

《基金经理媒体报道与个体投资者行为》附录

附录 1 样本代表性

本文从投资者行为、基金风格分布以及投资者特征分布三方面提供了本文样本数据代表性的部分证据。从投资者行为角度出发，据中国证券投资基金管理业年报（2021），全市场股票型和混合型基金在 2020 年总认申购金额为 71849 亿元，总赎回金额为 67442 亿元，使用本文方法计算的净申购率为 0.032，与本文同期样本 0.039 的净申购率基本一致；另外，使用基金层面申购赎回份额计算的基金份额净申购率变量均值为 -0.21，个体投资者逐笔数据加总计算的基金份额净申购率变量均值为 -0.22，二者基本一致；最后，将基金申购赎回份额加总到季度计算季度份额净申购率，基金总体数据与使用本文样本计算的指标相关系数达到 0.898。以上结论均说明本文样本在该研究主题下具有较好代表性。

基金的投资风格分布报告于附表 1，结果表明，本文选取的基金分布与全部主动权益基金分布整体较为接近。

附表 1 基金风格分布代表性分析

基金风格	基金风格占比 (%)	
	本文样本	全部主动权益基金
平衡型	61.89	66.22
成长型	18.14	10.63
增值型	9.13	9.97
稳健成长型	5.70	5.67
收益型	2.63	5.16
积极成长型	2.14	0.70
价值型	0.25	1.03
稳健增值型	0.12	0.43
保本增值型	0.00	0.20

从投资者特征分布出发，将本文投资者性别、交易频率和金融资产总规模特征分布与景顺长城基金、富国基金和交银施罗德基金于 2021 年 3 月发布的《公募权益类基金投资者盈利洞察报告》中三大基金公司全部个人投资者，及中国证券投资基金管理业协会发布的《全国公募基金市场投资者状况调查报告》（2020 年）总体投资者特征进行对比。投资者特征分布见附表 2，本文样本投资者特征与市场整体投资者特征分布接近，本文样本具有一定代表性。

附表 2 投资者特征分布代表性分析

基金投资者特征	具有该特征的投资者占比 (%)	
	本文样本	报告样本
性别		
男性	44.49	49.79
女性	55.51	50.21
月度交易频率	本文样本	报告样本
低于 5 次	83.98	83.66
5-10 次	7.92	7.52
10-20 次	4.30	3.28
20 次以上	3.80	5.54

金融资产总规模(元)	本文样本	报告样本
低于10万	36.15	34.00
10-50万	34.88	34.20
50-100万	12.18	17.40
100-300万	12.23	9.30
300万以上	4.56	5.00

注：报告样本中性别和交易频率分布数据来源于景顺长城基金、富国基金和交银施罗德基金联合发布的《公募权益类基金投资者盈利洞察报告》，金融资产总规模分布来源于中国证券投资基金业协会发布的《全国公募基金市场投资者状况调查报告》（2020年）。

附录2 主要变量符号及释义

附表3 主要变量符号及释义

变量符号	变量名	释义
<i>BuySellRatio</i>	净申购率	(月度基金申购金额-月度基金赎回金额)/(月度基金申购金额+月度基金赎回金额)
<i>Buy</i>	申购金额	投资者对基金的月度申购金额
<i>Sell</i>	赎回金额	投资者对基金的月度赎回金额
<i>DummyNews</i>	是否被报道	管理基金的基金经理在当月是否被媒体报道提及，是取值为1，否取值为0
<i>LnNews</i>	媒体报道数量	管理基金的基金经理在当月的媒体报道数量，加一取自然对数
<i>Performance</i>	基金业绩	以基金净现值NAV计算的基金价格增长率在同一基金投资风格中的百分比排序
<i>PerformanceSqr</i>	基金业绩平方项	基金业绩变量的平方
<i>LnTNA</i>	基金规模	基金资产净值取对数
<i>Fundrisk</i>	基金风险	基金月度日收益率标准差
<i>Fundflow</i>	基金资金净流入	基金季度净现金流入
<i>ExpenseRatio</i>	基金费用	基金半年度总费用占基金规模的比率
<i>LnAge</i>	基金存续时长	基金存续月度时长，加一取自然对数
<i>Mgender</i>	基金经理性别	女性基金经理为1，男性基金经理为0
<i>Mdegree</i>	基金经理学历	基金经理学历，1为本科，2为硕士，3为博士
<i>Mtenure</i>	基金经理管理时间	基金经理管理该基金的月份数，加一取自然对数

