

2022 年 03 月 27 日

金融工程研究团队

魏建榕（首席分析师）  
证书编号：S0790519120001

张翔（分析师）  
证书编号：S0790520110001

傅开波（分析师）  
证书编号：S0790520090003

高鹏（分析师）  
证书编号：S0790520090002

苏俊豪（分析师）  
证书编号：S0790522020001

胡亮勇（分析师）  
证书编号：S0790522030001

王志豪（研究员）  
证书编号：S0790120070080

盛少成（研究员）  
证书编号：S0790121070009

苏良（研究员）  
证书编号：S0790121070008

## 相关研究报告

《开源量化评论（47）-量体裁衣：可转债的分层组合方法》-2022.02.12

《开源量化评论（48）-宏观择时：多维度结合下的新视角》-2022.02.13

《开源量化评论（49）-为何近期北上持续撤离，南下持续抄底》-2022.03.15

《开源量化评论（50）-港股优选：技术面、资金面、基本面》-2022.03.23

## 自选股与点击量：投资者关注度的选股能力

——开源量化评论（51）

魏建榕（分析师）  
weijianrong@kysec.cn  
证书编号：S0790519120001

胡亮勇（分析师）  
huliangyong@kysec.cn  
证书编号：S0790522030001

### ● 投资者个股关注度高低具有稳健的负向选股能力

投资者对个股的关注度反映到具体行为上，可以归类为个股点击量、新闻点击量和自选股数量等，值越高表明个股受到投资者的关注程度越高。

点击量占比是投资者个股关注度的真实映射。个股点击量占比因子的 RankIC 历史均值达到-10%，五分组收益虽然不严格单调，但多空分组表现优异，多头年化收益率 11.57%，多空年化收益率 19.83%。

新闻点击量占比是投资者关注度的外部驱动。从测试结果来看，新闻点击量占比因子 RankIC 均值为-6.10%，五分组严格单调，多空收益在 2021 年 3 月后呈现了较大的回撤，多头年化收益率 11.42%，多空年化收益率 9.71%。

自选股票池很大程度上代表了投资者对上市公司的主动关注，而非被动接收，是投资者偏好的内在反映。从测试结果来看，自选股占比因子 RankIC 均值为-7.58%，多头年化收益率为 11.08%，五分组收益虽不完全单调，但多空收益更加稳定，年化收益率 16.51%，夏普比率 1.94。

自选股中关注度变化较低的个股未来表现具有相对优势。这个结果的含义是，影响个股未来表现的，是投资者关注度的高低，而非投资者关注的方向。

### ● 投资者关注度因子间高度相关，但与主流风格因子关注度相对较低

从各投资者关注度因子的绩效表现来看，个股点击量占比因子空头效应最显著，年化收益率为-8.7%，远高于其他因子；自选股占比因子多空对冲稳定性最好，夏普比率最高，达到 1.94；三类因子多头年化收益率近似，表明在多头端大家没有显著的差别。

从关联度来看，个股点击量占比因子、新闻点击量占比因子、自选股占比因子两两之间具有较高的相关性，其中点击量占比和新闻点击量占比因子的相关性更是高达 0.8；而投资者关注度因子与其他风格因子整体关联度较低，最高的为市值因子，平均相关性约 0.49。

### ● 基于投资者关注度的因子增强模型在小市值股票池内效果优异

从测试结果来看，在不同宽基成分股池子中，点击量占比因子均有显著的负向选股能力，其中点击量占比因子在中证 1000 指数成分股中表现最优，RankIC 均值为-0.11，年化 RankICIR 为-3.10。

以权重再平衡方法为例，空头端个股的剔除对不同宽基指数的增强效果不一。沪深 300 指数表现虽然有所提升，但超额收益并不稳定且幅度较小，测试期内累积超额收益仅 14%；中证 500 指数收益增强效果相对稳健，但幅度相对较小，年化超额收益率仅为 3.1%；在中证 1000 指数成分股中做尾部个股剔除，效果优异，测试期内累积超额收益率约 32%，年化超额收益率为 4.7%，超额收益最大回撤仅为 4.4%。

● 风险提示：模型测试基于历史数据，市场未来可能发生变化。

## 目 录

1、 投资者关注度因子测试：具有稳定的负向选股能力 .....	4
1.1、 个股点击量占比：关注度的综合映射 .....	4
1.2、 新闻点击量占比：外驱关注度的度量 .....	7
1.3、 自选股占比：内生关注度的度量 .....	8
1.4、 自选股占比变化：内生关注度的边际变动 .....	9
2、 投资者关注度因子比较：个股点击量占比因子相对占优 .....	10
2.1、 因子间绩效对比：不同因子表现差异主要源于空头端 .....	10
2.2、 因子间关联度：与主流大类风格因子呈现低关联度 .....	11
3、 投资者关注度因子实践：空头个股剔除能有效增强收益 .....	11
3.1、 在主流宽基指数中的表现：在中证 1000 指数中表现占优 .....	11
3.2、 指数增强：基于空头个股剔除的负向因子增强 .....	12
3.2.1、 权重再平衡 .....	12
3.2.2、 约束条件下的优化求解 .....	13
4、 风险提示 .....	14

## 图表目录

图 1： 个股单日最大点击量占比最高可达 8% .....	4
图 2： 原始点击量占比数据呈现右偏分布 .....	5
图 3： 不同行业点击量占比存在较大的差异 .....	5
图 4： 不同市值分组点击量占比在时序上变化相对稳定 .....	6
图 5： 市值越高，点击量占比越高（单位：亿元） .....	6
图 6： 高点击量占比个股潜在回报显著为负 .....	6
图 7： 高点击量占比个股分组大幅跑输其他分组 .....	6
图 8： 低资金用户的高点击量占比的个股未来表现相对更差 .....	7
图 9： 新闻点击量占比越高，潜在回报越低 .....	7
图 10： 新闻点击量占比因子组间区分度单调 .....	7
图 11： 自选股占比越高的个股潜在回报亦为负值 .....	8
图 12： 分组收益不完全单调，但超额收益相对稳定 .....	8
图 13： 低资金用户自选股占比越高的个股，潜在回报越低 .....	8
图 14： 自选股占比差分后的分布呈现尖峰现象 .....	9
图 15： 高新增比高剔除操作负向效应更显著 .....	9
图 16： 自选股占比的变化率呈现尖峰厚尾分布 .....	9
图 17： 自选股占比变化下高新增操作负向效应最强 .....	9
图 18： 个股点击量占比因子在预测能力和稳定性上相对占优 .....	10
图 19： 个股点击量占比和自选股占比因子具有高关联度 .....	11
图 20： 个股点击量占比因子在不同股票池表现不同 .....	12
图 21： 个股点击量占比因子在中证 1000 中表现最优 .....	12
图 22： 沪深 300 指数增强组合超额收益不稳定 .....	13
图 23： 中证 500 指数增强组合年化超额收益约 3.1% .....	13
图 24： 中证 1000 指数增强组合具有稳健超额收益，最大回撤仅 4.4% .....	13
图 25： 沪深 300 指数增强组合超额收益波动较大 .....	14
图 26： 中证 1000 指数增强组合累积超额收益达 50% .....	14

