

2021 年 7 月 7 日

## 业绩超预期 Plus 组合的构建

——开源量化评论（25）

### 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（研究员）

证书编号：S0790120020012

胡亮勇（研究员）

证书编号：S0790120030040

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（20）-北上资金行业配置的双轮驱动力》-2021.2.24

《开源量化评论（21）-从茅指数动力学研判抱团现状》-2021.3.30

《开源量化评论（22）-年报展望全扫描：基金经理如何看后市》-2021.4.6

《开源量化评论（23）-“金股+”组合的量化方案》-2021.4.26

《开源量化评论（24）-上市公司招聘数据的选股能力》-2021.5.12

魏建榕（分析师）

weijianrong@kysec.cn

证书编号：S0790519120001

胡亮勇（联系人）

huliangyong@kysec.cn

证书编号：S0790120030040

### ● 业绩预告、业绩快报和定期报告构成了投资者了解上市公司的三个窗口

在 2019 年之前，业绩预告披露数量整体呈现波动上升的趋势，后续随着监管机构对业绩预告披露要求的放松，上市公司业绩预告披露数量有一定程度的下降。业绩快报披露主要集中在中报和年报两个时期，其中又以年报的业绩快报为主。随着上市公司数量不断增加，且退市节奏相对较缓慢，定期报告披露数量逐年稳定上升。

沪深 300 指数成分股分析师覆盖度持续保持在 90% 以上的水平；中证 500 指数成分股分析师覆盖度在 70%~90% 之间浮动；随着注册制的实施，公司上市节奏加快，全市场个股的分析师覆盖率呈现了较大幅度的下滑，从 2016 年 Q3 的 75% 以上的覆盖率水平逐步下降到 2021 年 Q1 不足 50% 的覆盖率。

财务报告发布前后的事件收益中，披露时间最早的业绩预告未来超额收益表现最好，其次是业绩快报，公布时间最靠后的定期报告未来超额收益水平最低。

### ● 预期外盈余（SUE）因子和预期外百分比（PCT）因子均具有良好的选股表现

多报告融合后的 SUE 因子表现相对业绩预告 SUE 因子，虽然在多头端收益有了一定的下降，但是在分组收益的区分度上有了提升；从多空净值来看，相比定期报告亦出现明显改善。在测试区间期间，多报告融合后的多头年化收益率为 21.56%，夏普比率 0.72，多空对冲年化收益率为 14.15%，夏普比率为 3.46。

多报告融合后的业绩预期外百分比因子分组收益虽然整体单调，但分组四和分组五比较靠近，多头年化收益率 17.23%，夏普比率 0.72，多空对冲年化收益率 14.10%，夏普比率 2.68。

### ● 超预期股票池内收益增强大有可为

在测试期间，超预期个股在分析师覆盖股票池中占比大约处于 30%~40% 的区间，2020 年第一季度由于受新冠疫情的影响，超预期个股占比下滑到 20% 左右水平，受益于国内疫情率先控制，2021 年一季度在 2020 年低基数的基础上业绩超预期个股占比大幅上升到 60% 的水平。

超预期股票池内的选股，我们从多个不同的角度展开测试，包括基于财务数据的标准化预期外因子、基于价格数据的区间收益和跳空收益、基于交易行为的开源金工独家因子，以及基于资金流的大小单残差。结果表明，以上所选因子在超预期股票池内均具有良好的区分度。

基于不同维度等权合成得到的最终选股因子，选取其中得分最高的前 N 只股票构成投资组合，测试期内表现出了强劲的超额收益能力。投资组合持仓个股越高，超额收益水平相对越低。

选取合成因子得分最高的 30 只股票构建的超预期 30 组合，在测试期内多头收益率高达 43.13%，夏普比率为 1.53，月度胜率约 70%；对冲基准收益为 31.08%，夏普比率为 3.6，月度胜率超 80%。

● 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、 数据概览 .....	4
1.1、 财报数据：披露的节奏与数量 .....	4
1.2、 分析师一致预期数据：覆盖率与 alpha 特征 .....	4
2、 数据预处理 .....	5
2.1、 数据口径一致化 .....	6
2.2、 因子数据标准化 .....	6
2.3、 调仓日期定期化 .....	6
2.4、 避免前视偏差 .....	7
3、 标准化预期外盈余因子（SUE） .....	7
3.1、 业绩预告：时效性最高的业绩预告选股效果优异 .....	7
3.2、 定期报告：内容披露最完整的定期报告选股表现稍逊一筹 .....	8
3.3、 多报告融合：兼顾时效与完整性的多报告对冲表现改善显著 .....	8
4、 预期外百分比因子（PCT） .....	9
5、 超预期股票池的增强探索 .....	10
5.1、 标准化预期外因子 .....	11
5.2、 公布日前后的区间收益 .....	12
5.3、 超预期跳空 .....	13
5.4、 交易行为因子 .....	14
5.5、 大小单残差因子 .....	15
5.6、 国企与非国企 .....	15
6、 超预期 Plus 组合构建 .....	16
7、 附录：分析师一致预期数据调整方式对比 .....	18
7.1、 年度数据转成逐季累加模式 .....	19
7.2、 年度数据转成单季度模式 .....	19
7.3、 数据转换实例 .....	19
8、 风险提示 .....	20

## 图表目录

图 1： 业绩预告、业绩快报和定期报告每期财报发布数量差异显著 .....	4
图 2： 分析师一致预期归母净利润个股覆盖率 2017 年以来下滑明显 .....	5
图 3： 业绩预告发布前后的累积超额收益表现最佳 .....	5
图 4： 业绩预告 SUE 因子表现亮眼，多头收益显著 .....	8
图 5： 定期报告 SUE 因子分组收益单调，多空对冲走势稳健 .....	8
图 6： 多报告融合 SUE 因子表现亮眼，多空对冲夏普比率高达 3.46 .....	9
图 7： 业绩预告 PCT 因子分层效果优异 .....	10
图 8： 定期报告 PCT 因子多头组区分度有所下滑 .....	10
图 9： 多报告融合 PCT 因子多空对冲夏普比率为 2.68 .....	10
图 10： 超预期个股占比在 2020 年三季度以来增长比较明显 .....	11
图 11： 超预期样本池内个股表现优于基准组合中证 500 指数 .....	11
图 12： SUE 因子在超预期股票池内依然具有分层效果 .....	12
图 13： PCT 因子在超预期股票池内多头组分层效果下滑 .....	12
图 14： 超预期收益（OER）因子短区间呈现动量效应，长区间呈现反转效应 .....	12

图 15:	2017 年后，OER 因子在超预期股票池中具有稳定超额收益.....	13
图 16:	不同业绩报表超预期跳空前后的超额累积收益率表现不一.....	13
图 17:	业绩预告超预期跳空选股效果最佳 .....	14
图 18:	JUMP 因子在超预期股票池具有多头区分能力 .....	14
图 19:	理想反转因子在超预期股票池内表现优异，测试期内超额收益稳定.....	15
图 20:	大单残差因子在超预期股票池内具有空头区分度.....	15
图 21:	小单残差因子在超预期股票池内多空区分度显著.....	15
图 22:	非国企在超预期样本池内表现优于国企 .....	16
图 23:	持仓个股数量越少，超预期 Plus 组合净值越高.....	16
图 24:	超预期 30 组合历史表现惊艳，具有显著且稳定的超额收益.....	17
表 1:	分析师覆盖个股不同类型财报集中披露期位于不同的月份.....	6
表 2:	超预期 30 组合分年度绩效 .....	17
表 3:	超预期 30 组合最新明细（20210630） .....	18
表 4:	分析师一致预期数据的不同转换方式结果比较 .....	20

