响。首先研究按公司特征和社交媒体报道进行双重排序的投资组合的平均超额收益，然后研究这些投资组合的风险调整后的收益。

作者使用非参数法研究来检验股吧发帖数量的影响，包括规模、B/M、流动性、波动性、价格、机构持有率、营业额、交易量、上个月的最大日回报率 R_{max} 、上个月的月回报率 R_{t-1} 、毛利率率和分析师覆盖率。根据 Fang 和 Peress(2009)的方法，作者首先根据这些公司特征对股票进行双重排序，然后根据股吧发帖数量对每个组中的股票进行双重排序。表 2 展示了按公司特征和股吧发帖数量双重排序的 3×3 市值加权(VW)投资组合的平均 1 个月前相对无风险利率的超额回报。研究发现随着股吧发帖数量的增加，平均回报率下降。“LH”列展示了多空投资组合的平均回报（在每个规模的三分位数中做多最低数发布的投资组合和做空最高发布数的投资组合），在低的三分位数中每月收益高达 2.07%。结果表明**社交媒体效应存在于除了顶部五分之一以外的不同规模的公司中**。在其他公司特征中同样可以观察到类似的社交媒体效应规律。**收益不显著的多空组合是指公司规模大、B/M 高、流动性高、波动性低、换手率低、情绪高的股票组合**。这些股票可能有更多的分析师报道、更多的大众媒体报道和更少的信息不对称，从而获得更多的投资者认知。因此，社交媒体对此类股票的影响有限。同样，高回报率吸引投资者关注股票信息，

这也增加了投资者的认知度。因此，社交媒体在提供知名度较低的股票信息方面起着不可或缺的作用，社交媒体效应对这些股票的溢价明显更高

图表 2 (3 × 3) VW 投资组合的每月超额收益

Factors	Gubapost				Factors	Gubapost			
	Low	Medium	High	L-H		Low	Medium	High	L-H
<i>Size × Gubapost portfolio</i>					<i>B/M × Gubapost portfolio</i>				
Small	0.0361	0.0277	0.0154	0.0207***	Low	0.0214	0.0130	0.0015	0.0199***
Medium	0.0260	0.0178	0.0085	0.0175***	Medium	0.0251	0.0142	0.0070	0.0182***
Big	0.0147	0.0141	0.0059	0.0087	High	0.0211	0.0169	0.0085	0.0126*
<i>Liquidity × Gubapost portfolio</i>					<i>Volatility × Gubapost portfolio</i>				
Low	0.0289	0.0278	0.0196	0.0093***	Low	0.0218	0.0183	0.0096	0.0121*
Medium	0.0197	0.0170	0.0135	0.0063**	Medium	0.0236	0.0137	0.0072	0.0164**
High	0.0146	0.0112	0.0046	0.0100*	High	0.0241	0.0091	-0.0034	0.0274***
<i>Price × Gubapost portfolio</i>					<i>Institution holding × Gubapost portfolio</i>				
Low	0.0230	0.0192	0.0077	0.0152**	Less	0.0267	0.0177	0.0094	0.0173**
Medium	0.0232	0.0136	0.0082	0.0150**	Medium	0.0207	0.0163	0.0106	0.0101
High	0.0219	0.0126	0.0036	0.0183**	More	0.0214	0.0138	0.0067	0.0147**
<i>Turnover × Gubapost portfolio</i>					<i>Dollar trading volume × Gubapost portfolio</i>				
Low	0.0204	0.0157	0.0099	0.0105	Low	0.0268	0.0238	0.0196	0.0071**
Medium	0.0244	0.0165	0.0085	0.0159*	Medium	0.0179	0.0170	0.0092	0.0087***
High	0.0247	0.0114	0.0014	0.0233***	High	0.0160	0.0102	0.0040	0.0120**
<i>R_{max} × Gubapost portfolio</i>					<i>R_{t-1} × Gubapost portfolio</i>				
Low	0.0236	0.0165	0.0085	0.0151**	Low	0.0249	0.0236	0.0100	0.0148**
Medium	0.0244	0.0148	0.0099	0.0144*	Medium	0.0246	0.0146	0.0082	0.0165**
High	0.0193	0.0132	0.0072	0.0122	High	0.0166	0.0114	0.0041	0.0125
<i>Gross profitability × Gubapost portfolio</i>					<i>Analyst coverage × Gubapost portfolio</i>				
Low	0.0212	0.0121	0.0072	0.0139**	Low	0.0265	0.0162	0.0063	0.0202***
Medium	0.0221	0.0127	0.0085	0.0136*	Medium	0.0221	0.0146	0.0073	0.0148**
High	0.0225	0.0157	0.0058	0.0167**	High	0.0197	0.0146	0.0075	0.0122

资料来源：华安证券研究所整理

接下来，作者通过构建社交媒体因子(SMF)并检验社交媒体对股票横截面收益率的影响来确定与投资者认知相关的风险。与 Fang 和 Peress(2009)类似，作者根据股吧发帖数量将所有样本股票分为三组，计算股吧发帖数量股票投资组合的较低 1/3 减去 1 个月前较高股票投资组合的市值加权回报 1/3 作为 SMF 的值。每个月计算多空投资组合的回报，得到一个时间序列，它代表了社交媒体效应的风险因子。为了比较，作者还根据 Fang 和 Peress(2009)构建了一个大众媒体因子(MMF)，并为 Fama 和 French(1993)的三因子模型、Carhart(1997)的四因子模型、Fama 和 French(2015)的五因子模型以及 Pastor 和 Stambaugh(2003)的流动性因子(LIQ)构建了因子。表 17 列出了所有因子及其定义。表 3 展示了这些因子的汇总统计数据。在样本期内，SMF 占每月风险溢价 1.54%，高于规模 (1.45%)、增长 (0.21%) 等因素的溢价。表 3 中的 B 部分展示了 SMF 与其他因子之间的相关系数矩阵，表明 SMF 与许多已知因素 (如 SMB 和 HML) 高度相关。因此，非常重要的是检验社交媒体效应是否具有显著的风险调整回报。