表 1: 投资者关注度因子均表现出稳定的负向选股能力.....	10
---------------------------------	----

随着移动互联网的普及，手机端交易软件的推出大幅提升了投资者的交易便捷性。区别于机构投资者在 Wind 终端具有相对数量优势，同花顺和东方财富等第三方交易软件提供商则更多吸引了个人投资者的使用。

在大数据时代，数据即是价值。交易软件记录的海量投资者的交易行为，是否蕴含着丰富的  $\alpha$  信息源，是本篇报告的关注重点。我们尝试从交易软件记录的各类投资者行为中挖掘出潜藏的特征，以作为未来投资的参考依据。

以同花顺交易软件为例，其通过对原始数据进行脱敏处理后，统计了几大类典型的投资者行为，包括但不限于投资者个股点击量、新闻点击量和自选股数量等多种行为。同时，为了区分不同资金量级下的投资者行为之间是否存在显著差异，同花顺也进行了相应的切分处理，为我们更细致地审视投资者间的差异提供了更加清晰的视角。

全文总共分为三个部分，第一部分我们根据投资者的行为特征，构建了三大类投资者关注度因子，并测试了各关注度因子的表现，发现投资者关注度因子普遍具有稳定的负向选股能力；在第二部分，我们对比了各关注度因子间以及和主流风格因子的关联度，从结果来看，投资者关注度因子间具有较高的关联度，可以理解为是投资者同一类行为特征的不同表述方式，但与主流风格因子，整体关联度相对较低；最后我们构建了以投资者关注度特征为  $\alpha$  源的指数增强模型，并以表现相对占优的个股点击量占比因子作为示例进行说明。

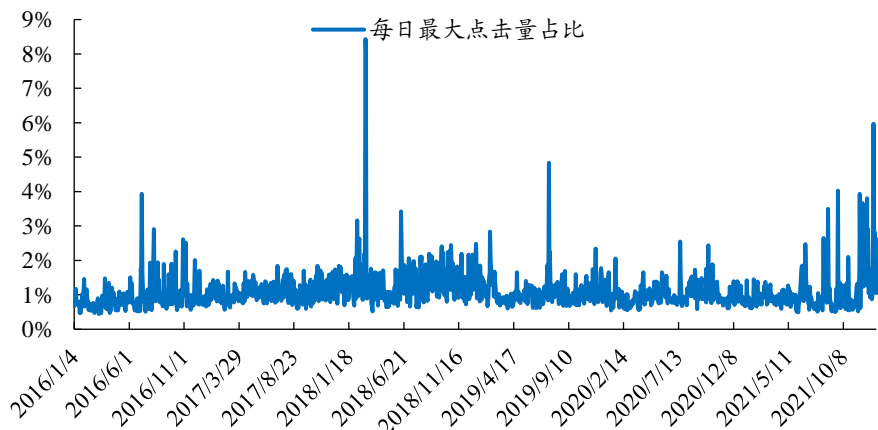
## 1、投资者关注度因子测试：具有稳定的负向选股能力

### 1.1、个股点击量占比：关注度的综合映射

个股点击量能够较好地反映市场上个股当前时刻的受关注程度，当前受关注程度相对更高的个股通常具有更高的点击次数，其可以理解为投资者对个股关注程度的映射。以此为切入点，我们构建了个股点击量占比因子，其计算方法为个股当天点击量占全市场所有个股当天点击量的比例。

通过观察每个交易日个股点击量占比的最大值时序图可知，单个股票最大点击量占比的平均值约为 1.1%，极值能够达到全市场所有个股点击量的 8% 以上。

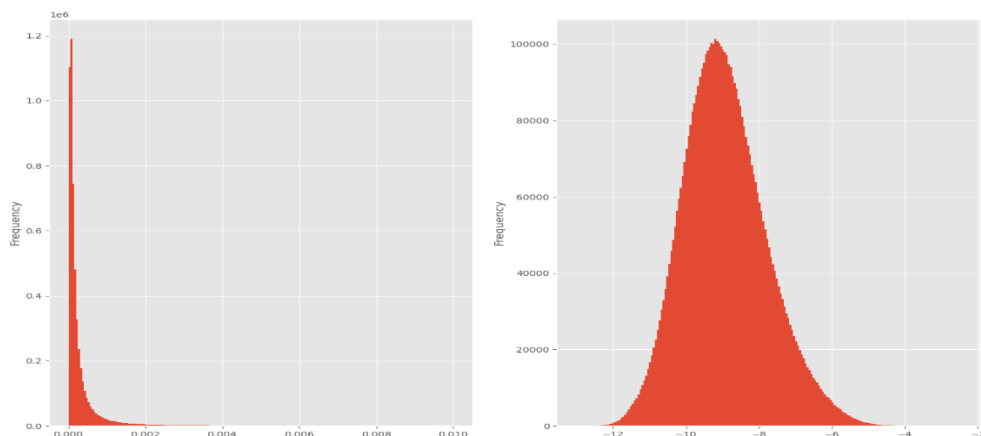
图1：个股单日最大点击量占比最高可达 8%



数据来源：同花顺 iFind、开源证券研究所

考虑到全市场几千支股票，个股点击量占比数据大概率呈现厚尾形态。为了验证猜测，我们绘制了点击量占比数据的分布图，其呈现显著的右偏分布（为了方便显示，我们对个股点击量占比超过历史上单日最大点击量均值的数值予以剔除）。为此，我们尝试对原始数据进行取对数处理，可以看到转换后的新数据近似服从正态分布。

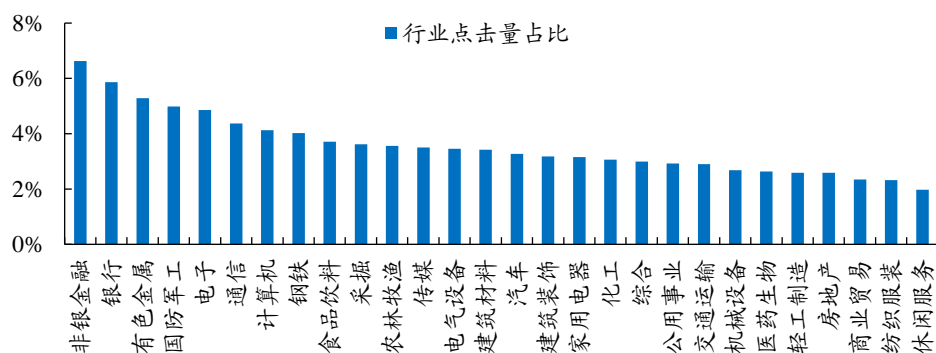
图2：原始点击量占比数据呈现右偏分布



数据来源：同花顺 iFind、开源证券研究所

从行业点击量占比分布来看，各行业之间亦存在较明显的差异。其中，行业点击量占比计算过程为：首先计算每个月末行业内个股点击量占比的均值，再对时序上的行业均值取平均值处理，最后在截面上进行归一化，得到各行业在全区间内的点击量占比情况。可以看到，非银金融、银行、有色金属和国防军工行业点击量占比相对较高，达到5%以上，商业贸易、纺织服装和休闲服务点击量占比则相对靠后，均低于2.5%。