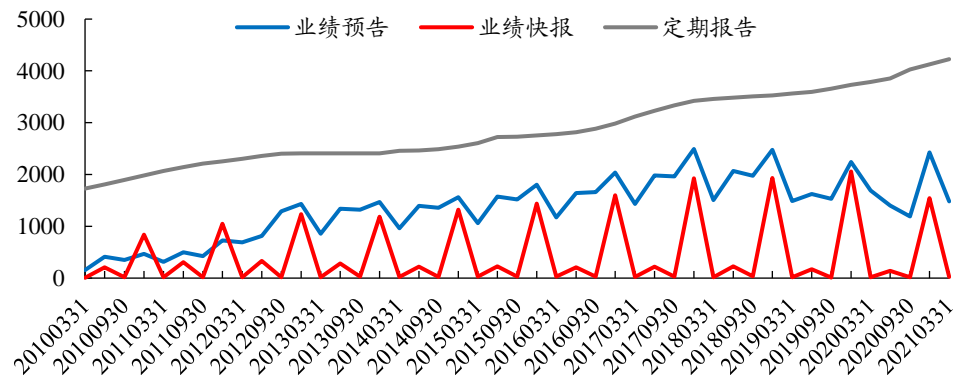
## 1、数据概览

### 1.1、财报数据：披露的节奏与数量

业绩预告、业绩快报和定期报告是投资者了解上市公司经营业绩的直接渠道。按照发布时间，一般业绩预告领先于业绩快报，业绩快报早于定期报告。按照发布内容，业绩预告披露的是上市公司经营业绩的范围；业绩快报可以理解为即将发布的定期报告的核心指标的精简版，与未来实际披露的数值差异较小；定期报告则事无巨细地披露了上市公司所有相关情况，是投资者全面了解上市公司经营状况的主要通道。

在 2019 年之前，业绩预告披露数量整体呈现波动上升的趋势，后续随着监管机构对业绩预告披露要求的放松，上市公司业绩预告披露数量有一定程度的下降。业绩快报披露主要集中在中报和年报两个时期，其中又以年报的业绩快报为主。随着上市公司数量不断增加，且退市节奏相对较缓慢，定期报告披露数量逐年稳定上升。

**图1：业绩预告、业绩快报和定期报告每期财报发布数量差异显著**



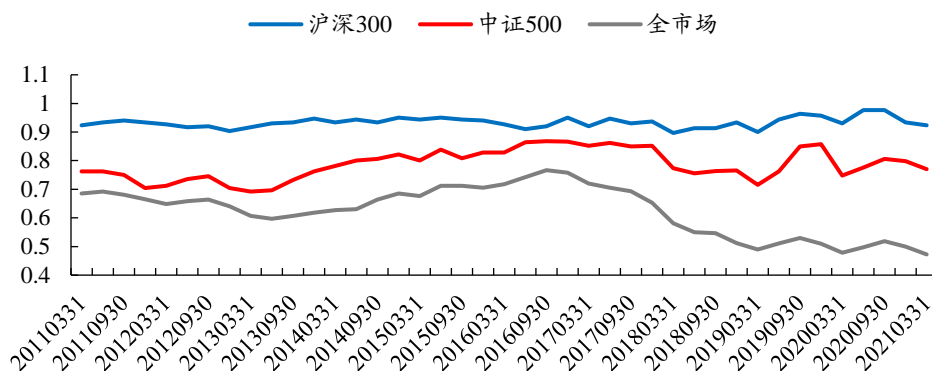
数据来源：Wind、开源证券研究所

### 1.2、分析师一致预期数据：覆盖率与 alpha 特征

分析师一致预期数据为不同券商分析师对相同覆盖标的给出的预期值的均值。分析师一致预期数据通常涵盖多个财务指标，如预期每股收益、预期营业收入、预期净利润等，本文选取的指标为分析师一致预期归母净利润。

2011 年以来，分析师一致预期归母净利润覆盖率逐季变化如图 2 所示。沪深 300 指数成分股分析师覆盖度持续保持在 90%以上的水平；中证 500 指数成分股分析师覆盖度在 70%~90%之间浮动；随着注册制的实施，公司上市节奏加快，全市场个股的分析师覆盖率呈现了较大幅度的下滑，从 2016 年 Q3 的 75%以上的覆盖率水平逐步下降到 21 年 Q1 不足 50%的覆盖率。

图2：分析师一致预期归母净利润个股覆盖率 2017 年以来下滑明显

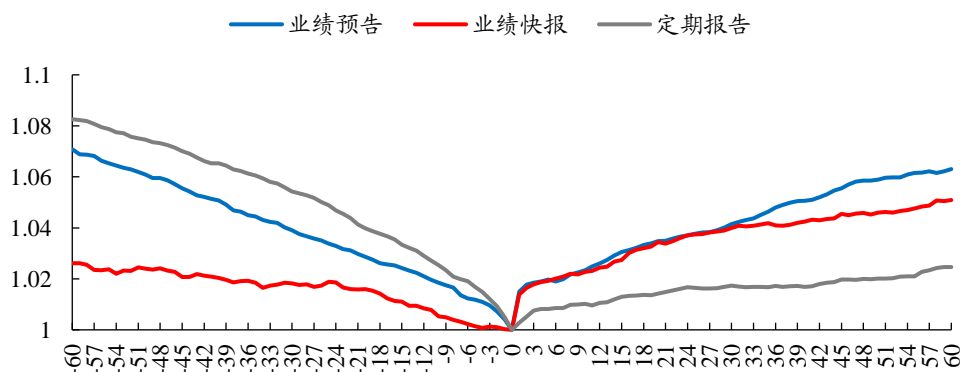


数据来源：Wind、开源证券研究所

分析师一致预期数据通常与公司财报披露的数据相比较时更能体现其用处，业绩超预期判断便是其主要作用之一。**业绩超预期的判断标准为：上市公司发布的财报数据超过了最新一期券商分析师对该公司一致预期归母净利润的均值。**囿于券商分析师并非对所有的上市公司都进行覆盖，我们对部分上市公司无法进行业绩是否超预期的判断。

对于业绩超预期个股，我们统计了个股分别在业绩预告、业绩快报和定期报告发布前后（announce date）60个交易日累积超额收益的表现。可以看到，披露时间最早的业绩预告未来超额收益表现最好，其次是业绩快报，公布时间最靠后的定期报告未来超额收益水平最低。

图3：业绩预告发布前后的累积超额收益表现最佳



数据来源：Wind、开源证券研究所

盈余公告发布一段时期以后，个股依然能产生超越基准的稳定收益，表明市场对盈余公告未能充分反应。从图3测算结果可知，业绩超预期个股未来超额收益率的衰减是相对比较缓慢的，这为我们构建因子选股提供了数据支撑。

## 2、数据预处理

在构建选股因子之前，我们首先需要对原始数据进行相应的处理。数据处理主要集中在以下四个方面：

- （1）数据口径一致化
- （2）因子数据标准化



(3) 调仓日期定期化

(4) 避免前视偏差

## 2.1、数据口径一致化

数据口径一致化指的是财务报告数据和分析师一致预期数据保持直接可比。首先是指标一致性，本文选取的指标均为归母净利润（下文简称净利润）；其次，数据统计的口径需要保持一致。

分析师一致预期净利润数据是按照自然年给出的，每个季度下的预测值均对应全年的净利润目标，而财务报表中公布的净利润则是按照财报期进行累加的，比如Q1季度公布的净利润即为第一个季度的净利润情况，而Q3季度公布的净利润则为前三个季度的累计净利润情况。为此我们需要对分析师一致预期数据进行调整，使其口径保持与财报披露数据一致，从而保证二者直接可比。具体调整方式请参考附录部分，本文采用方法二。