图表 3 风险因素的汇总统计

Panel A: Summary statistics									
	SMF	MMF	Mkt	SMB	HML	UMD	RMW	CMA	LIQ
Mean	0.0154	0.0136	0.0077	0.0145	0.0021	0.0016	-0.0058	0.0045	0.0060
Std dev.	0.0636	0.0579	0.0738	0.0544	0.0454	0.0505	0.0312	0.0195	0.0258
Median	0.0135	0.0132	0.0097	0.0138	0.0025	0.0081	-0.0073	0.0036	0.0026
Min	-0.3039	-0.2513	-0.2511	-0.1999	-0.1648	-0.1974	-0.0868	-0.0378	-0.0393
Max	0.2905	0.2274	0.1670	0.2015	0.1915	0.1463	0.0914	0.0636	0.0880
Skewness	-0.6760	-0.7222	-0.4337	-0.2451	0.3486	-0.5066	0.2782	0.5100	1.1003
Kurtosis	15.9914	10.4936	4.7582	6.9479	8.7261	6.0287	4.1680	3.9467	4.9241
Panel B: Correlation matrix									
	SMF	MMF	Mkt	SMB	HML	UMD	RMW	CMA	LIQ
SMF	1.0000								
MMF	0.8759	1.0000							
Mkt	0.1403	0.1627	1.0000						
SMB	0.8652	0.8924	0.3113	1.0000					
HML	-0.8925	-0.7842	-0.1851	-0.8144	1.0000				
UMD	0.3665	0.1717	-0.0373	0.3209	-0.3120	1.0000			
RMW	-0.6657	-0.7686	-0.4047	-0.8659	0.5886	-0.3214	1.0000		
CMA	-0.0562	0.1768	0.1721	0.1991	0.2180	0.0693	-0.4606	1.0000	
LIQ	-0.3946	-0.2223	0.1570	-0.1423	0.3470	-0.2721	-0.0233	0.3077	1.0000

资料来源：华安证券研究所整理

为了研究新的 SMF 是否确实为之前构建的因子增加了解释力，作者按照 Fama 和 French (2015) 以及 Fang 和 Peress (2009) 通过对其他因子回归 SMF 进行了因子冗余测试。如果 SMF 提供了新信息，则回归的截距将是显著的。表 4 在第(1)中分别展示了 Fama-French 三因子、Carhart 四因子和 Fama-French 五因子的回归结果；作者将第(2)列中的流动性因子和第(3)列中的大众媒体因子相加。作者从 Fama-French 的三个因子的回归中看到它们与 SMF 相关，更重要的是截距显著为正，系数为 0.01。对于含有和不包含流动性因子的四因子和五因子模型，作者观察到相同的结果。在控制市场、规模、账面市值比、动量、盈利能力和投资因素后，显著的正截距仍然存在，几乎没有变化。这些结果表明，**股吧发帖数量衡量的风险无法用 Fama-French 和 Carhart 模型中的因子来解释**。此外，在列 (2) 中控制流动性风险后截距仍然显著，这表明社交媒体因子不能用流动性风险来解释。当加上大众媒体因子时，看到它有一个显著的正系数；截距从 0.010 下降到 0.009，但仍然显著。总之，作者发现 SMB、HML、LIQ 和 MMF 与 SMF 显著相关，因此 SMF 提供了超出现有已知因子的新信息并衡量了新型风险。

图表 4 SMF 的因子冗余测试

	(1)			(2)			(3)		
Mkt	-0.088** (-2.032)	-0.081* (-1.860)	-0.085* (-1.861)	-0.064 (-1.564)	-0.061 (-1.492)	-0.064 (-1.501)	-0.035 (-0.907)	-0.020 (-0.521)	-0.031 (-0.758)
SMB	0.537*** (5.395)	0.520*** (5.167)	0.588*** (3.407)	0.600*** (6.330)	0.590*** (6.067)	0.599*** (3.712)	0.305** (2.363)	0.197 (1.427)	0.360** (2.158)
HML	-0.753*** (-6.521)	-0.744*** (-6.430)	-0.715*** (-4.559)	-0.608*** (-5.208)	-0.608*** (-5.181)	-0.600*** (-3.998)	-0.558*** (-5.072)	-0.548*** (-5.088)	-0.482*** (-3.348)
UMD		0.069 (1.077)			0.031 (0.507)			0.115* (1.915)	
RMW			0.045 (0.179)			-0.016 (-0.056)			0.032 (0.156)
CMA			-0.059 (-0.243)			-0.023 (-0.079)			-0.145 (-0.654)
LIQ				-0.392*** (-3.220)	-0.379*** (-3.029)	-0.392*** (-3.136)	-0.356*** (-3.119)	-0.299** (-2.593)	-0.342*** (-2.925)
MMF							0.336*** (3.151)	0.418*** (3.706)	0.359*** (3.221)
Intercept	0.010*** (2.919)	0.010*** (2.943)	0.010*** (2.689)	0.011*** (3.415)	0.011*** (3.394)	0.011*** (3.240)	0.010*** (3.370)	0.010*** (3.375)	0.009*** (3.019)
R ²	0.863	0.865	0.863	0.883	0.883	0.883	0.900	0.906	0.901

资料来源：华安证券研究所整理

接下来研究社交媒体效应的经济意义。作者计算风险使用 Fama-Macbeth 两步回归的 SMF 溢价。作者使用 25 个 Size-股吧发帖数量投资组合作为左侧(LHS)投资组合，并用因子模型的时间序列回归来估计每个因子的 beta。然后，在每个时期对估计的 beta 进行投资组合回报的横截面回归，以获得风险溢价。