图3：不同行业点击量占比存在较大的差异

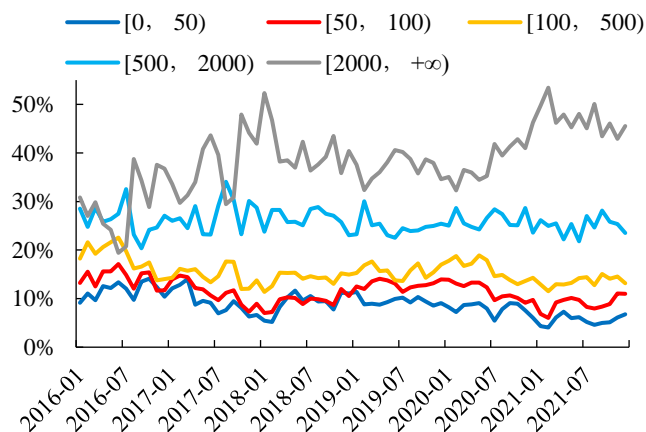


数据来源：同花顺 iFind、开源证券研究所

我们按照 50 亿、100 亿、500 亿和 2000 亿将个股市值分为从低到高五个组别，并根据行业分组下点击量占比的计算方式统计了市值分组下点击量占比的情况。可以看到在时序上不同市值分组下点击量占比处于相对稳定的位置，高市值分组点击量占比相对更高，其中市值规模在 500 亿以上的个股被点击的概率显著高于其他分组。数据背后的直观解释是，大市值股票通常被投资者更熟知，被点击的概率更大，小市值股票通常受限于全 A 几千支标的，投资者精力有限，导致被遗忘在角落里的概率更高。

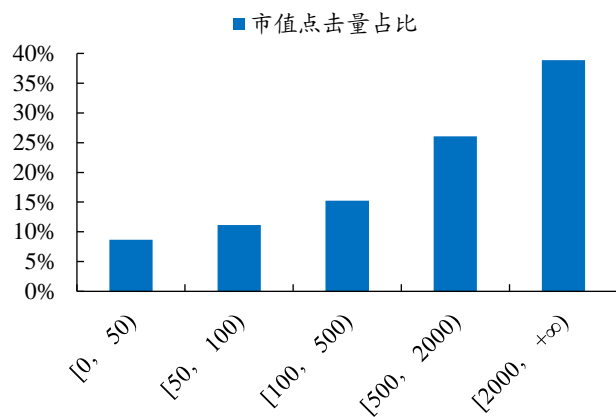


图4：不同市值分组点击量占比在时序上变化相对稳定



数据来源：同花顺 iFind、开源证券研究所

图5：市值越高，点击量占比越高（单位：亿元）

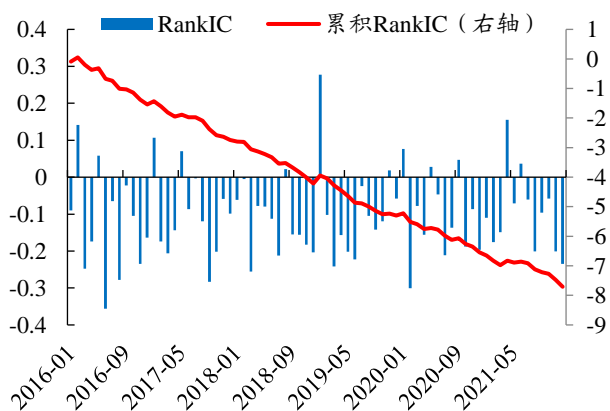


数据来源：同花顺 iFind、开源证券研究所

为了避免选择偏差，我们首先将自然月内的个股点击量占比取均值，然后在此基础上进行取对数处理，得到每个月末个股点击量占比。根据前文分析可知个股点击量占比在不同行业和市值分组下存在较大差异，因此我们在进行因子测试前，分别进行市值和行业中性化处理，同时为了避免次新股和新股的影响，我们剔除了上市不足一年的个股。如无特别说明，以下因子测试均采用此流程。

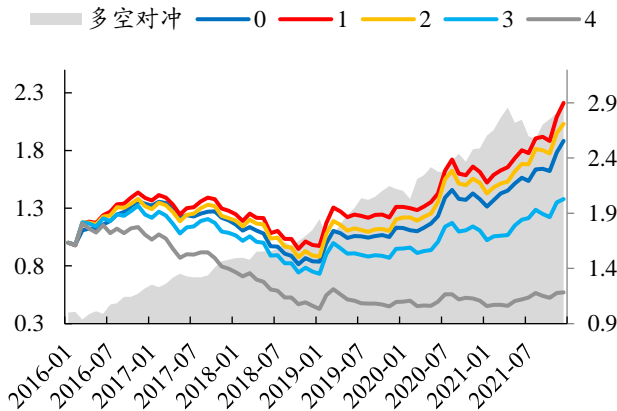
谚语有云，“人多的地方不要去”。对于个股而言，高点击量占比是否意味着个股的市场关注度过高，投资者情绪达到极端值，导致相关个股未来的潜在回报为负？从测试结果来看，符合前述猜测，即高点击量占比个股所在的分组通常未来的收益表现相对不佳。点击量占比因子的 RankIC 历史均值达到-10%，五分组收益虽然不严格单调，但多空分组表现优异，其中多头年化收益率 11.57%，多空年化收益率 19.83%。

图6：高点击量占比个股潜在回报显著为负



数据来源：同花顺 iFind、开源证券研究所

图7：高点击量占比个股分组大幅跑输其他分组

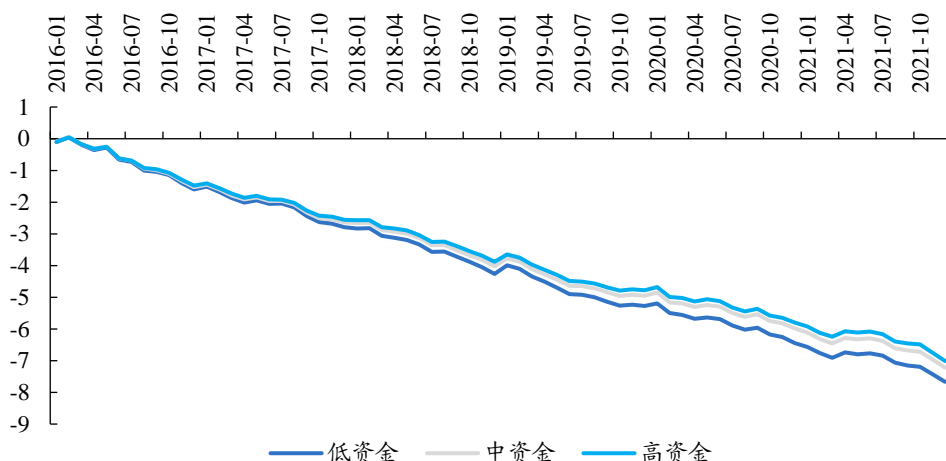


数据来源：同花顺 iFind、开源证券研究所

更进一步，高资金量级投资者的表现是否更加理性？其个股点击分布是否与低资金量级的投资者存在显著差别？为此，我们按照用户资金量的大小，将投资者分为低、中、高三种类型，其中低资金投资者持仓规模为 10~50 万元，中资金投资者规模为 50~100 万元，高资金投资者规模为 100 万元以上。