## 2.2、因子数据标准化

因子数据标准化主要是为了不同上市公司之间的指标直接可比，由于不同上市公司净利润量级存在较大差异，如果简单比较不同上市公司净利润超出一致预期数据的绝对值，容易产生偏差，为此我们需要基于净利润量级进行调整。

## 2.3、调仓日期定期化

不同上市公司发布财务报告的时间虽然集中在特定的时间段，但具体日期亦有较大差异。如果按照事件驱动的模式进行调仓，对于规模庞大的机构投资者会是比较繁琐的事情。为此，我们尝试将不同时间公布的业绩数据进行定期化处理。虽然不同上市公司财报发布日期不一，但主要集中在特定的时间段。

我们对分析师覆盖的上市公司不同类型财务报告的发布日期进行月度归总，根据表1可知，业绩预告发布较密集的月份有1、3、4、7、8、10六个月份，发布日期的分布整体而言相对比较离散；业绩快报发布相对密集的月份仅有1、2和7月，其中1、2月发布的主要是上一年年报的快报，7月发布的对应当年中报的快报，且业绩快报主要集中在年报层面，中报涉及较少；定期报告发布频率较高的月份是3、4、8和10月四个月份，其中3月份主要是上年年报，4月份一般同时包含当年的一季报数据和上一年的年报数据，8月份则对应当年的半年报，而10月份则为三季报集中披露期。

表1：分析师覆盖个股不同类型财报集中披露期位于不同的月份

月份	1	2	3	4	5	6	7	8	9	10	11	12
业绩预告	607	42	263	780	2	20	483	475	13	1013	2	12
业绩快报	103	828	55	88	0	0	159	15	0	18	0	0
定期报告	8	81	763	2920	12	11	131	1754	1	1979	0	0

数据来源：Wind、开源证券研究所

可以看到不同月份不同类型财报的披露数量存在较大差异，为了尽可能保证样本池可用证券数量充足，我们采用前向填充的方法对报告真空期的月份进行数据补齐，填充数据时仅采用最新可获取的上一期财报的数据，避免跨报告期填充。

## 2.4、避免前视偏差

前视偏差指在策略开发过程中采用了在当下无法获取到的信息，即未来数据。以财报数据为例，通常大部分上市公司的一季报集中在四月份发布，如果我们需要根据一季报的财务数据进行选股，那么我们最快只能在四月底进行调仓，而非财报报告期对应的月份，即三月底。

更容易被忽略的过程是我们在进行数据口径一致化的过程中，当我们在 T 年将上市公司一季度的分析师一致预期数据转化成统一口径的时候，我们在转化过程中需要依赖于 T-1 年的年报数据，但通常在实际中，很多上市公司在当前节点尚未披露 T-1 年的年报数据，为此，我们需要注意避免此类错误。

## 3、标准化预期外盈余因子（SUE）

早在 1968 年，Ball and Brown(1968)便提出了标准化预期外盈余（Standardized Unexpected Earnings, SUE）因子的构建方式。

$$SUE = \frac{R_t - E_t}{\sigma(R_t - E_t)}$$

其中  $R_t$  表示财报披露的净利润水平， $E_t$  表示预期的净利润水平， $\sigma(R_t - E_t)$  表示预测偏差的波动率水平。

在原始预期外盈余因子（下文简称 SUE）构建模型中，盈利预测数据主要是通过时间序列模型进行样本外预测得到。根据时间序列模型预测得到的净利润水平的精确度，依赖于时间序列数据的稳定性。为了避免模型偏差，本文尝试使用更加高频的券商分析师一致预期数据来替换模型预测数据。相比于统计方法得到的预测值，分析师一致预期数据具有更准确的信息含量，不足之处在于覆盖度相对较低，其在全市场样本池的覆盖度近年逐渐下滑到 50% 左右的水平。但由于 A 股上市公司绝对数量可观，分析师覆盖样本池个股数量依然比较充足。

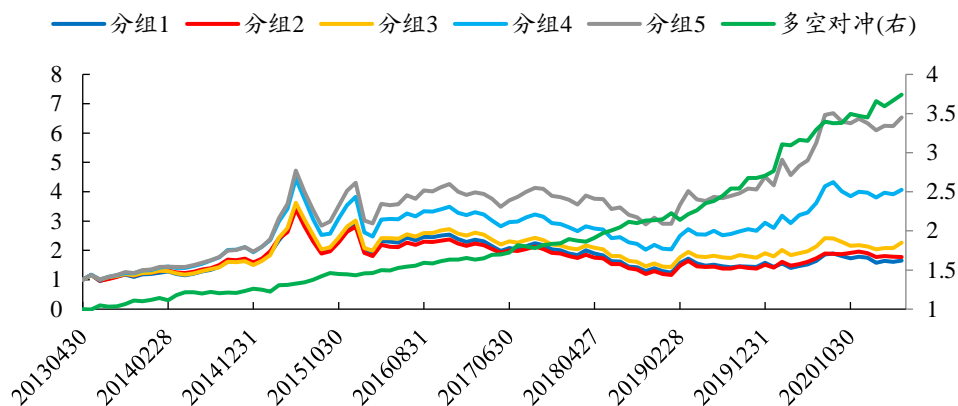
使用分析师一致预期数据替换时间序列模型估算的预测值，究竟表现如何，我们分别从不同类型的财务报告进行展开讨论。最后，我们将不同类型的财务报告按照披露时间先后顺序进行整合，并测试汇总后的财报数据选股表现。

在进行因子测试前，我们对市值和行业进行中性化处理，调仓频率为每个月最后一个交易日，交易手续费设置为双边千三，默认在每月最后一个交易日扣除。

### 3.1、业绩预告：时效性最高的业绩预告选股效果优异

业绩预告作为最早发布的财务报告，其披露数据以区间呈现的方式虽然在精确度上有所欠缺，但由于时效性最高，往往超额收益最显著。从测试区间来看，业绩预告 SUE 因子表现亮眼，多头组在测试期内年化收益率为 26.43%，夏普比率为 0.81，多空对冲年化收益率高达 17.93%，夏普值为 2.24。

**图4：业绩预告 SUE 因子表现亮眼，多头收益显著**



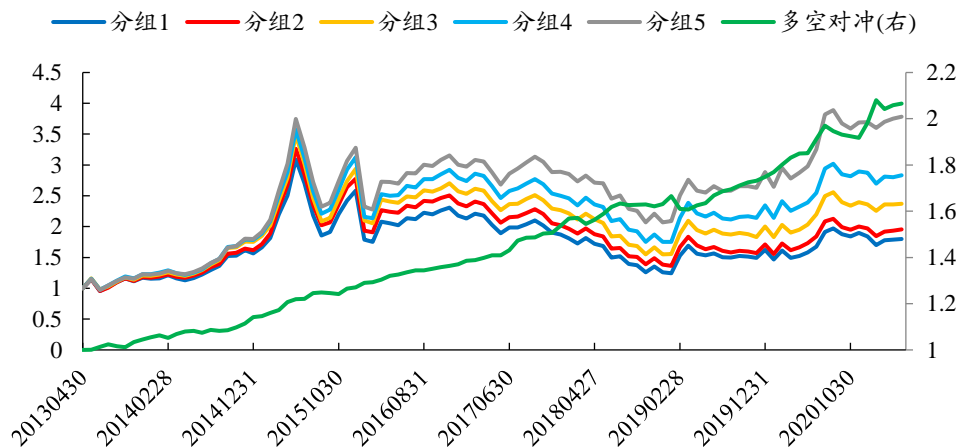
数据来源：Wind、开源证券研究所

### 3.2、定期报告：内容披露最完整的定期报告选股表现稍逊一筹

定期报告相对业绩预告和业绩快报有着最详细的数据披露，但时效性相对较差。对于没有业绩预告或业绩快报发布的相关个股，定期报告便成为我们了解上市公司财务经营状况的主要来源。