图表 5 Fama-Macbeth 回归的因子风险溢价

	(1)			(2)			(3)		
	FF 3	Carhart 4	FF 5	FF 3	Carhart 4	FF 5	FF 3	Carhart 4Variable	FF 5
SMF				0.020*** (4.049)	0.020*** (4.379)	0.019*** (4.173)			
MMF							0.030* (1.778)	0.030* (1.773)	0.022 (1.528)
Mkt	-0.028 (-1.475)	-0.032 (-1.556)	0.001 (0.054)	0.025 (1.157)	0.021 (1.054)	0.033 (1.673)	-0.023 (-1.140)	-0.026 (-1.180)	0.004 (0.219)
SMB	0.013*** (3.898)	0.013*** (3.871)	0.012*** (4.369)	0.018*** (5.975)	0.019*** (6.601)	0.016*** (5.806)	0.013*** (3.922)	0.013*** (3.859)	0.012*** (4.359)
HML	-0.007 (-0.962)	-0.008 (-0.984)	-0.011 (-1.607)	0.014 (1.664)	0.017* (2.028)	0.007 (0.830)	-0.002 (-0.194)	-0.003 (-0.256)	-0.007 (-0.765)
UMD		-0.005 (-0.229)			-0.024 (-1.469)			-0.012 (-0.531)	
RMW			0.01 (1.663)			-0.001 (-0.212)			0.008 (1.106)
CMA			0.002 (0.326)			0.011 (1.615)			0.006 (0.693)
Intercept	0.035* (1.750)	0.039* (1.810)	0.006 (0.331)	-0.018 (-0.835)	-0.016 (-0.782)	-0.026 (-1.290)	0.029 (1.352)	0.032 (1.376)	0.002 (0.121)
R ²	0.490	0.498	0.662	0.688	0.737	0.765	0.509	0.513	0.673

资料来源：华安证券研究所整理

表 5 显示了 Fama-Macbeth 回归结果。列(1)展示了 Fama-French 三因子模型、Carhart 四因子模型和 Fama-French 五因子模型的结果。可以看到三因子模型的 R² 范围从 49%到 66%。使用 SMF 时，R² 增加 10-20%，表明 SMF 有效地解释了投资组合回报。还注意到，SMF 的风险溢价为 0.02，在 1%的水平上显著，年化回报率约为 26.8%。值得注意的是 Fang 和 Peress(2009)的大众媒体效应的 R² 改进不是与 SMF 一样高。结果表明，社交媒体效应具有经济意义，因为它产生了单位风险的高溢价。

研究表明，社交媒体效应对资产定价的影响在经济上和统计上都是显著的。然后通过比较使用和不使用 SMF 的因子模型的性能来进一步评估影响。作者使用 Gibbonsetal(1989)的 GRS 统计量。GRS 检验的本质是确定投资组合收益对因子收益回归的截距是否显著，这也是资产定价模型的关键点。GRS 检验的原假设是回归的截距与零无法区分。为了充分捕捉投资组合的不同特征，作者根据 5×5 Size-B/M、5×5 Size-股吧发帖数量和 2×4×4Size-B/M-股吧发帖数量构建了三种 LHS 投资组合。作者使用四个指标来衡量模型的性能：第一个指标 p(GRS)是 GRS 统计量的 p 值；第二个指标是 A| α |，即截距的平均绝对值，这是未由风险因素解释的超额收益 (FamaandFrench,2015)；第三个衡量指标是总平均绝对定价误差 (MAPE)，由 $|\alpha| + \frac{1}{N} |\epsilon|$ 衡量 Adrianetal.(2014)；最后一个衡量指标 adj.R² 是时间序列回归的平均调整后的 R²，同时可以衡量定价能力。显然，较大的 p(GRS)和 adj.R² 值以及较小的 A| α |和 MAPE 值表示更好的模型性能。表 6 提供了四个指标的统计数据。

图表 6 因子模型的 GRS 检验

	(1)				(2)			
	$p(GRS)$	$A \alpha_i $	MAPE	adj. R^2	$p(GRS)$	$A \alpha_i $	MAPE	adj. R^2
<i>25 Size-B/M portfolios</i>								
FF 3	0.0127	0.0027	0.0173	0.9597	0.0351	0.0029	0.0171	0.9610
Carhart 4	0.0099	0.0027	0.0169	0.9614	0.0293	0.0029	0.0168	0.9628
FF 5	0.0125	0.0028	0.0172	0.9598	0.0353	0.0027	0.0168	0.9612
<i>25 Size-Gubapost portfolios</i>								
FF 3	0.0060	0.0061	0.0222	0.9427	0.0272	0.0043	0.0191	0.9525
Carhart 4	0.0032	0.0061	0.0221	0.9436	0.0180	0.0042	0.0189	0.9537
FF 5	0.0150	0.0057	0.0217	0.9431	0.0485	0.0040	0.0186	0.9532
<i>32 Size-B/M-Gubapost portfolios</i>								
FF 3	0.0714	0.0058	0.0237	0.9649	0.2069	0.0045	0.0213	0.9684
Carhart 4	0.0818	0.0058	0.0233	0.9651	0.2234	0.0046	0.0208	0.9682
FF 5	0.1429	0.0056	0.0233	0.9641	0.3227	0.0044	0.0209	0.9678

资料来源：华安证券研究所整理

在表 6 中，第(1)列展示了 Fama-French 三因子模型、Carhart 四因子模型和 Fama-French 五因子模型的统计数据，而第(2)列显示由 SMF 增强的相同模型的统计数据。该表显示含有 SMF 的模型性能更好。例如，在 Size-股吧发帖数量投资组合中，将 SMF 添加到 Fama-French 三因子模型会将 GRS 统计量的 p 值从 0.0060 变为 0.0272。结果表明，SMF 提高了资产定价模型的性能，并对横截面回报具有定价能力。即使在独立于股吧发帖数量的 Size-B/M 组中，SMF 也能在一定程度上提高模型性能。

总而言之，作者证明了社交媒体效应存在于不同的公司特征中；社交媒体因子提供了传统风险因子无法传达的新信息，并产生显著的正风险溢价。包含社交媒体因素的资产定价模型在解释股票横截面收益率方面表现更好。接下来探索社交媒体的经济解释。