我们分别统计了三种类型投资者的累计 RankIC 变化情况，从测试结果来看，高中低三类投资者的个股点击量占比表现具有一致性，均呈现稳定的负向选股能力，但低资金投资者点击量占比高的个股未来表现相对更差，即负向指示效果更显著。

图8：低资金用户的高点击量占比的个股未来表现相对更差



数据来源：同花顺 iFind、开源证券研究所

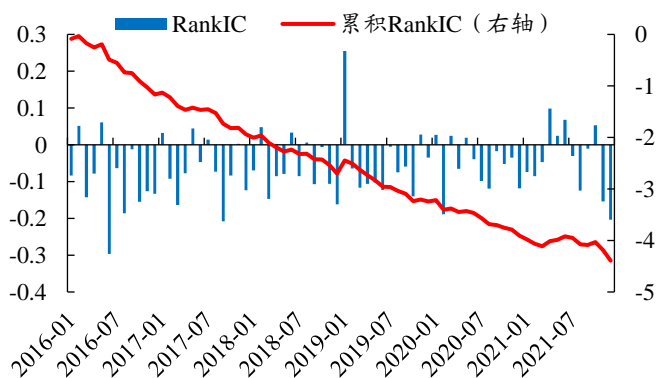
前文，我们测试了个股点击量占比因子的表现，个股点击触发通常源于三个途径：第一，被点击个股源于行情软件对相关新闻消息的推送，投资者出于好奇心所引发的点击；第二，被点击个股本身处于自选股池，其点击触发源于投资者对个股每日的例行跟踪；第三，被点击个股既处于自选股池，又被行情软件推送相关新闻，进而引发投资者的点击操作。根据前两个影响因素，我们分别构建了新闻点击量占比和自选股占比因子。那么，这两类因子是否潜藏着收益差别呢？

## 1.2、新闻点击量占比：外驱关注度的度量

资本市场中每天充斥着上市公司各种新闻消息，既涵盖上市公司官方发布的公告，也包括财经媒体发布的关于上市公司的各种信息。在浩如烟海的新闻信息中，部分上市公司的相关新闻信息曝光度更高，更容易受到投资者的关注，被投资者点击的概率更大，但投资者点击率更高的新闻对应的个股未来表现会更好吗？

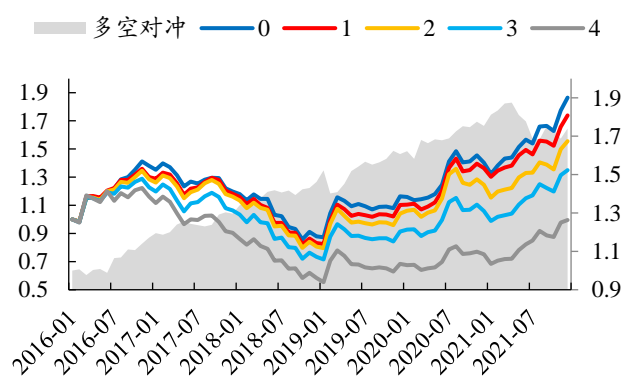
基于此，我们构建了新闻点击量占比因子，其计算方法为个股新闻点击量除以A股所有上市公司新闻点击量合计值。从测试结果来看，新闻点击量占比因子 RankIC 均值为-6.10%，五分组严格单调，多空收益在2021年3月后呈现了较大的回撤，多头年化收益率11.42%，多空年化收益率9.71%。

图9：新闻点击量占比越高，潜在回报越低



数据来源：同花顺 iFind、开源证券研究所

图10：新闻点击量占比因子组间区分度单调



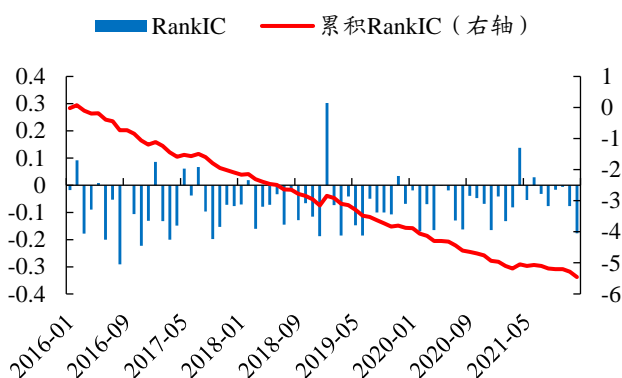
数据来源：同花顺 iFind、开源证券研究所

### 1.3、自选股占比：内生关注度的度量

自选股票池方便了投资者对关注个股的持续跟踪，其很大程度上代表了投资者对上市公司的主动关注，而非被动接收，是投资者偏好的内在反映。当某个标的股票被多数投资者加入自选股票池时，个股未来表现又会如何呢？为此，我们构建了自选股占比因子，其计算逻辑为：个股自选股用户数 / 自选股用户数合计（注：当同一个用户自选股中有多个股票时，每个股票都会统计该用户）。

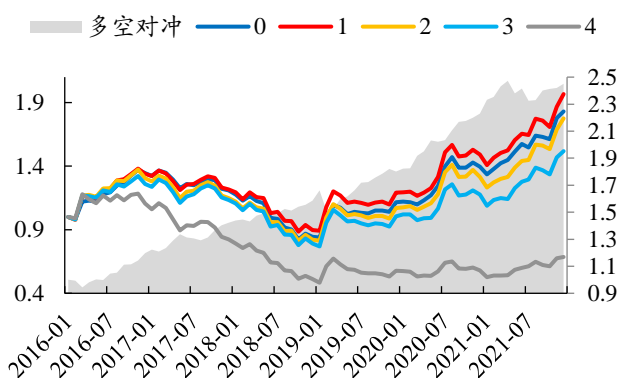
从测试结果来看，自选股占比因子行为特征与个股点击量占比和新闻点击量占比因子表现近似，均呈现出显著的负向选股能力，其 RankIC 均值为-7.58%，多头年化收益率为 11.08%，五分组收益虽不完全单调，但多空收益更加稳定，年化收益率 16.51%，夏普比率 1.94。