定期报告数据分组收益呈现了较好的单调性，在整个测试期内，多头年化收益率为 18.09%，夏普比率为 0.63，多空对冲年化收益率为 9.49%，夏普比率为 2.35。相比于业绩预告数据，定期报告多头和多空对冲年化收益均稍逊一筹。

**图5：定期报告 SUE 因子分组收益单调，多空对冲走势稳健**



数据来源：Wind、开源证券研究所

### 3.3、多报告融合：兼顾时效与完整性的多报告对冲表现改善显著

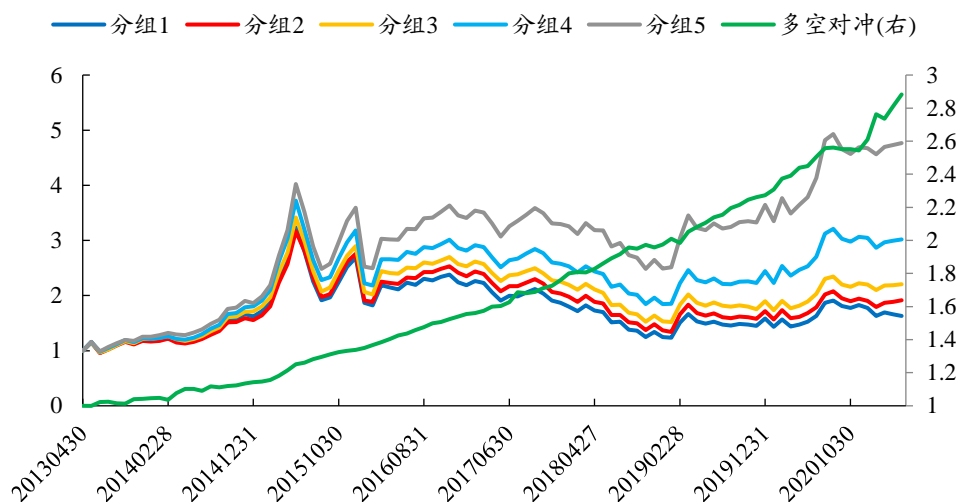
业绩快报由于数据量较少，且集中在中报和年报期间披露，导致不适合单独作为股票池构建相关因子。为了兼顾及时性和全面性，我们尝试将不同类型的财务报告融合到一起构建业绩超预期因子。当不同类型报告在同一月份发布时，选取最新的报告数据参与因子构建。

多报告融合后的 SUE 因子表现相对业绩预告 SUE 因子，虽然在多头端收益有了一定的下降，但是多头分层效果更显著；从多空净值来看，相比定期报告亦出现明显



显改善。在测试区间期间，多报告融合后的多头年化收益率为 21.56%，夏普比率 0.72，多空对冲年化收益率为 14.15%，夏普比率为 3.46。

图6：多报告融合 SUE 因子表现亮眼，多空对冲夏普比率高达 3.46



数据来源：Wind、开源证券研究所

总体看来，基于个股财报数据和分析师一致预期数据构建的标准化预期外盈余因子在不同报告样本池中均具有良好的表现，其中业绩预告对应的多头表现最佳，多报告融合对应的多空对冲夏普比率最高。

#### 4、预期外百分比因子（PCT）

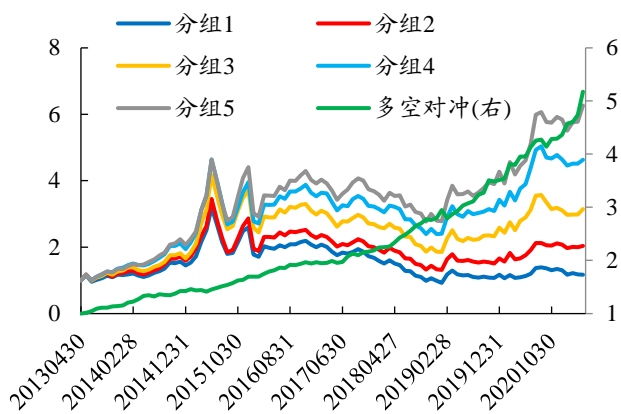
预期外百分比因子（下文简称 PCT）构建方式与 SUE 因子几无差异，最大的不同在于分母部分的变化，但本质均是为了剔除个股净利润量级的差异对结果的影响。在计算预期外百分比时，当分析师预测数据为负时，需要进行取绝对值调整，同时剔除分析师预测数据为 0 值的部分。

$$PCT = \frac{R_t - E_t}{\text{abs}(E_t)}, E_t \neq 0$$

预期外百分比因子在业绩预告样本池中表现优异，多头年化收益率为 25.78%，夏普比率 0.79，多空对冲年化收益率 22.81%，夏普比率 2.93。

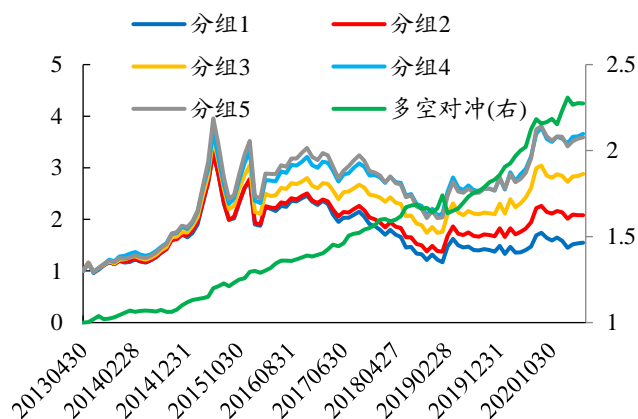
预期外百分比因子在定期报告样本池中表现有所下滑，多头年化收益 17.31%，夏普比率为 0.58，多空对冲年化收益率为 10.81%，夏普比率为 2.10。

图7：业绩预告 PCT 因子分层效果优异



数据来源：Wind、开源证券研究所

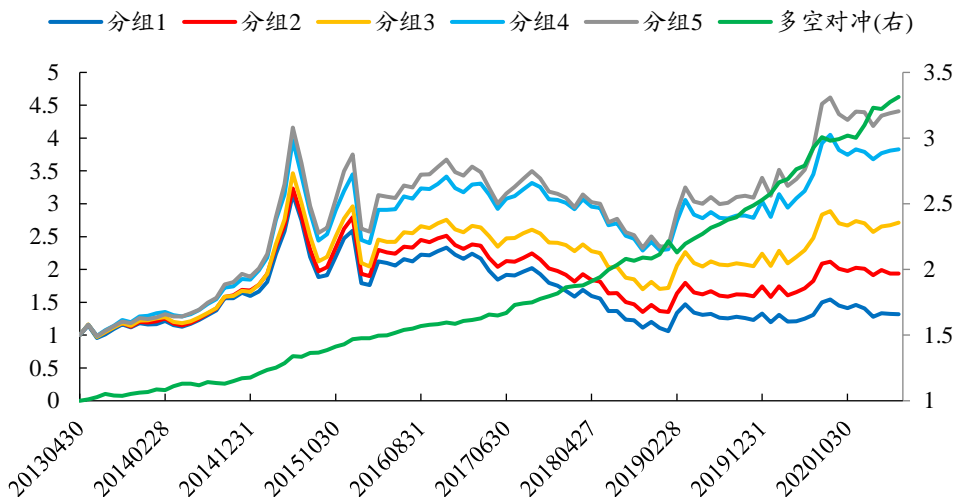
图8：定期报告 PCT 因子多头组区分度有所下滑



数据来源：Wind、开源证券研究所

多报告融合的业绩预告外百分比因子分组收益虽然整体单调，但分组四和分组五比较靠近，多头年化收益率 17.23%，夏普比率 0.72，多空对冲年化收益率 14.10%，夏普比率 2.68。