3 社交媒体溢价的经济解释

社交媒体溢价可以为不完善的分散化提供补偿。Merton(1987)提出的投资者认知假说认为，在不完全信息环境下，投资者只知道可用股票的一个子集，因此认知度较低的股票需要提供风险溢价来补偿其持有人的不完全分散化。因此，识别度较低的股票应该具有较高的回报率。在实证研究中，由于“投资者认知度”是无形的，大多数文献以股东数作为指标来表示上期投资者认知度（Bodnaruk and Ostberg,2009;Zhu and Jiang,2018）。然而，股东数(shareholder base)仅每季度报告一次，因此之前的研究不得不从不同的角度选择代理变量。其中包括媒体报道(Fang and Peress,2009;Barber and Odean,2008)、广告投资(Grullonetal.,2004;Chemmanur and Yan,2010)以及公司名称(Green and Jame,2013)。

在作者的研究中，关键变量股吧发帖数量衡量关于股吧股票的帖子数量。投资者要想发帖，必须知道股票的名称或代码，这样才能保证对股票有一定的认知度。正是从这个角度，作者认为社交媒体帖子的数量是投资者对股票认知度的直接衡量标准。此外，社交媒体为数以千计的投资者提供了一个便捷的识别渠道，信息成本低。股吧中每个帖子的平均读者数和中位数分别为 2068 和 1240，表明帖子中的信息将传播到约 1000 至 2000 名投资者。因此，更多的社交媒体帖子意味着信息将

到达更多的投资者，从而提高投资者的认知度。对股票的认知导致投资者在其投资组合分配中考虑该股票。因此，社交媒体上的帖子应该与股东数和其他投资者认知呈正相关。

因此，作者测试了股吧发帖数量是否与之前研究中使用的其他投资者认知指标呈正相关，包括股东数量和增长、大众媒体报道和广告成本。中国要求上市公司在季度报告中披露股东人数等财务指标。对于每个季度，作者对股吧发帖数量的对数代理进行横截面回归，并控制公司规模、公司年龄、股价（上一年公司股票中位数价格的倒数）、回报和波动性（过去三个月股票回报的平均标准差），如 Bodnaruk and Ostberg (2009)。表 7 报告了样本季度的平均估计值。

图表 7 投资者认知和股吧发帖数量

	(1) ln(Shareholder)	(2) Δ Shareholder	(3) ln(Mass media)	(4) ln(Advertisement)
ln(Gubapost)	0.454*** (25.018)	0.031*** (5.604)	0.212*** (23.111)	0.079** (2.781)
ln(Mass media)	-0.022** (-2.120)	0.013*** (5.075)	-	0.149*** (3.494)
ln(Size)	0.397*** (28.015)	0.003 (0.637)	0.521*** (29.646)	-0.096*** (-3.064)
ln(Age)	0.230*** (13.985)	-0.019*** (-3.791)	-0.124*** (-6.148)	0.170*** (3.603)
1/(share price)	5.699*** (22.331)	-0.072 (-0.790)	-1.554*** (-5.392)	-4.827*** (-3.568)
Return	-0.367*** (-8.756)	-0.194*** (-8.454)	0.198** (2.568)	0.210 (0.938)
Volatility	-10.978*** (-6.813)	7.769*** (10.063)	32.054*** (12.759)	-27.309*** (-3.286)
Intercept	-3.876*** (-14.301)	-0.387*** (-3.926)	-10.522*** (-20.903)	2.907*** (3.336)
R ²	0.816	0.155	0.485	0.045

资料来源：华安证券研究所整理

表 7 中的列(1)表明，社交媒体发帖数量与股东数量之间存在显著的正相关关系。股吧发帖数量增加一个对数单位对应于股东数增加 0.454 个对数单位。由于股东数是一个持续性较强的变量，作者还检验了股吧发帖数量是否与股东数的变化呈正相关。第(2)列显示，股吧发帖数量与股东数变化呈显著正相关。股吧发帖数量增加一个对数单位将使股东总数增加 3.1%。此外，作者使用了 Fang 和 Peress(2009)与 Barber 和 Odean(2008)的媒体报道，Grullon et al.(2004)的广告投资以及 Chemmanur 和 Yan(2010)作为投资者认知的代理。同样，作者在第(3)和(4)列中发现了显著的正相关关系。这些结果表明，股吧发帖数量与传统的投资者认知代理正相关。

此外，与其他投资者认知度衡量指标相比，股吧发帖数量衡量指标具有一些优势。第一，媒体报道和广告投资是建立在信息传播的基础上的，不能保证投资者收到信息和认知股票。而股吧发帖数量则是直接衡量，投资者在 Guba-forum 发帖前必须对一只股票有一定的认知度。其次，股票的股东数反映了在其投资组合中持有该股票的投资者数量。但是，它忽略了认知该股票并可能在未来将其添加到其投资组合中的潜在投资者。股吧发帖数量反映了当前和未来投资者的股票认知度。第三，股吧发帖数量的历史数据和实时数据都可以轻松获取，不受季度披露限制。

作者认为社交媒体上的帖子扩大了投资者的认知度并增加了股东数。然而，拥有庞大股东数的股票可能会在社交媒体上获得更多帖子，这可能会带来内生性问题。因此，作者将股票按股东数和股吧发帖数量双重排序为 2×2 个子样本，并对每个子样本实施了前文的回归。表 8 展示了结果。