图11：自选股占比越高的个股潜在回报亦为负值



数据来源：同花顺 iFind、开源证券研究所

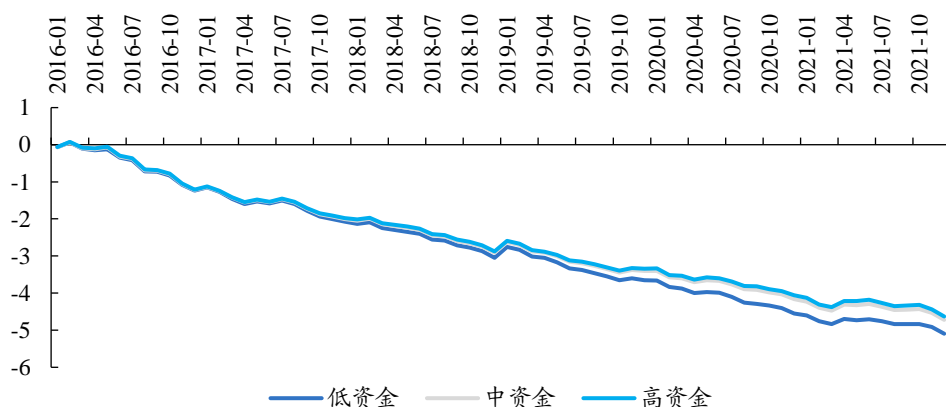
图12：分组收益不完全单调，但超额收益相对稳定



数据来源：同花顺 iFind、开源证券研究所

类比点击量占比因子，我们亦根据用户资金量大小将自选股占比因子分为低、中、高三组，分别进行测试，观察不同资金量的用户其自选股占比是否蕴含着不同的信息。从累计 RankIC 的结果来看，低资金用户的自选股池代表的信息更偏负面，符合预期。

图13：低资金用户自选股占比越高的个股，潜在回报越低



数据来源：同花顺 iFind、开源证券研究所



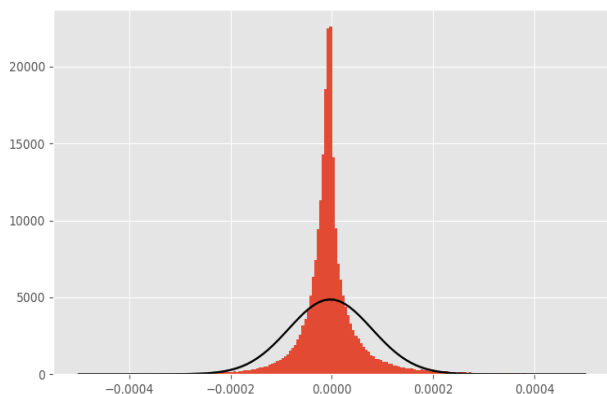
## 1.4、自选股占比变化：内生关注度的边际变动

自选股票池是一个偏静态的动作，新增或剔除操作或许更能反映投资者对个股的偏好变化。新增或剔除操作是一体两面的关系，若个股当前自选股占比高于前序自选股占比，则认为是新增操作，反之，则认为是剔除操作。为了计算新增或剔除操作的表现，我们对原始的自选股占比因子进行差分 and 取变化率处理。

做差分处理时，将当前月份自选股占比均值减去上个月份自选股占比均值；计算变化率时，我们将当前月份自选股占比均值除以上个月份自选股占比均值后减去 1。注意，这里没有考虑个股自选股新增或剔除绝对数量之间的前后差异，仅从相对维度来切入的。当自选股占比差分或变动数值高于零值时，认为是新增操作，反之，则归类到剔除操作。

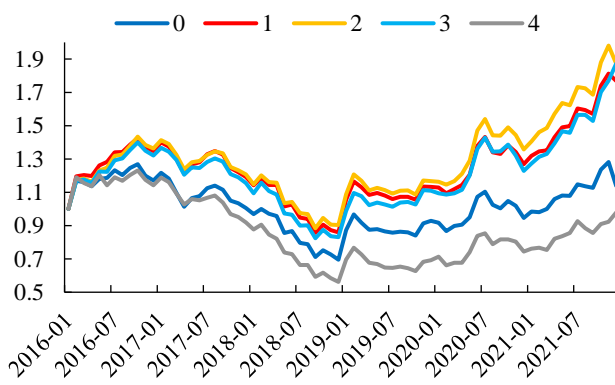
自选股占比因子进行差分操作后，分布呈现尖峰形态。在分组收益层面，可以看到高新增或剔除自选股的操作均呈现负向选股效应，其中高新增操作负向效果更显著。

图14：自选股占比差分后的分布呈现尖峰现象



数据来源：同花顺 iFind、开源证券研究所

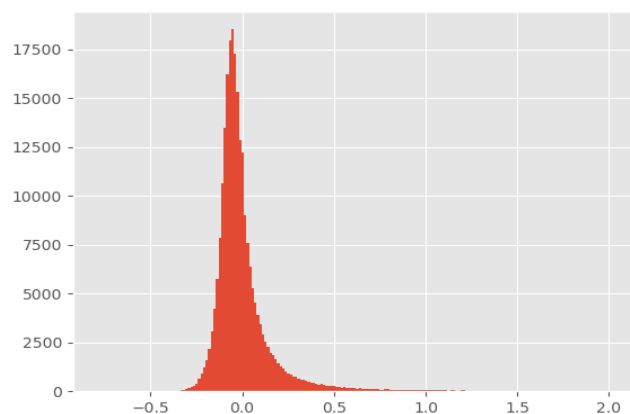
图15：高新增比高剔除操作负向效应更显著



数据来源：同花顺 iFind、开源证券研究所

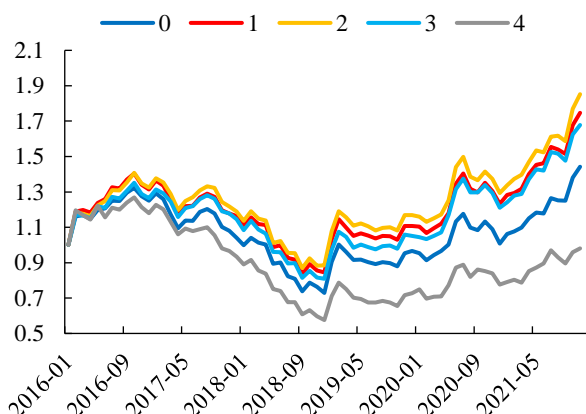
自选股占比因子进行变化率计算时，自选股占比的变动值为正（新增）或为负（剔除），亦均是偏负向的效应，两种操作对应的分组的累计收益率均低于自选股占比变动量绝对值较小的分组。另一方面，高新增操作比高剔除操作，负向效用更加显著。

图16：自选股占比的变化率呈现尖峰厚尾分布



数据来源：同花顺 iFind、开源证券研究所

图17：自选股占比变化下高新增操作负向效应最强



数据来源：同花顺 iFind、开源证券研究所

从两种构建方法可以看到，自选股中关注度变化较低的个股未来表现具有相对优势。这个结果的含义是，影响个股未来表现的，是投资者关注度的高低，而非投资者关注度的方向。

## 2、投资者关注度因子比较：个股点击量占比因子相对占优

### 2.1、因子间绩效对比：不同因子表现差异主要源于空头端

从各投资者关注度因子的绩效表现来看，个股点击量占比因子空头效应最显著，年化收益率为-8.7%，远高于其他因子；自选股占比因子多空对冲稳定性最好，夏普比率最高，达到 1.94；三类因子多头年化收益率近似，表明在多头端大家没有显著的差别。