图9：多报告融合 PCT 因子多空对冲夏普比率为 2.68



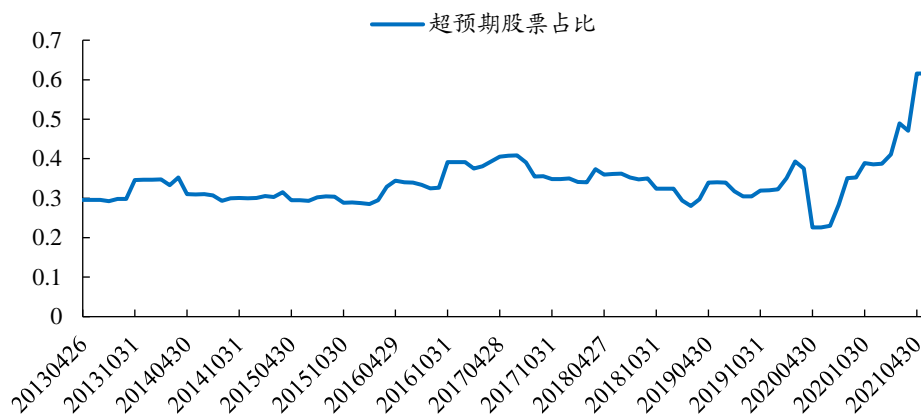
数据来源：Wind、开源证券研究所

无论是 SUE 因子，还是 PCT 因子，在测试区间均表现了优异的选股能力。SUE 因子和 PCT 因子的正向选股能力，也意味着“业绩超预期股票”具有相对更高的收益预期。

## 5、超预期股票池的增强探索

我们统计了每期超预期个股占比，在测试期间，超预期个股在分析师覆盖股票池中占比大约处于 30%~40% 的区间，2020 年第一季度由于受新冠疫情的影响，超预期个股占比下滑到 20% 左右水平，受益于国内疫情率先控制，2021 年一季度在 2020 年低基数的基础上业绩超预期个股占比大幅上升到 60% 的水平。

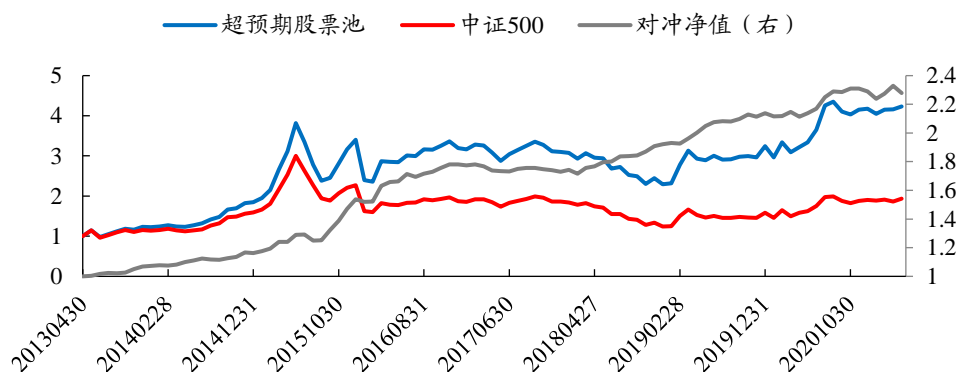
图10: 超预期个股占比在 2020 年三季度以来增长比较明显



数据来源: Wind、开源证券研究所

在多报告融合样本池中(下文如无特别说明,均在多报告融合样本池中进行测试),我们对超预期个股按照月频调仓的方式进行组合构建,并与业绩基准(中证 500 指数)进行对比。超预期股票池组合从 2013 年 4 月以来,年化收益率 19.75%,对冲中证 500 后的年化收益率 10.84%,最大回撤 3.71%,收益波动比为 1.81。

图11: 超预期样本池内个股表现优于基准组合中证 500 指数



数据来源: Wind、开源证券研究所

从回测效果来看,超预期股票池相比基准具有一定的超额收益,但超额幅度并不是特别显著,且近期超额收益有走平趋势。为此,我们尝试在超预期股票池内进行股票精选,以构建更加优异的投资组合。

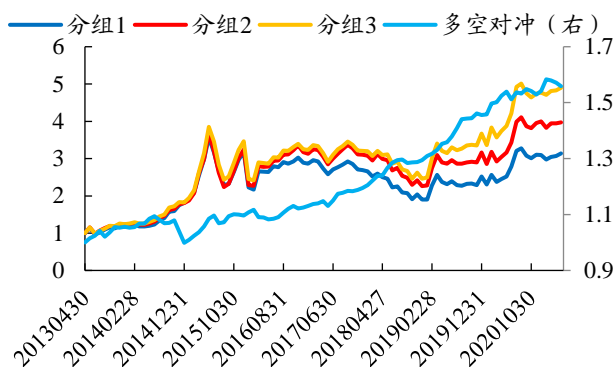
超预期股票池内的选股,我们尝试从多个不同的角度展开,包括基于财务数据的标准化预期外因子、基于价格数据的区间收益和跳空收益、基于交易行为的开源金工独家因子,以及基于资金流的大小单残差。

### 5.1、标准化预期外因子

从上述预期外盈余因子和预期外百分比因子的表现来看,二者均具有正向选股能力,当个股超预期幅度越大,未来表现越好。为了更好地筛选超预期标的,我们首先基于 SUE 和 PCT 因子在超预期股票池内进行股票分组测试。

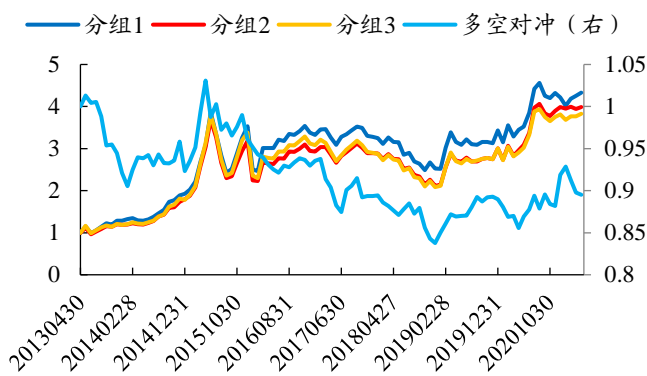
从测试结果来看,预期外百分比因子在超预期股票池内的表现比较差,基本上没有进一步的分层效果。而 SUE 因子在超预期股票池内的分组表现较好,其适合用来对超预期标的进行进一步的区分。

图12: SUE 因子在超预期股票池内依然具有分层效果



数据来源: Wind、开源证券研究所

图13: PCT 因子在超预期股票池内多头组分层效果下滑



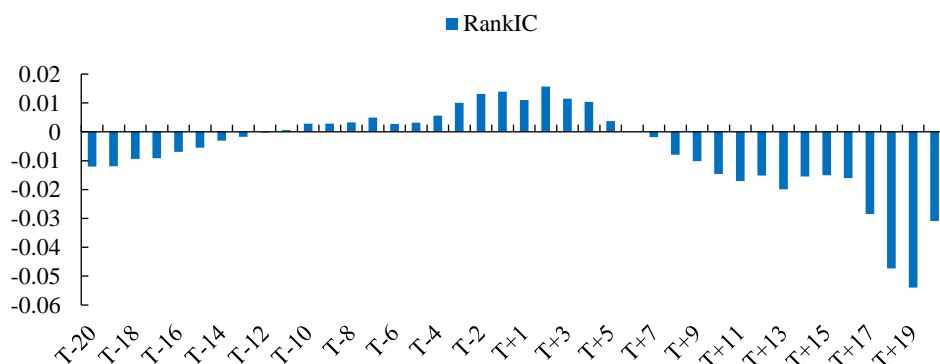
数据来源: Wind、开源证券研究所

## 5.2、公布日前后的区间收益

超预期外的业绩包含了两个维度的信息，业绩未达预期和业绩超预期。对于业绩未达预期的个股，其超额收益来源于对相关个股做空，但受限于 A 股对机构投资者的做空限制，实际操作存在一定的障碍。对于业绩超预期个股，市场的表现也呈现出不同的模式，一种是业绩公布后股价反应并不强烈，表明当前股价已经提前反映了业绩超预期表现，其可能的解释是市场中总有一些知情交易者在交易；另一种是业绩公布后的次日股价发生向上跳空，表明业绩超出市场预期尚未被市场完全定价。