图表 8 按股东数和股吧发帖数量分组的股东数回归

	(1) Low Gubapost Low Shareholders	(2) Low Gubapost High Shareholders	(3) High Gubapost Low Shareholders	(4) High Gubapost High Shareholders
ln(Gubapost)	0.408*** (25.749)	0.161*** (10.599)	0.189*** (13.654)	0.320*** (11.610)
ln(Mass media)	-0.021** (-2.259)	-0.021*** (-3.958)	-0.009 (-1.251)	-0.007 (-0.856)
ln(Size)	0.239*** (24.342)	0.295*** (19.638)	0.182*** (13.647)	0.373*** (30.586)
ln(Age)	0.174*** (11.655)	0.100*** (9.112)	0.122*** (7.736)	0.141*** (9.751)
1/(share price)	6.040*** (16.203)	3.287*** (22.915)	4.279*** (17.603)	3.863*** (17.401)
Return	-0.387*** (-9.981)	-0.334*** (-3.941)	-0.221*** (-6.563)	-0.154*** (-4.530)
Volatility	-9.155*** (-6.356)	-4.209** (-2.671)	-1.067 (-1.043)	-9.581*** (-6.469)
Intercept	0.174 (0.633)	2.065*** (5.527)	3.303*** (7.217)	-1.242*** (-6.460)
R ²	0.650	0.484	0.418	0.723

资料来源：华安证券研究所整理

在表 8 的每个子样本中，社交媒体发帖数量与股东数量之间仍然存在显著的正相关关系。值得注意的是，在第（2）和（3）列中，对于高股东数和低股吧发帖数量的子样本以及低股东数和高股吧发帖数量的子样本，关系仍然显著，这表明在按股东数和股吧发帖数量双重排序的子样本中，正相关关系是稳健的。这进一步支持了股吧发帖数量是投资者认知的良好代理的论点。

在文献中，还有对社交媒体效应的其他解释。Kim 和 Kim(2014)从情绪的角度解释了社交媒体对横截面回报的影响。为了评估股吧发帖数量中可能的情绪影响，我们使用文本分析的朴素贝叶斯分类方法将股吧发帖数量分为三组（正面、中性和负面）（Kim and Kim(2014); Leung and Ton,2015）。作者对 149569708 个帖子进行分类；情绪中性的比例最高（约 74.8%），负面帖子占 16.6%，正面帖子占 8.6%。如果社交媒体效应与认知的相关性比情绪更密切，应该观察社交媒体溢价在积极、消极和中性情绪之间的关系。与表 2 类似，表 9 中报告了在不同情绪下按大小和股吧发帖数量双重排序的 25 个市值加权投资组合的平均 1 个月前回报相对无风险利率的超额收益。

图表 9 25 个大众投资组合在不同情绪组中的每月超额收益

Panel A Size \times Negative Gubapost portfolio							
		low		Gubapost		High	L-H
		1	2	3	4		
Size	small	0.037	0.034	0.028	0.021	0.010	0.027***
	2	0.032	0.026	0.021	0.016	0.002	0.030***
	3	0.027	0.022	0.017	0.011	0.007	0.020***
	4	0.021	0.017	0.013	0.014	0.007	0.013***
	big	0.017	0.015	0.016	0.013	0.004	0.013***
	S-B	0.020**	0.019**	0.012	0.008	0.006	
Panel B Size \times Neutral Gubapost portfolio							
		low		Gubapost		High	L-H
		1	2	3	4		
Size	small	0.039	0.033	0.026	0.021	0.014	0.025***
	2	0.034	0.024	0.023	0.016	0.000	0.034***
	3	0.027	0.021	0.017	0.013	0.006	0.021***
	4	0.021	0.016	0.015	0.011	0.009	0.012**
	big	0.013	0.014	0.014	0.015	0.006	0.008
	S-B	0.026***	0.019**	0.012*	0.006	0.008	
Panel C Size \times Positive Gubapost portfolio							
		low		Gubapost		High	L-H
		1	2	3	4		
Size	small	0.038	0.035	0.026	0.020	0.017	0.021***
	2	0.033	0.027	0.021	0.016	0.000	0.034***
	3	0.025	0.026	0.016	0.010	0.008	0.017***
	4	0.021	0.015	0.016	0.012	0.008	0.014***
	big	0.013	0.014	0.013	0.013	0.006	0.006
	S-B	0.025***	0.021***	0.013*	0.007	0.010	

资料来源：华安证券研究所整理

表 9 中的结果表明，社交媒体溢价存在于不同的情感群体中。表 9 中的 B 组显示，情感中立的社交媒体帖子的溢价与表 2 中整个社交媒体样本的溢价具有相似的规律。在每个组中，都有一个相似的规律：随着股吧发帖数量的增加，股票投资组合平均回报显著下降。也就是说，只要发帖量增加，无论情绪如何变化，相应的投资者认知度就会提高，溢价就会下降。结果表明，**社交媒体效应不是情绪，而是由投资者认知引起的。**

作者还研究了其他情绪代理变量，例如(Bakeretal.,2012)、交易量(Scheinkman and Xiong,2003;Bakeretal.,2012)和过去几个月的最大每日回报(Balietal.,2011;Fong and Toh,2014)。他们发现过去的高情绪（高换手率、成交量和 Rmax）导致股价高估，从而导致下个月的低回报。作者进一步测试这个，看看社交媒体效应是否可以包含在情感效应中。我们用 1/3 低换手率股票的投资组合的回报减去 1/3 高换手率股票的投资组合的回报来构建 TurnoverF。类似地构造 VolumeF 和 MAXF。作者用这些情绪因素对 SMF 进行回归，并控制常用风险因素，以观察其是否会产生显著的正收益。

图表 10 针对情感因素的 SMF 的因素冗余测试

	(1)	(2)	(3)
TurnoverF	0.201 (1.558)	0.210* (1.784)	0.228* (1.765)
VolumeF	0.436*** (3.135)	0.451*** (3.270)	0.417*** (3.032)
MAXF	-0.010 (-0.084)	-0.027 (-0.261)	-0.031 (-0.271)
Mkt	0.024 (0.400)	0.035 (0.640)	0.023 (0.393)
SMB	0.215 (1.278)	0.098 (0.580)	0.245 (1.277)
HML	-0.535*** (-3.574)	-0.523*** (-3.524)	-0.494*** (-2.868)
UMD		0.117*** (2.808)	
RMW			-0.055 (-0.303)
CMA			-0.149 (-0.689)
MMF	0.103 (0.956)	0.183* (1.713)	0.125 (1.376)
LIQ	-0.354*** (-4.720)	-0.297*** (-4.804)	-0.347*** (-4.546)
Intercept	0.006** (2.451)	0.006** (2.344)	0.006** (2.267)
R ²	0.928	0.934	0.929