附录3 描述性统计

附表4 主要变量描述性统计

变量	均值	标准差	最小值	中位数	最大值
<i>BuySellRatio</i>	0.0553	0.3606	-1.0000	0.0000	1.0000
<i>Buy</i>	1056.6570	5868.8703	0.0000	0.0000	5.00e+04
<i>Sell</i>	932.7111	6133.1177	0.0000	0.0000	5.65e+04
<i>DummyNews</i>	0.4783	0.4995	0.0000	0.0000	1.0000
<i>LnNews</i>	0.7890	1.0530	0.0000	0.0000	4.0943
<i>Performance</i>	0.5799	0.3069	0.0004	0.6381	1.0000

<i>PerformanceSqr</i>	0.4305	0.3197	0.0000	0.4072	1.0000
<i>Fundrisk</i>	0.6085	1.2001	-3.2913	0.8309	2.6586
<i>LnTNA</i>	21.3665	1.6078	17.0878	21.5863	24.0029
<i>Fundflow</i>	-11.4862	10.8192	-41.3000	-10.7000	10.8800
<i>LnAge</i>	3.7857	0.9776	1.4663	3.9753	5.3254
<i>ExpenseRatio</i>	0.0167	0.0190	0.0043	0.0115	0.1383
<i>Mgender</i>	0.1327	0.3393	0.0000	0.0000	1.0000
<i>Mdegree</i>	2.1140	0.3843	1.0000	2.0000	3.0000
<i>Mtenure</i>	3.2249	0.8852	0.3830	3.3393	4.6092

附录 4 内生性分析

首先，采用工具变量法来缓解内生性问题。工具变量 *LnStyleNews* 和 *LnCategoryNews* 为同一投资风格下，以及同一 Wind 基金二级投资分类下除管理本基金的基金经理以外全部基金经理的媒体报道数量，加一取自然对数。工具变量 *LnNewsSimilarFund* 为相似基金的基金经理媒体报道数量，加一取自然对数。回归结果报告于附表 5。如第（1）列、第（3）列和第（5）列所示，工具变量与该基金的基金经理媒体报道数量显著正相关。由弱工具变量检验结果可知，工具变量的 *Kleibergen-Paap Wald rk F* 统计量分别为 1367.435、1215.594 和 499.68，远大于临界值 16.38，可以拒绝弱工具变量假设。附表 5 第（2）列、第（4）列和第（6）列表明，使用工具变量法缓解内生性问题后，主回归结论不变。其次，基于本文样本探究从未受到过媒体报道的基金经理，首次被媒体报道对投资者对基金净申购率带来的影响。本文识别了基金经理第一次被报道的时间，且保留基金经理首次报道时间在 2019 年 9 月以后的基金经理子样本。基于这一子样本识别处理组和对照组，并识别出所有基金经理第一次被媒体报道的月份，构建基金经理第一次被媒体报道的冲击变量 *DID*。参考 Cengiz et al. (2019) 采用堆叠 *DID* 方法 (*Stacked DID*) 进行回归。结果见附表 5 第（7）列，变量的系数显著为正，结果表明基金经理第一次被媒体报道后，投资者会增加对其管理基金的净申购率，主回归结论保持稳健。