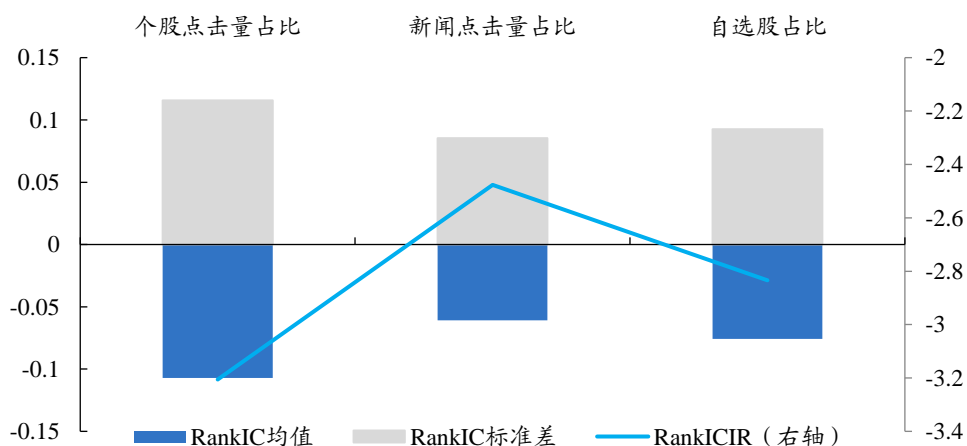
表1：投资者关注度因子均表现出稳定的负向选股能力

指标名称	个股点击量占比			新闻点击量占比			自选股占比		
组别	多头	空头	对冲	多头	空头	对冲	多头	空头	对冲
年化收益率	11.6%	-8.7%	19.8%	11.4%	-0.1%	9.7%	11.1%	-6.0%	16.5%
年化波动率	17.0%	23.5%	11.2%	18.1%	22.9%	8.5%	17.0%	22.5%	8.5%
夏普比率	0.68	-0.37	1.76	0.63	0.01	1.14	0.65	-0.27	1.94
最大回撤	-40.2%	-63.4%	-9.8%	-39.2%	-54.7%	-10.2%	-40.0%	-59.2%	-8.8%
胜率	50.7%	43.7%	73.2%	52.1%	47.9%	64.8%	47.9%	38.0%	78.9%

数据来源：同花顺 iFind、开源证券研究所

从未来收益预测能力来看，个股点击量占比因子表现相对占优，RankIC 均值的绝对值超过 10%，年化 RankICIR 达到-3.2，因子表现十分稳定；自选股占比因子，RankIC 均值的绝对值亦达到 8%左右的水平。新闻点击量占比因子预测能力和因子稳定性均相对较弱。

图18：个股点击量占比因子在预测能力和稳定性上相对占优



数据来源：同花顺 iFind、开源证券研究所

## 2.2、因子间关联度：与主流大类风格因子呈现低关联度

前文中我们测试了不同投资者关注度因子的表现，从测试结果来看，投资者在个股上的相关行为整体呈现非常显著的负向选股效果。为了比较各因子之间的异同，我们计算了点击量占比因子、新闻点击量占比因子和自选股占比因子和主流风险因子之间的关联度。

图19：个股点击量占比和自选股占比因子具有高关联度

因子名称	个股点击量占比	新闻点击量占比	自选股占比
个股点击量占比	1.00	0.59	0.80
新闻点击量占比	0.59	1.00	0.61
自选股占比	0.80	0.61	1.00
残差波动率	0.39	0.18	0.20
动量	0.30	0.23	0.25
盈利	-0.03	0.02	-0.02
成长	0.05	0.05	-0.03
流动性	0.47	0.23	0.41
市值	0.46	0.50	0.51
杠杆	0.03	0.08	0.11
估值	-0.11	-0.05	-0.05
Beta	0.36	0.24	0.33

数据来源：同花顺 iFind、开源证券研究所

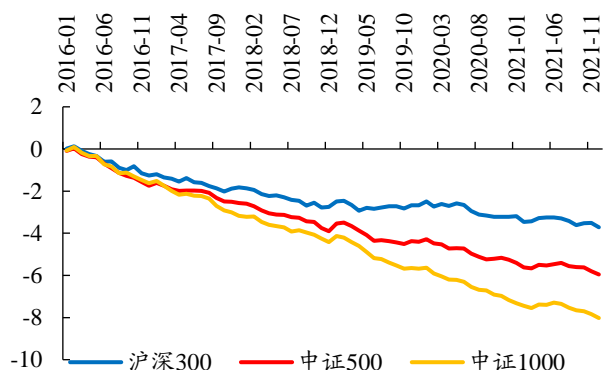
从关联度来看，个股点击量占比因子、新闻点击量占比因子、自选股占比因子两两之间具有较高的相关性，其中个股点击量占比和新闻点击量占比因子的相关性更是高达 0.8，可以推断大部分的个股点击操作来源于投资者的自选股，而非新闻消息的推送；而投资者关注度因子与其他风格因子整体关联度较低，最高的为市值因子，平均相关性约 0.49。

## 3、投资者关注度因子实践：空头个股剔除能有效增强收益

### 3.1、在主流宽基指数中的表现：在中证 1000 指数中表现占优

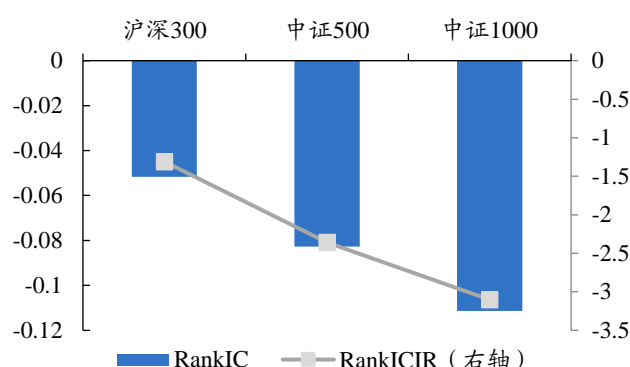
以上各投资行为因子的测试是在全市场股票池内进行，这里我们以表现相对占优的个股点击量占比因子为例，尝试在主流宽基成分股中进行效果测试。从测试结果来看，在不同宽基成分股池中，个股点击量占比因子均有显著的负向选股能力。其中，个股点击量占比因子在中证 1000 指数成分股中表现最优，RankIC 均值为-0.11，年化 RankICIR 为-3.10，其次为中证 500 指数，最次为沪深 300 指数。