资料来源：华安证券研究所整理

表 10 报告了回归结果。在每一列中，作者控制 MMF 和 LIQ 以及其他常用因素，如表 4 的第(3)列。结果表明情绪效应与社交媒体效应正相关。作者观察到 VolumeF 和 SMF 之间的正相关。总的来说，回归的截距都显著为正，这表明**社交媒体效应不能包含在情绪效应中**。

总之，作者发现股吧发帖数量与股东数量呈正相关，并且这种关系在按股东数量和股吧发帖数量进行双重排序的子样本中是稳健的。作者还发现社交媒体效应不能包含在情绪效应中，这支持了股吧发帖数量是衡量投资者认知度而不是衡量情绪的观点。因此，作者得出结论，与文献中提出的其他方法相比，股吧发帖数量是衡量投资者认知度的更好方法；因此，作者的论证支持了默顿提出的投资者认知假说。

4 社交媒体效应溢价

作者发现社交媒体帖子较少的股票会获得回报溢价。作者研究了社交媒体发布率低的多头股票和社交媒体发布率高的空头股票投资组合的样本外的盈利能力。作者在社交媒体帖子上构建十分位数投资组合，并获得这些投资组合的市值加权回报。表 11 报告了投资组合的月度表现。

图表 11 股吧发帖数量投资组合的月度回报

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Decile	Return	Alpha(3)	Alpha(4)	Alpha(5)	Alpha (5+LIQ)	Odds	Std	Sharpe Ratio
D1(low)	0.0267	0.0117***	0.0118***	0.0120***	0.0131***	0.7077	0.0984	0.2713
D2	0.0214	0.0098***	0.0098***	0.0094***	0.0104***	0.7538	0.0895	0.2391
D3	0.0202	0.0097***	0.0097***	0.0098***	0.0103***	0.6615	0.0896	0.2248
D4	0.0176	0.0078***	0.0077***	0.0075***	0.0082***	0.5077	0.0903	0.1947
D5	0.0140	0.0060**	0.0060**	0.0059**	0.0067**	0.5846	0.0875	0.1597
D6	0.0149	0.0064*	0.0063*	0.0061*	0.0069**	0.5077	0.0849	0.1756
D7	0.0183	0.0114***	0.0112***	0.0112***	0.0107***	0.5538	0.0827	0.2211
D8	0.0096	0.0023	0.0023	0.0027	0.0030	0.5231	0.0802	0.1200
D9	0.0093	0.0030*	0.0029*	0.0028	0.0024	0.4154	0.0747	0.1244
D10(high)	0.0025	-0.0034	-0.0034	-0.0030	-0.0036	0.2923	0.0768	0.0325
D1-D10	0.0242** (2.2387)	0.0150*** (2.804)	0.0152*** (2.880)	0.0150** (2.653)	0.0166*** (3.065)	0.6308	0.0872	0.2777
Tercile1-Tercile3	0.0154** (1.9573)	0.0099*** (2.919)	0.0100*** (2.943)	0.0096*** (2.689)	0.0108*** (3.240)	0.5846	0.0636	0.2428

资料来源：华安证券研究所整理

表 11 的第(1)列报告了每个投资组合的一个月前市值加权超额收益（相对于无风险利率）。股吧发帖数量最低 10% 的投资组合的平均每月超额收益为 2.6%。随着发布量的增加，超额收益随着投资组合的减少而减少。最低 1/10 和最高 1/10 股吧发帖数量之间的回报率之差为每月 2.4%，非常显著。此外，作者将结果与按三分位数分组的投资组合的收益进行比较，多空投资组合仍有 1.5% 左右的显著收益。控制三、四、五和流动性因素，在列(2)-(5)中报告了风险调整后的回报和截距(alpha)。对于大多数投资组合，alpha 始终为正，并且在经济上和统计上都显著不等于零。alpha 值介于 1.3%（第(5)列中的最低发布组合）和 -0.36%（第(5)列中的最高发布组合）之间，溢价为 1.66%。表 11 第(6)列报告了投资组合表现优于市场投资组合的胜率，表明在样本中 70% 的月份中，社交媒体发布水平最低的投资组合的多头头寸可以跑赢大盘市场组合。多空组合的胜率是 63%，这表明在大多数情况下，多空组合的表现优于市场组合。最后两列反映了投资组合的风险，即投资组合的标准差和夏普比率。

作者进一步检验了这种盈利能力是否可以在不同的公司特征子样本中持续下去。通过在最低 10% 和最高 10% 的空头头寸上建立投资组合，使用公司规模、账

面市值比和动量作为公司的特征，并将它们分类为三个子样本。在每个子样本中，报告了由不同因子模型估计的投资组合的异常收益(alpha)。

图表 12 按公司特征分组的组合回报

FF 3	Carhart 4	FF 5	LIQ
<i>Panel A: By Firm Size</i>			
		Small	
0.025*** (4.265)	0.025*** (4.267)	0.022*** (3.683)	0.023*** (3.828)
		Medium	
0.023*** (3.953)	0.023*** (3.948)	0.020*** (3.428)	0.021*** (3.437)
		Big	
0.012*** (2.689)	0.012*** (2.672)	0.011** (2.463)	0.012*** (2.770)
<i>Panel B: By Book-to-Market</i>			
		Low	
0.019*** (3.486)	0.019*** (3.499)	0.018*** (3.151)	0.019*** (3.540)
		Medium	
0.017*** (2.965)	0.017*** (3.009)	0.014** (2.485)	0.016*** (2.855)
		High	
0.010* (1.933)	0.010* (1.934)	0.009* (1.755)	0.010* (1.851)
<i>Panel C: By Past 12-Month Momentum</i>			
		Low	
0.011** (2.304)	0.011** (2.310)	0.011** (2.111)	0.011** (2.205)
		Medium	
0.012* (1.952)	0.012* (1.938)	0.011* (1.784)	0.012* (1.858)
		High	
0.021*** (2.897)	0.021*** (2.869)	0.020** (2.653)	0.022*** (2.837)