附表 5 投资者基金申购赎回数据交易特征

变量	<i>LnNews</i>	<i>BuySell</i>	<i>LnNews</i>	<i>BuySell</i>	<i>LnNews</i>	<i>BuySell</i>	<i>BuySell</i>
		<i>-Ratio</i>		<i>-Ratio</i>		<i>-Ratio</i>	<i>-Ratio</i>
	IV:		IV:		IV:		<i>Stacked</i>
	<i>LnStyleNews</i>		<i>LnCategoryNews</i>		<i>LnSimilarFundNews</i>		<i>DID</i>
	第一阶段	第二阶段	第一阶段	第二阶段	第一阶段	第二阶段	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>LnNews</i>		0.0341*** (3.6804)		0.0431*** (4.4609)		0.0530*** (4.4169)	
<i>LnStyleNews</i>		0.1376*** (36.9756)					
<i>LnCategoryNews</i>			0.1266*** (34.8612)				
<i>LnNewsSimilarFund</i>					0.2609*** (22.3535)		
<i>DID</i>						0.0227* (1.9034)	

控制变量	Yes	Yes	Yes	Yes	Yes	Yes	Yes
个体固定效应	Yes	Yes	Yes	Yes	Yes	Yes	Yes
年一月固定效应	Yes	Yes	Yes	Yes	Yes	Yes	Yes
投资风格固定效应	Yes	Yes	Yes	Yes	Yes	Yes	Yes
样本量	37318	37318	37318	37318	37318	37318	23944
F/R ²	1367.2014	0.0044	1215.3042	0.0018	499.6801	0.0020	0.2013

注：第一阶段报告 Kleibergen-Paap Wald rk F 统计量，第二阶段报告 R²统计量。

附录 5 其他稳健性检验

(一) 排除相关政策干扰

本文通过国务院政策文件库和中国政府网政府信息公开平台手动收集了样本期内全部政策公告。经逐条阅读筛选，获得 3 条符合条件的政策作为本文政策控制变量¹，分别影响软件与信息技术服务行业、畜牧业以及汽车制造业。构建衡量基金是否受到政策影响的虚拟变量 *Treat*。若基金在交易月份滞后一季度的前十大重仓股中含软件与信息技术服务行业、畜牧业以及汽车制造业个股时，*Treat* 取值为 1，否则为 0²。构建政策产生影响的时间虚拟变量 *Post*，交易月份晚于相应政策实施时间取值为 1。由于 *Post* 变量被年一月固定效应吸收，将 *Treat* 变量和二者相乘构建的 *Policy* 变量加入主回归模型中重新进行回归。此外，政策实施时间最早为 2020 年 8 月，剔除晚于这一时点的交易月份重新进行回归。回归结果见附表 6，与主回归结果基本一致。

附表 6 控制政策影响效应

变量	BuySellRatio							
	控制政策变量1		控制政策变量2		控制政策变量3		剔除2020年8月后样本	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DummyNews	0.0238*** (5.3017)		0.0244*** (5.4381)		0.0245*** (5.4751)		0.0287*** (4.8994)	
LnNews		0.0125*** (4.2660)		0.0122*** (4.1621)		0.0122*** (4.1567)		0.0164*** (3.9645)
Treat	-0.0093 (-1.5834)	-0.0118** (-1.9848)	0.0120 (1.2126)	0.0110 (1.1090)	-0.0085 (-1.1556)	-0.0078 (-1.0582)		
Policy	-0.0049 (-0.5718)	-0.0047 (-0.5440)	-0.0106 (-0.4292)	-0.0094 (-0.3799)	0.0415*** (2.8791)	0.0389*** (2.7079)		
控制变量和固定效应	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
样本量	37318	37318	37318	37318	37318	37318	20936	20936
Adj. R ²	0.1623	0.1623	0.1622	0.1624	0.1624	0.1624	0.1262	0.1261

(二) 替换变量度量方式

¹ 筛选过程：首先通过国务院政策文件库和中国政府网政府信息公开平台获取 2019 年 9 月到 2020 年 12 月政策共 193 条。其次，对全部政策公告逐条仔细阅读，筛选出可能会对本文结果产生影响的行业或产业政策，最终获得 3 条符合条件的政策，即 2020 年 8 月印发的《新时期促进集成电路产业和软件产业高质量发展的若干政策》，2020 年 9 月印发的《国务院办公厅关于促进畜牧业高质量发展的意见》以及 2020 年 11 月印发的《新能源汽车产业发展规划（2021—2035 年）》。