图20: 个股点击量占比因子在不同股票池表现不同



数据来源: 同花顺 iFind、开源证券研究所

图21: 个股点击量占比因子在中证 1000 中表现最优



数据来源: 同花顺 iFind、开源证券研究所

### 3.2、指数增强：基于空头个股剔除的负向因子增强

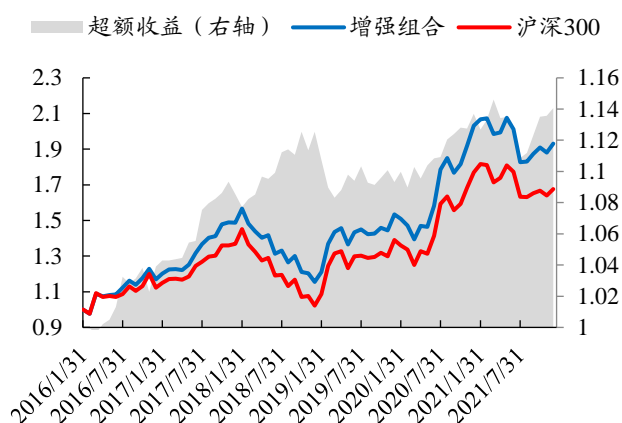
考虑到投资者关注度因子的强预测能力，我们希望能将其进行投资落地，一个很朴素的想法便是应用于指数增强的框架下，获取超额稳健回报。但问题是，以上投资者关注度因子的选股能力主要体现在空头端，多头超额收益获取能力相对较弱且稳定性不足。若直接按照多头端因子构建增强组合的方式进行处理，可以预见的是效果并不出众。为了尽可能保留空头端蕴含的稳健信息，我们在构建增强模型前对股票池个股进行筛选，即对空头端个股预先进行剔除处理。

#### 3.2.1、权重再平衡

由前文测试结果可知，个股点击量占比因子的空头组（第五组）具有显著的负向选股能力，我们希望将这些表现低于市场均值的个股从样本池内剔除。具体而言，在每期调仓之际，我们在基准指数成分股内剔除掉空头端分组对应的个股，由于部分个股从基准指数成分股中剔除，导致当期指数权重之和不为 1，我们对剩余个股做权重再平衡处理。我们亦尝试过在全市场股票池内先做尾部剔除，再与基准指数成分股取交集，最后进行权重再平衡，发现测试结果表现相对更优，有兴趣的读者可以尝试一下。

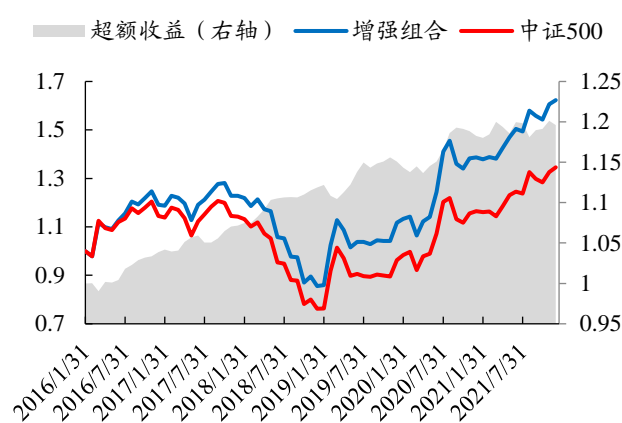
从事前剔除效果来看，空头端个股的剔除对不同宽基指数的增强效果不一。沪深 300 指数表现虽然有所提升，但超额收益并不稳定且幅度较小，测试期内累积超额收益仅 14%；剔除空头端个股后，中证 500 指数收益增强效果相对稳健，但幅度相对较小，年化超额收益率仅为 3.1%。

图22: 沪深 300 指数增强组合超额收益不稳定



数据来源：同花顺 iFind、开源证券研究所

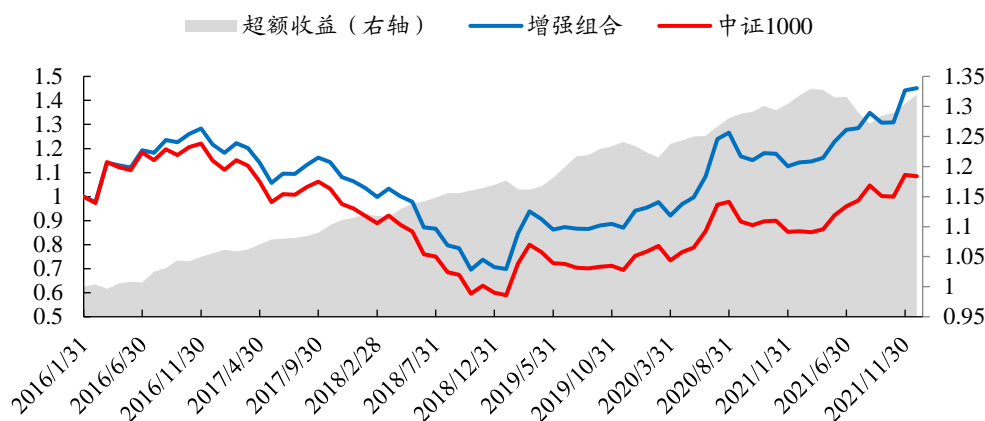
图23: 中证 500 指数增强组合年化超额收益约 3.1%



数据来源：同花顺 iFind、开源证券研究所

在中证 1000 指数成分股中做尾部个股剔除，效果优异，测试期内累积超额收益率约 32%，年化超额收益率为 4.7%，超额收益最大回撤仅为 4.4%。

图24: 中证 1000 指数增强组合具有稳健超额收益，最大回撤仅 4.4%



数据来源：同花顺 iFind、开源证券研究所

### 3.2.2、约束条件下的优化求解

剔除空头端个股再进行权重再平衡的方法虽然能够实现对原始指数的跟踪，但当所剔除的个股在风格或行业上分布不均匀的时候，会导致权重再平衡的方法无法有效控制暴露度的偏离，为此我们尝试在设定约束条件的基础上来进行最优权重求解。

与权重再平衡在指数成分股内进行不同，优化模型的选股在全市场股票池内，但会控制成分股在最终持仓中的权重占比。为了充分利用个股点击量占比因子的稳健负向选股信息，我们在优化求解前亦预先剔除空头端个股，在此基础上再进行最优权重求解。由于个股点击量占比因子表现为负向选股能力，所以因子值越低对应未来预期收益率越高，最大化预期收益率的目标函数可以转化为最小化因子暴露度。



$$\min \alpha^T w$$

$$s_l \leq X_s(w - w_b) \leq s_u$$

$$h_l \leq X_h(w - w_b) \leq h_u$$

$$w_l \leq w - w_b \leq w_u$$

$$b^T w \geq c_l$$

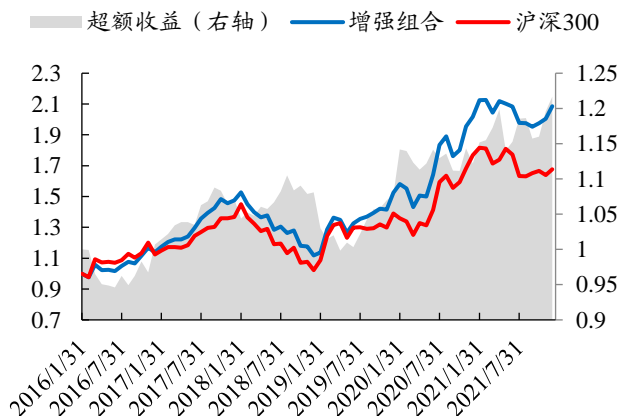
$$1^T w = 1$$

$$0 \leq w \leq 1$$

其中， $\alpha$ 表示因子暴露度， $w$ 表示待优化权重， $X_s$ 表示风格暴露度矩阵， $X_h$ 表示行业哑变量矩阵， $s_u$ 和 $s_l$ 分别表示风格暴露度的偏离上下限， $h_u$ 和 $h_l$ 分别表示行业偏离上下限， $w_u$ 和 $w_l$ 分别表示权重偏离的上下限， $b$ 表示股票池个股是否位于指数成分的示性向量， $c_l$ 表示成分股权重之和下限。具体而言，约束条件的参数 $s_l, h_l$ 设置为-0.01， $s_u, h_u$ 设置为0.01， $c_l$ 设置为0.8，为了避免无解， $w_l, w_u$ 分别设置为-0.05, 0.05。需要注意的是，在求解过程中，如果分组数过少，对特定基准指数可能存在无解的情况，这个时候可以适当增加分组数量，比如从常见的五分组转变为更多分组。