我们以财报公告日为界限，在超预期样本池中，分别向前和向后回溯 N 个交易日，计算个股在 N 个交易日内的区间超额收益率，并计算该区间超额收益率与未来持仓收益率的 RankIC。可以看到，公告日附近的区间超额收益率与未来持仓收益率的 RankIC 值相对较高。离公告日发布时点较近时，区间收益与未来持仓收益多为正值，呈现动量效应，而离公告日发布时点较远时，区间收益与未来持仓收益多为负值，表现为反转效应。

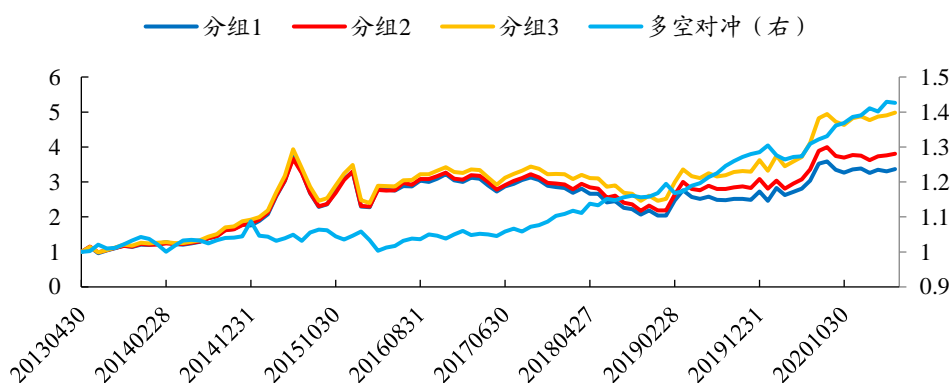
图14: 超预期收益 (OER) 因子短区间呈现动量效应，长区间呈现反转效应



数据来源: Wind、开源证券研究所

为了避免刻意选择，我们选取公告日前后一个交易[T-1, T+1]的超额收益之和构建超预期收益因子 OER(Over Expectation Return)。其中，公告前一个交易日收益用来表征知情交易者的相关信息，后一个交易日收益用来表征市场反应情况。超预期收益因子市值行业中性化后回溯结果如图 15 所示：2016 年之前，OER 因子在超预期股票池内的区分度不显著，但之后开始逐渐有了稳定的区分效果。

图15: 2017年后，OER因子在超预期股票池中具有稳定超额收益



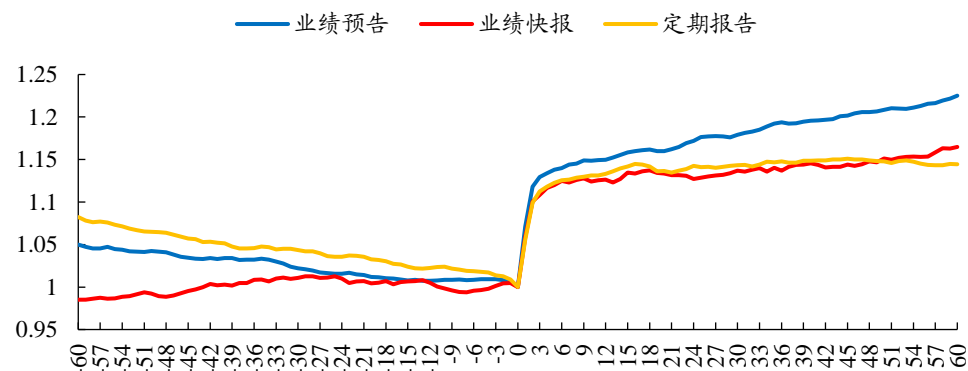
数据来源: Wind、开源证券研究所

### 5.3、超预期跳空

在报告发布时除了区间累计超额收益之外，投资者比较关注的还有报告发布后的下一个交易日股价是否出现跳空现象。一般而言，股价跳空幅度越大，表明个股业绩超预期幅度越高。

我们将发布财务报告后的第二个交易日发生股价跳空的个股单独提取出来，观察其财务业绩公告日期前后 60 个交易日的累积超额收益表现。可以看到，公告发布后一个交易日若股价发生跳空现象，则其幅度平均在 5% 以上。其中，业绩预告跳空幅度最大，平均值超 7%，业绩快报次之，定期报告跳空幅度最小。在业绩预告发布后的第二个交易日，首日跳空股票次日亦有 4% 左右的涨幅。

图16: 不同业绩报表超预期跳空前后的超额累积收益率表现不一

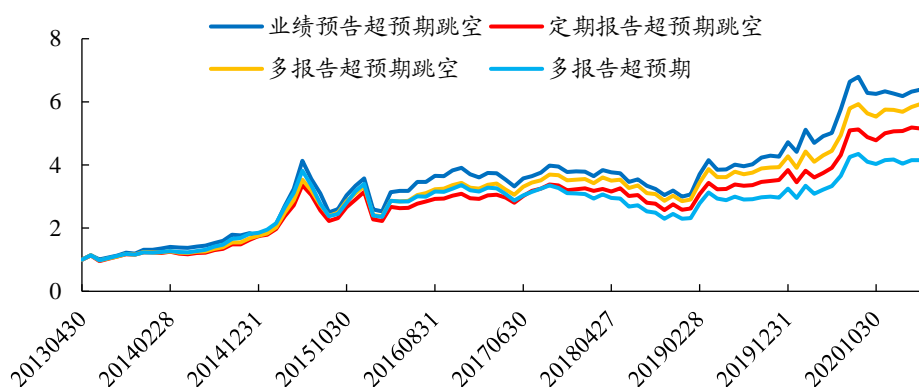


数据来源: Wind、开源证券研究所

我们分别针对业绩预告、定期报告以及多报告融合的数据进行测试。具体方法为在超预期股票池内，每月末将发布报告后第二天股价发生跳空的股票挑选出来，等权持有到下个月做再平衡，回测结果如图 17 所示。可以看到，时效性最佳的业绩预告具有最佳的超额收益，定期报告由于发布最晚，超额收益相对业绩预告有一定的下降，多报告融合的超额收益位于两者之间。他们三者的收益全都大于原始超预期股票池的表现。



图17: 业绩预告超预期跳空选股效果最佳

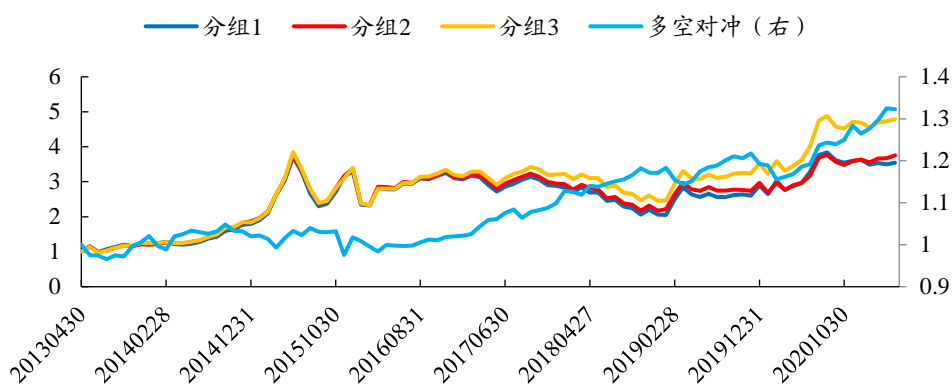


数据来源: Wind、开源证券研究所

虽然超预期跳空选股策略具有较好的表现,但由于每期所选个股数不确定,我们尝试将其因子化。首先使用股票发布公告后第二个交易日最低价和前一个交易日收盘价计算股票跳空收益率,同样计算中证 500 指数跳空收益率,使用股票跳空收益率减去指数跳空收益率得到跳空 (JUMP) 因子。减去指数收益率主要是为防止不同时间点市场强度不同而造成的影响。