² 本文行业分类均选择证监会行业分类 2012 年版。

替换被解释变量为净申购额变量 $NetInflow$, 即投资者每月基金申购金额与基金赎回金额之差除以 1000。替换解释变量为基金经理媒体报道数量在同一投资风格基金中的排序变量 $RankNews$ 和高媒体报道数量虚拟变量 $NewsHigh$ 。基金经理被媒体报道次数高于同一投资风格所有基金经理媒体报道的中位数, $NewsHigh$ 取 1, 反之则取 0。替换变量后重新进行回归, 结果报告于附表 7 第 (1) - (4) 列, 与主回归结果一致。

(三) 更换回归模型

本文使用 Tobit 模型替换基准回归 OLS 模型进行重新回归。另外, 构建虚拟变量 $Traded$ 来衡量投资者是否进行交易, 当月发生交易取为 1, 反之取值为 0。使用 Logit 模型将被解释变量是否进行交易 $Traded$ 对解释变量进行回归, 结果报告于附表 7 第 (5) - (8) 列, 与主回归结果相一致。

(四) 改变变量频率

改变样本数据频率到日度进行稳健性检验。被解释变量为日度净申购率, 解释变量 $DummyNews$ 为交易日期前 30 日基金经理是否被媒体报道提及, $LnNews$ 定义为交易日期前 30 日基金经理被媒体报道提及的数量加一再取对数, 其余控制变量频率和构建方法不变。结果报告于附表 7 第 (9)、(10) 列, 估计结果与基准回归结果一致。

附表 7 替换核心变量、回归模型以及数据频率的稳健性检验

变量	<i>NetInflow</i>		<i>BuySellRatio</i>		<i>BuySellRatio</i>		<i>Traded</i>		<i>BuySellRatio</i>	
	替换被解释变量		替换解释变量		Tobit 模型回归		Logit 模型回归		替换为日度数据	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>DummyNews</i>	0.3312*** (3.1717)				0.0243*** (5.8256)		0.2505*** (4.3207)		0.0033*** (6.7982)	
<i>LnNews</i>		0.1132* (1.9432)				0.0122*** (5.4798)		0.1563*** (4.7435)		0.0019*** (5.0740)
<i>RankNews</i>			0.0446*** (4.7365)							
<i>NewsHigh</i>				0.0214*** (4.7177)						
控制变量和固定效应	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
样本量	37318	37318	37318	37318	37371	37371	22585	22585	487677	487677
<i>Adj. R</i> ²	0.0289	0.0287	0.1622	0.1621	0.2731	0.2730	0.2693	0.2695	0.0258	0.0258

(五) 扩充回归样本

本文进一步获取 2015 年 1 月到 2019 年 8 月的基金投资者申购数据, 使用 2015 年 1 月到 2020 年 12 月样本检验基金经理媒体报道对投资者月度申购金额总额 Buy 的影响。采用泊松回归方法进行回归, 结果见附表 8 第 (1)、(2) 列。另外, 本文将基金经理团队管理的基金样本加回重新进行基准回归, 结果见附表 8 第 (3)、(4) 列, 本文结论依然稳健。

附表 8 扩充样本的稳健性检验

变量	<i>Buy</i>		<i>BuySellRatio</i>	
	样本时长扩充		基金经理样本扩充	
	(1)	(2)	(3)	(4)
<i>DummyNews</i>	0.4796*** (6.4476)		0.0297*** (7.4696)	

LnNews		0.1052*** (2.8593)		0.0060** (2.5684)
控制变量和固定效应	Yes	Yes	Yes	Yes
样本量	64900	64900	51347	51347
伪 $R^2/Adj. R^2$	0.4255	0.4219	0.1949	0.1941