资料来源：华安证券研究所整理

表 12 报告了多空策略在不同公司特征子样本中的显著风险调整回报。作者注意到 Fang 和 Peress(2009)检验了媒体报道效应的非流动性假设。非流动性假设认为该效应反映了错误定价，可以通过套利交易消除。类似地，作者检查社交媒体效应是否支持非流动性假设。如果异常收益集中在流动性最差的股票中，则可以通过套利消除社交媒体效应，这支持了流动性不足的假设。作者使用股吧发帖数量最低 10% 的多头股票和最高 10% 的空头股票，使用 Amihud 的非流动性比率和交易量作为非流动性的代理。在按这些指标排序的样本中检查投资组合的风险调整回报。表 13 的面板 A 和 B 报告了不同因子模型下投资组合的风险调整回报。报告的 alpha 在所有非流动性子样本中都显著为正。例如，面板 A 第一列中的 alpha 范围从 1.2% 到 1.7%，这表明没有足够的证据支持非流动性假设。

作者还研究了情绪效应对社交媒体效应交易策略盈利能力的影响。和以前一样，我们根据成交量、Rmax 和交易量等情绪指标对股票进行排序(Scheinkman and Xiong, 2003; Baker et al., 2012; Bali et al., 2011; Fong and Toh, 2014)。表 13 的面板 B 和 C 报告了不同因子模型下投资组合的风险调整回报。显著正的 alpha 表明多空社交媒体效应策略的超额收益不能用情绪效应来解释。

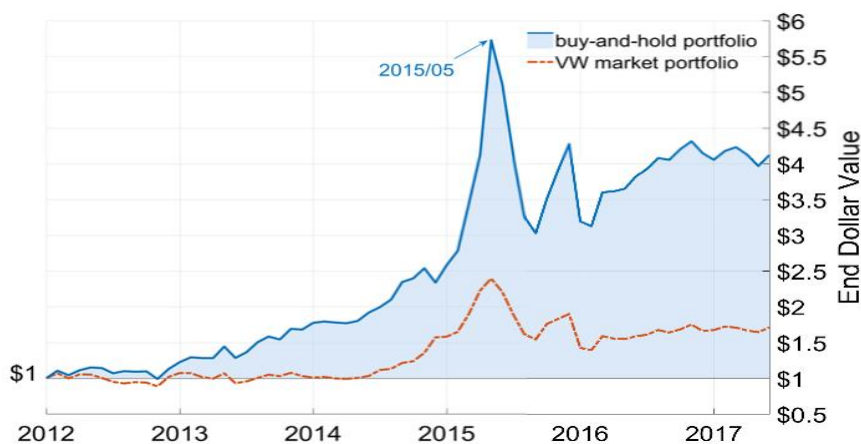
图表 13 按流动性和情绪特征分组的交易利润

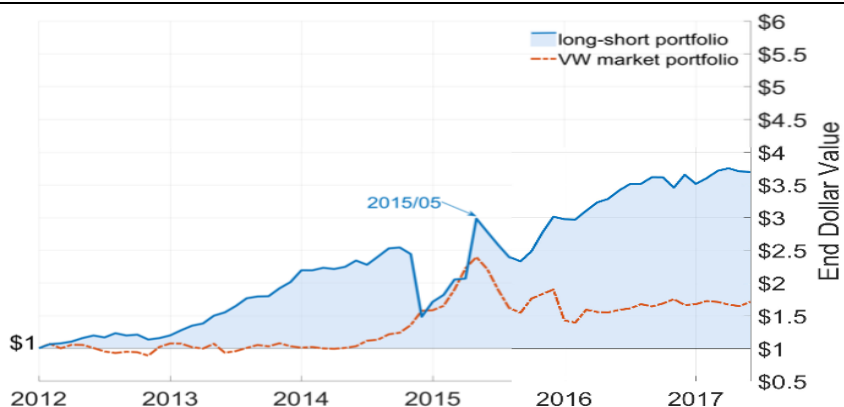
FF 3	Carhart 4	FF 5	LIQ
<i>Panel A: By Amihud's Illiquidity Ratio</i>			
0.012** (2.203)	0.012** (2.184)	0.010* (1.822)	0.012** (2.061)
0.015*** (2.995)	0.016*** (3.000)	0.012** (2.407)	0.014*** (2.739)
0.017*** (3.028)	0.017*** (3.020)	0.016*** (2.720)	0.016** (2.654)
<i>Panel B: By Dollar Trading Volume</i>			
0.015*** (3.108)	0.015*** (3.238)	0.014*** (2.845)	0.016*** (3.285)
0.013** (2.306)	0.013** (2.331)	0.011* (1.892)	0.012** (2.147)
0.014** (2.518)	0.014** (2.493)	0.013** (2.215)	0.014** (2.426)
<i>Panel C: By Turnover</i>			
0.011** (2.265)	0.011** (2.247)	0.012** (2.439)	0.013*** (2.679)
0.014** (2.437)	0.014** (2.652)	0.011* (1.911)	0.012** (2.022)
0.025*** (3.179)	0.025*** (3.280)	0.022*** (2.798)	0.024*** (2.955)
<i>Panel D: By R_{max}</i>			
0.012*** (2.731)	0.012*** (2.709)	0.013*** (2.886)	0.013*** (2.890)
0.011* (1.761)	0.011* (1.756)	0.011 (1.645)	0.012* (1.887)
0.017** (2.467)	0.018** (2.450)	0.014* (1.982)	0.014** (2.031)

资料来源：华安证券研究所整理

作者还制定了两种交易策略来测试社交媒体溢价驱动的盈利能力是否具有持久性。第一个是**买入并持有策略**，以做多社交媒体帖子中最低 10% 的股票。第二种策略采用多空策略，即做多股吧发帖数量最低十分位数的股票，做空最高十分位数的股票。根据更新的股吧发帖数量每月重新平衡两个投资组合。图 1 显示了两种策略与市场投资组合相比的累积回报。