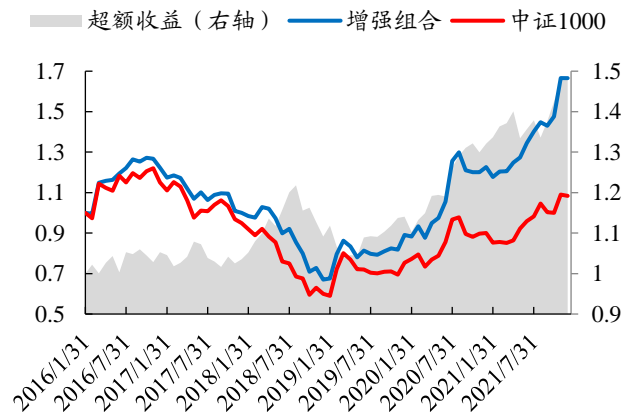
从优化效果来看，在测试期内，沪深300指数增强组合累积超额收益约20%，相比权重再平衡策略有所提升，但幅度有限，超额收益波动依然较大。测试期内，中证1000指数增强组合累积超额收益率达到50%，高出权重再平衡策略约17%，年化超额收益率为6.8%，但超额收益率最大回撤达到14.6%，高于权重再平衡策略。中证500指数增强效果居于二者之间，此处未予展示。

图25：沪深300指数增强组合超额收益波动较大



数据来源：同花顺 iFind、开源证券研究所

图26：中证1000指数增强组合累积超额收益达50%



数据来源：同花顺 iFind、开源证券研究所

#### 4、风险提示

模型测试基于历史数据，市场未来可能发生变化。

## 特别声明

《证券期货投资者适当性管理办法》、《证券经营机构投资者适当性管理实施指引（试行）》已于2017年7月1日起正式实施。根据上述规定，开源证券评定此研报的风险等级为R3（中风险），因此通过公共平台推送的研报其适用的投资者类别仅限定为专业投资者及风险承受能力为C3、C4、C5的普通投资者。若您并非专业投资者及风险承受能力为C3、C4、C5的普通投资者，请取消阅读，请勿收藏、接收或使用本研报中的任何信息。因此受限于访问权限的设置，若给您造成不便，烦请见谅！感谢您给予的理解与配合。

## 分析师承诺

负责准备本报告以及撰写本报告的所有研究分析师或工作人员在此保证，本研究报告中关于任何发行商或证券所发表的观点均如实反映分析人员的个人观点。负责准备本报告的分析师获取报酬的评判因素包括研究的质量和准确性、客户的反馈、竞争性因素以及开源证券股份有限公司的整体收益。所有研究分析师或工作人员保证他们报酬的任何一部分不曾与，不与，也将不会与本报告中具体的推荐意见或观点有直接或间接的联系。

## 股票投资评级说明

	评级	说明
证券评级	买入（Buy）	预计相对强于市场表现 20%以上；
	增持（outperform）	预计相对强于市场表现 5% ~ 20%；
	中性（Neutral）	预计相对市场表现在 - 5% ~ + 5%之间波动；
	减持（underperform）	预计相对弱于市场表现 5%以下。
行业评级	看好（overweight）	预计行业超越整体市场表现；
	中性（Neutral）	预计行业与整体市场表现基本持平；
	看淡（underperform）	预计行业弱于整体市场表现。
备注：评级标准为以报告日后的 6~12 个月内，证券相对于市场基准指数的涨跌幅表现，其中 A 股基准指数为沪深 300 指数、港股基准指数为恒生指数、新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）、美股基准指数为标普 500 或纳斯达克综合指数。我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议；投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者应阅读整篇报告，以获取比较完整的观点与信息，不应仅仅依靠投资评级来推断结论。		

## 分析、估值方法的局限性说明

本报告所包含的分析基于各种假设，不同假设可能导致分析结果出现重大不同。本报告采用的各种估值方法及模型均有其局限性，估值结果不保证所涉及证券能够在该价格交易。

## 法律声明

开源证券股份有限公司是经中国证监会批准设立的证券经营机构，已具备证券投资咨询业务资格。

本报告仅供开源证券股份有限公司（以下简称“本公司”）的机构或个人客户（以下简称“客户”）使用。本公司不会因接收人收到本报告而视其为客户。本报告是发送给开源证券客户的，属于机密材料，只有开源证券客户才能参考或使用，如接收人并非开源证券客户，请及时退回并删除。

本报告是基于本公司认为可靠的已公开信息，但本公司不保证该等信息的准确性或完整性。本报告所载的资料、工具、意见及推测只提供给客户作参考之用，并非作为或被视为出售或购买证券或其他金融工具的邀请或向人做出邀请。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。客户应当考虑到本公司可能存在可能影响本报告客观性的利益冲突，不应视本报告为做出投资决策的唯一因素。本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。本公司未确保本报告充分考虑到个别客户特殊的投资目标、财务状况或需要。本公司建议客户应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。在任何情况下，本公司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。若本报告的接收人非本公司的客户，应在基于本报告做出任何投资决定或就本报告要求任何解释前咨询独立投资顾问。

本报告可能附带其它网站的地址或超级链接，对于可能涉及的开源证券网站以外的地址或超级链接，开源证券不对其内容负责。本报告提供这些地址或超级链接的目的纯粹是为了客户使用方便，链接网站的内容不构成本报告的任何部分，客户需自行承担浏览这些网站的费用或风险。

开源证券在法律允许的情况下可参与、投资或持有本报告涉及的证券或进行证券交易，或向本报告涉及的公司提供或争取提供包括投资银行业务在内的服务或业务支持。开源证券可能与本报告涉及的公司之间存在业务关系，并无需事先或在获得业务关系后通知客户。

本报告的版权归本公司所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本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。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

## 开源证券研究所

### 上海

地址：上海市浦东新区世纪大道1788号陆家嘴金控广场1号楼10层  
邮编：200120  
邮箱：research@kysec.cn

### 深圳

地址：深圳市福田区金田路2030号卓越世纪中心1号楼45层  
邮编：518000  
邮箱：research@kysec.cn

### 北京

地址：北京市西城区西直门外大街18号金贸大厦C2座16层  
邮编：100044  
邮箱：research@kysec.cn

### 西安

地址：西安市高新区锦业路1号都市之门B座5层  
邮编：710065  
邮箱：research@kysec.cn