附录 6 排除替代性解释

(一) 投资者学习效应

为排除投资者学习效应，本文用基金经理选股能力指标 $TMStkSlct$ 衡量基金经理管理能力，用基金持仓换手率 $Turnover$ 衡量信息不对称程度 (Solomon et al., 2014)，检验基金经理媒体报道是否对管理能力更强和信息不对称程度更高的基金的投资者净申购率影响更大。若媒体报道起信息扩散作用，促进投资者学习基金真实管理能力，做出更加理性的选择，则在基金经理管理能力更高和换手率更高的基金中，基金经理媒体报道对个体投资者净申购率的影响更显著。将上述变量与解释变量的交乘项加入模型进行回归，检验结果报告于附表 9 第(1)、(2)列。结果中交乘项的回归系数均不显著，说明在基金经理管理能力不同以及基金经理持仓换手率不同的基金之间，投资者受到基金经理媒体报道影响无差异，基金经理媒体报道对投资者净申购率的影响难以被投资者学习效应解释。

(二) 媒体情绪效应

首先讨论基金经理正面报道数量和负面报道数量对个体投资者投资行为的影响。若媒体报道情绪能够影响个体投资者的投资决策，媒体报道语调对申购和赎回影响方向相反 (Fang et al., 2014)。通过词典法判断基金经理媒体报道的情绪¹，将解释变量替换为当月正面报道的数量 $PositiveNews$ 和当月负面报道的数量 $NegativeNews$ 进行回归。进一步地，将被解释变量净申购率替换为申购金额 Buy 和赎回金额 $Sell$ ，进行泊松回归，结果报告于附表 9 第(3)-(5)列。结果显示，正面报道显著正向影响净申购率，负面报道影响不显著。同时正面报道增加，申购增加，而赎回并不会减少，与申购和赎回反向变动的假设不一致。其次，构建媒体报道情绪变量加入回归模型，单独检验情绪对投资者净申购率的影响。构建正面报道数量占全部报道数量的比例 $PositiveRatio$ 作为情绪代理变量，加入模型中进行回归。此外，使用上述媒体报道情感分析方法对每月基金经理全部媒体报道进行文本分析，构建基金经理在某月全部媒体报道的情绪指标 $AllNewsSentiment$ 。将情绪变量与解释变量 $LnNews$ 构建交乘项，加入模型中进行回归。结果见附表 9 第(6)、(7)列，正面报道比例和交乘项系数均不显著，说明媒体情绪因素不影响投资者净申购，排除媒体情绪效应的替代性解释。

附表 9 排除替代性解释

变量	$BuySellRatio$		$BuySellRatio$		Buy	$Sell$	$BuySellRatio$	$BuySellRatio$
	投资者学习效应				媒体情绪效应			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
$LnNews$	0.0121*** (4.1397)	0.0077*** (2.5987)				0.0122*** (4.1464)	0.0090** (2.4659)	
$TMStkSlct$	0.0030 (0.7684)							
$LnNews \times TMStkSlct$	0.0049							

¹ 词典中金融情绪词典部分来源于姜富伟等 (2021)，停用词，否定词和程度副词等功能词汇来源于玻森数据 (<http://bosonnlp.com>)。对每一篇文章分词后去掉停用词，通过金融情绪词典判断词语是正面或负面词语，并使用程度副词对词语进行权重调整，最终所有词语进行加权得分计算得到文本情绪分值。

	(1.0897)			
<i>Turnover</i>	0.0000			
	(0.1308)			
<i>LnNews</i> × <i>Turnover</i>	0.0000			
	(0.2970)			
<i>PositiveNews</i>		0.0127***	0.0955**	0.0305
		(4.2935)	(2.5054)	(0.7520)
<i>NegativeNews</i>		-0.0059	-0.1618	0.2152
		(-0.5146)	(-1.2795)	(1.4218)
<i>PositiveRatio</i>			0.0009	
			(0.0425)	
<i>AllNewsSentiment</i>				0.0427*
				(1.9086)
<i>LnNews</i> × <i>AllNewsSentiment</i>				0.0042
				(0.1515)
控制变量和固定效应	Yes	Yes	Yes	Yes
样本量	37288	36100	37318	37318
<i>Adj. R</i> ² / <i>R</i> ²	0.1622	0.1652	0.1621	0.4085
				0.2867
				0.1621
				0.1622