JUMP 因子在超预期股票池内的回测表现如图 18 所示,在超预期样本池内,JUMP 因子在 2017 年之前表现乏善可陈,2017 年之后开始有显著的超额收益。

图18: JUMP 因子在超预期股票池具有多头区分能力

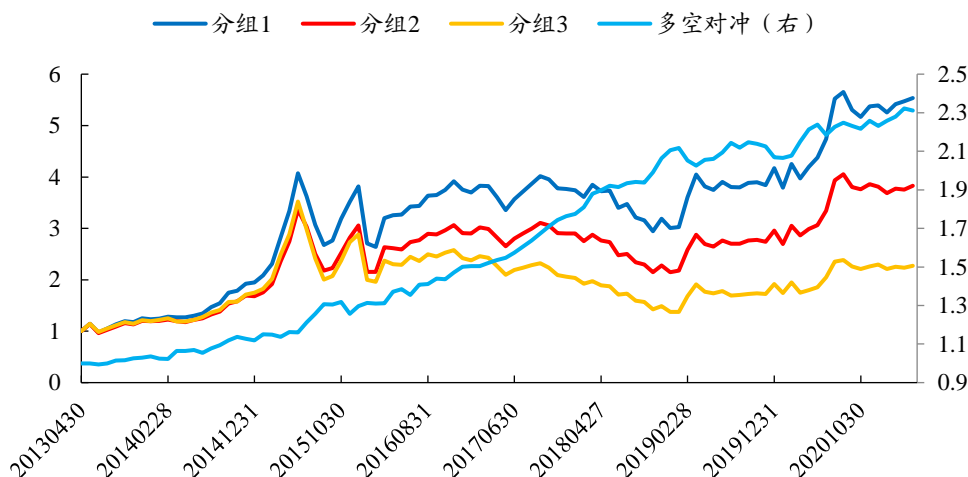


数据来源: Wind、开源证券研究所

## 5.4、交易行为因子

对于发生业绩超预期的股票来说,通常其交易活跃度会有不同程度的提升,而交易活跃度与交易行为因子息息相关。通过图 14 的测试结果可知,公告日 7 个交易日后的累计收益呈现反转效应,所以我们测试了开源金工独家的理想反转因子在超预期股票池内的表现。可以看出,理想反转因子在全区间内对超预期股票池具有显著的分层能力。理想反转因子详细的构建方法,请参见开源金融工程团队研究报告《A 股反转之力的微观来源》。

图19: 理想反转因子在超预期股票池内表现优异，测试期内超额收益稳定

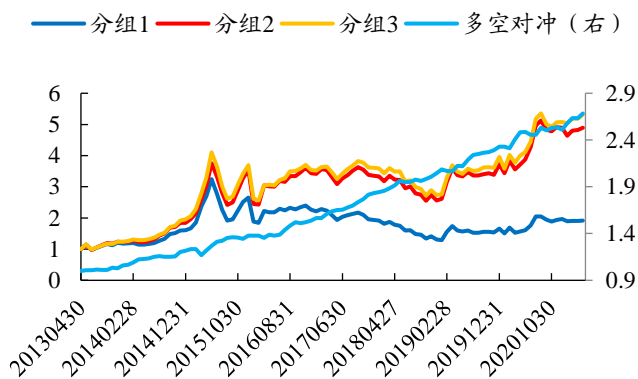


数据来源: Wind、开源证券研究所

## 5.5、大小单残差因子

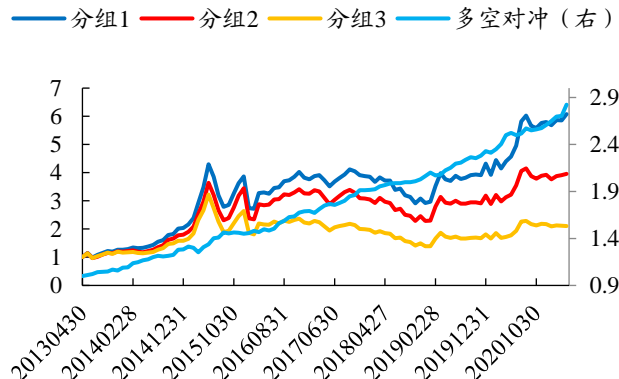
大小单残差因子展示了个股资金流的强弱关系，反映了超预期股票池内的个股被资金青睐的程度。从这二者在超预期股票池内的分组回测可以看出，该因子在超预期股票池内非常有效，而且小单残差因子比大单残差因子的效果表现更好。关于大小单残差因子的构建，参见开源金融工程团队研究报告《大单与小单资金流的 alpha 能力》。

图20: 大单残差因子在超预期股票池内具有空头区分度



数据来源: Wind、开源证券研究所

图21: 小单残差因子在超预期股票池内多空区分度显著

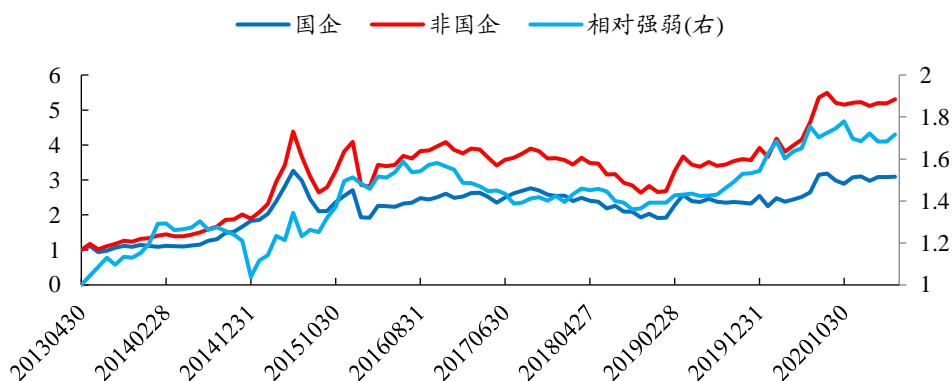


数据来源: Wind、开源证券研究所

## 5.6、国企与非国企

对于企业而言，其本身是国企或非国企在市场主动性和激励机制方面存在一定的差异。一般而言，市场对非国企的财务表现反应更加激烈，对国企的财务表现则相对温和。为了验证这一想法我们将超预期股票池分为国企和非国企股票池进行测试，结果如下所示。从图中可以看出在超预期股票池内，非国企企业的表现较好于国企企业，胜率为 62.5%，这和我们之前的猜想是一致的，但是其相对强弱曲线的波动性比较高，单纯的使用非国企进行后续的筛选可能会错过一定的机会，所以后文我们不将其作为有效的筛选指标。