图表 14 策略的累积回报





资料来源：华安证券研究所整理

从图表 14 中看到，这两种策略在整个样本期间实现了约 300% 的回报。尽管出现回撤，但它们的表现似乎优于市场投资组合。接下来，将投资组合持有期延长至 12 个月。表 15 报告了两种策略的年度回报。表 15 显示，买入持有策略可以实现近 40% 的年化收益率和 22.6% 的超额收益，对市场投资组合的胜率为 90%。此外，以沪深 300（规模最大的 300 家公司）、中证 500（规模中等的 500 家公司）和中证 1000（规模较小的 1000 家公司）三个市场指数作为衡量投资组合表现的基准。多头策略的超额收益分别为 27.3%、19.5%、12.9%。对于多空策略，年化收益率约为 32%，夏普比率为 1.168。

图表 15 社交媒体效应策略年度表现

	buy-and-hold portfolio	long-short portfolio
Mean num of Stocks	200	200 long & 200 short
Annualized Return	0.3966*** (6.8972)	0.3219*** (8.5848)
Excess Return relevant to VW-Market Portfolio	0.2265*** (9.0259)	0.1518** (2.0469)
Excess Return relevant to CSI 300 Index Return	0.2734*** (8.7269)	0.1987*** (2.7099)
Excess Return relevant to CSI 500 Index Return	0.1950*** (7.0832)	0.1203* (1.7039)
Excess Return relevant to CSI 1000 Index Return	0.1290*** (4.3559)	0.0544 (0.8066)
Odds	0.9074	0.7593
Sharpe Ratio	0.9386	1.1682

资料来源：华安证券研究所整理

总体而言，这两种策略的表现表明，由社交媒体溢价构建的投资组合能够产生显著的经济利润。因此，投资组合经理和市场从业者应该在他们的投资组合和风险管理中考虑社交媒体溢价。

5 结论

过往对资产定价中社交媒体影响的研究主要是在时间序列影响和情绪影响方面进行的。作者重点研究社交媒体对股票横截面收益的影响。使用股吧发帖数量作为衡量中国股票市场社交媒体报道的指标，作者发现社交媒体溢价显著。通过构建基

于社交媒体溢价的社交媒体因子进一步识别与投资者认知相关的风险，并发现社交媒体因子提供了著名的 Fama-French 因子和 Fang 和 Peress(2009)的大众媒体因子未涵盖的新信息，这个因子不能包含在情绪影响中。由社交媒体因子增强的 Fama-French 因素模型能够产生显著的风险调整后回报。风险定价测试发现，每单位风险，社交媒体效应可以带来市场投资组合收益 2% 的变化。GRS 统计数据还表明，社交媒体因子增强了现有因子模型的性能。

为什么社交媒体因子可以提高著名的 Fama-French 因子模型对资产横截面收益的解释力？作者表明该机制是**与投资者认知相关的风险**。在信息不完全市场中，投资者对所有股票的认知度有限，其投资组合不完全分散化，持有此类投资组合需要得到补偿。作者发现，与文献中现有的相比，社交媒体帖子的数量是投资者认知度的更好衡量标准。因此，作者为支持默顿（1987）的投资者认知假说提供了直接证据。此外，作者还发现基于社交媒体溢价的交易策略可以产生可观的利润，这表明市场从业者应该在他们的投资组合和风险管理中考虑社交媒体效应。

文献来源：

作者核心内容摘选自 Xiangtong Meng, Wei Zhang 和 Youwei Li 等人在 International Review of Financial Analysis 上的论文 Social media effect, investor recognition and the cross-section of stock returns。

风险提示：

作者结论基于历史数据与海外文献进行总结；不构成任何投资建议。

重要声明

分析师声明

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格，以勤勉的执业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，本报告所采用的数据和信息均来自市场公开信息，本人对这些信息的准确性或完整性不做任何保证，也不保证所包含的信息和建议不会发生任何变更。报告中的信息和意见仅供参考。本人过去不曾与、现在不与、未来也将不会因本报告中的具体推荐意见或观点而直接或间接接收任何形式的补偿，分析结论不受任何第三方的授意或影响，特此声明。

免责声明

华安证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。本报告中的信息均来源于合规渠道，华安证券研究所力求准确、可靠，但对这些信息的准确性及完整性均不做任何保证。在任何情况下，本报告中的信息或表述的意见均不构成对任何人的投资建议。在任何情况下，本公司、本公司员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。华安证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

本报告仅向特定客户传送，未经华安证券研究所书面授权，本研究报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。如欲引用或转载作者内容，务必联络华安证券研究所并获得许可，并需注明出处为华安证券研究所，且不得对作者进行有悖原意的引用和删改。如未经本公司授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。本公司并保留追究其法律责任的权利。

投资评级说明

以本报告发布之日起 6 个月内，证券（或行业指数）相对于同期沪深 300 指数的涨跌幅为标准，定义如下：

行业评级体系

- 增持—未来 6 个月的投资收益率领先沪深 300 指数 5%以上；
- 中性—未来 6 个月的投资收益率与沪深 300 指数的变动幅度相差-5%至 5%；
- 减持—未来 6 个月的投资收益率落后沪深 300 指数 5%以上；

公司评级体系

- 买入—未来 6-12 个月的投资收益率领先市场基准指数 15%以上；
- 增持—未来 6-12 个月的投资收益率领先市场基准指数 5%至 15%；
- 中性—未来 6-12 个月的投资收益率与市场基准指数的变动幅度相差-5%至 5%；
- 减持—未来 6-12 个月的投资收益率落后市场基准指数 5%至；
- 卖出—未来 6-12 个月的投资收益率落后市场基准指数 15%以上；
- 无评级—因无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使无法给出明确的投资评级。市场基准指数为沪深 300 指数。