附录 7 明星基金经理的媒体报道影响

本文进一步探究基金经理媒体报道对个体投资者的影响，在明星基金经理与非明星基金经理之间是否有差异，以及明星基金经理媒体报道是否具有溢出效应。首先，讨论明星基金经理与非明星基金经理的媒体报道对个体投资者的影响是否有差异。参考林树等（2009），将当月基金业绩在同一投资风格中排名前 5% 的基金定义为明星基金，定义明星基金经理为其管理基金中包含明星基金的基金经理。本文还考虑将过去 5 年管理基金业绩¹在同一投资风格中排名前 5% 的基金经理定义为明星基金经理。将是否为明星基金经理的虚拟变量与解释变量相乘构建交乘项加入模型进行回归，结果见附表 10 第（1）、（3）列。交乘项系数均显著为正，表明明星基金经理媒体报道对于个体投资者的影响高于非明星基金经理。其次，进一步分析明星基金经理媒体报道是否具有溢出效应。本文保留基金经理为明星基金经理的样本，针对每一只明星基金经理管理的基金，构建投资者对同基金公司非明星基金经理所管理基金的平均净申购率变量，并对明星基金经理媒体报道数量进行回归，结果见附表 10 第（2）、（4）列。基金经理媒体报道数量系数显著为正，表明基金经理媒体报道具有正向溢出效应。

附表 10 明星基金经理与基金经理媒体报道影响

变量	<i>BuySellRatio</i>	<i>NotStarBuySellRatio</i>	<i>BuySellRatio</i>	<i>NotStarBuySellRatio</i>
	明星基金经理定义一		明星基金经理定义二	
	(1)	(2)	(3)	(4)
<i>LnNews</i>	0.0111*** (3.7293)	0.0589* (1.7198)	0.0104*** (3.3955)	0.0378* (1.7368)
<i>StarManager</i>	-0.0017		-0.0089	

¹ 计算基金经理过往 5 年业绩的方法：首先，在每个月，识别出每只基金的基金经理在过去五年内管理的所有基金，并计算这些基金在过去 60 个月的累计收益率；其次，将基金经理管理的所有基金的过往收益率，以基金规模为权重，计算加权平均收益率，得到基金经理过往 5 年的业绩指标。

	(-0.1436)	(-1.1345)
<i>LnNews × StarManager</i>	0.0199*	0.0192**
	(1.9419)	(2.4989)
控制变量和固定效应	Yes	Yes
样本量	37318	278
<i>Adj. R²</i>	0.1623	0.1582
	0.1623	0.1623
	0.1664	0.1664

参考文献

- [1] 姜富伟、孟令超和唐国豪, 2021, 《媒体文本情绪与股票回报预测》, 《经济学 (季刊)》第 4 期, 第 1323~1344 页。
- [2] 林树、李翔、杨雄胜和 Onkit Tam, 2009, 《他们真的是明星吗?——来自中国证券基金市场的经验证据》, 《金融研究》第 5 期, 第 107~120 页。
- [3] Cengiz, D., D. Arindrajit, L. Attila, and Z. Ben, 2019, “The Effect of Minimum Wages on Low-Wage Jobs*”, *The Quarterly Journal of Economics*, 134(3), pp. 1405~1454.
- [4] Fang, L. H., J. Peress and L. Zheng, 2014, “Does Media Coverage of Stocks Affect Mutual Funds’ Trading and Performance? ”, *The Review of Financial Studies*, 27(12), pp. 3441~3466.
- [5] Solomon, D. H., E. Soltes and D. Sosyura, 2014, “Winners in the Spotlight: Media Coverage of Fund Holdings as a Driver of Flows”, *Journal of Financial Economics*, 113(1), pp. 53~72.