图22：非国企在超预期样本池内表现优于国企

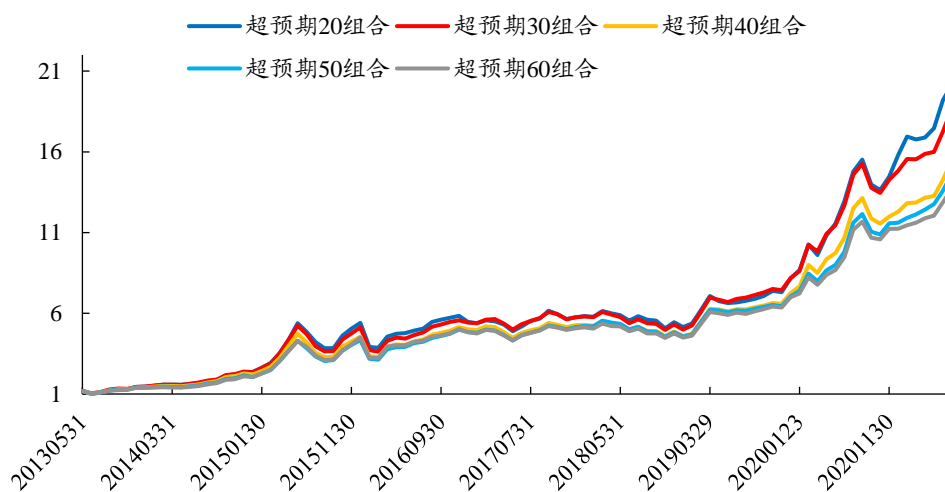


数据来源：Wind、开源证券研究所

## 6、超预期 Plus 组合构建

根据前文的探索结果，我们最终采取如下的方案进行最后精选组合的构建，在超预期股票池内首先剔除上市不满 60 天，停牌，ST，一字涨跌停的股票得到初选池，然后将 SUE、OER([T-1,T+1])、JUMP、理想反转、小单残差等权重排序加总得到最终的筛选因子，最终使用该筛选因子在初选池选取前 N 只股票作为**超预期 Plus 组合**。

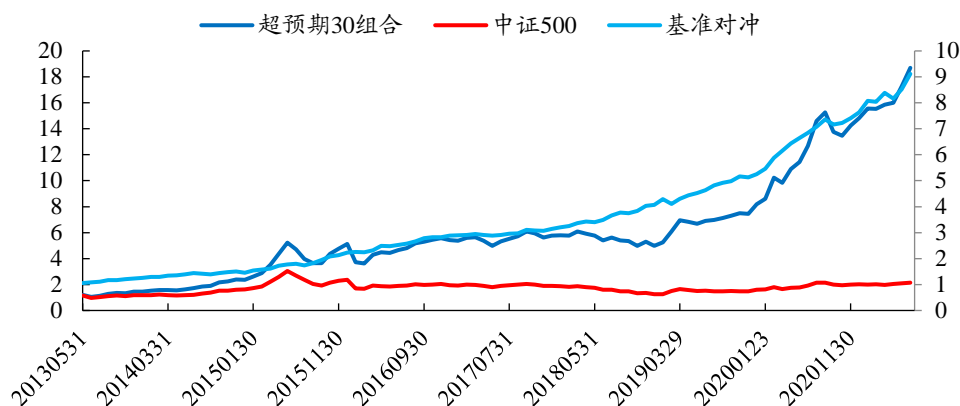
图23：持仓个股数量越少，超预期 Plus 组合净值越高



数据来源：Wind、开源证券研究所

可以发现，当组合持仓个股数量越少时，超预期 Plus 组合的收益越高。为了保证持仓个股数量适中，我们选取持仓个股数为 30 的组合作为最终推荐持仓，其测试期内表现如图 24 所示。超预期 30 组合多头收益率达到 43.13%，夏普比率为 1.53，月度胜率约 70%；对冲基准收益为 31.08%，夏普比率为 3.6，月度胜率超 80%。最新一期的超预期 30 组合持仓见表 3。

图24: 超预期30组合历史表现惊艳，具有显著且稳定的超额收益



数据来源: Wind、开源证券研究所

表2: 超预期30组合分年度绩效

年份	类别	年化收益率	年化波动率	收益波动比	最大回撤	月度胜率
2013	投资组合	75.61%	34.41%	2.20	12.94%	75.00%
(20130430起)	基准组合	26.22%	31.78%	0.83	15.66%	62.50%
	对冲净值	40.46%	6.94%	5.83	0.00%	100.00%
	投资组合	61.70%	14.76%	4.18	1.86%	83.33%
2014	基准组合	39.52%	13.62%	2.90	5.17%	83.33%
	对冲净值	16.21%	7.63%	2.12	3.31%	75.00%
	投资组合	117.68%	45.95%	2.56	30.36%	75.00%
2015	基准组合	45.28%	41.57%	1.09	36.82%	66.67%
	对冲净值	53.30%	10.88%	4.90	3.70%	91.67%
	投资组合	5.79%	36.66%	0.16	2.59%	66.67%
2016	基准组合	-18.10%	34.13%	-0.53	5.00%	50.00%
	对冲净值	29.28%	7.65%	3.83	0.54%	75.00%
	投资组合	6.36%	16.13%	0.39	11.54%	58.33%
2017	基准组合	-1.91%	12.33%	-0.16	10.25%	41.67%
	对冲净值	8.80%	5.72%	1.54	2.03%	66.67%
	投资组合	-13.57%	16.18%	-0.84	18.27%	33.33%
2018	基准组合	-33.77%	14.11%	-2.39	32.98%	16.67%
	对冲净值	29.45%	6.26%	4.71	0.79%	83.33%
	投资组合	64.09%	20.82%	3.08	3.95%	75.00%
2019	基准组合	26.32%	24.48%	1.08	11.02%	58.33%
	对冲净值	29.22%	8.91%	3.28	4.21%	83.33%
	投资组合	81.32%	27.97%	2.91	11.82%	75.00%
2020	基准组合	26.75%	22.43%	1.19	9.90%	75.00%
	对冲净值	45.27%	8.46%	5.35	2.68%	91.67%
	投资组合	59.08%	12.49%	4.73	0.09%	83.33%
(20210630止)	基准组合	11.61%	8.20%	1.42	2.39%	66.67%
	对冲净值	42.48%	13.21%	3.22	2.75%	66.67%
	投资组合	43.13%	28.17%	1.53	30.65%	68.37%
全区间	基准组合	9.70%	25.48%	0.38	58.61%	57.14%

对冲净值	31.08%	8.63%	3.60	4.21%	81.63%
------	--------	-------	------	-------	--------

数据来源：Wind、开源证券研究所

表3：超预期 30 组合最新明细（20210630）

股票代码	股票名称	所属一级行业	PE-TTM	流通市值（亿元）
603416.SH	信捷电气	电气设备	23.91	57.90
300516.SZ	久之洋	电子	74.81	59.78
002139.SZ	拓邦股份	电子	31.38	171.10
300782.SZ	卓胜微	电子	126.85	1027.19
601339.SH	百隆东方	纺织服装	15.62	80.85
600461.SH	洪城环境	公用事业	9.53	66.11
603678.SH	火炬电子	国防军工	38.84	309.84
603010.SH	万盛股份	化工	18.10	101.39
603203.SH	快克股份	机械设备	25.65	52.76
300756.SZ	金马游乐	机械设备	64.36	10.16
002268.SZ	卫士通	计算机	74.57	162.99
002705.SZ	新宝股份	家用电器	18.64	208.98
300737.SZ	科顺股份	建筑材料	19.50	147.29
300492.SZ	华图山鼎	建筑装饰	337.78	60.42
600039.SH	四川路桥	建筑装饰	7.60	226.38
600298.SH	安琪酵母	农林牧渔	29.99	448.14
601311.SH	骆驼股份	汽车	13.86	129.78
603089.SH	正裕工业	汽车	12.41	19.34
300733.SZ	西菱动力	汽车	133.29	26.06
603179.SH	新泉股份	汽车	37.84	105.24
603833.SH	欧派家居	轻工制造	35.63	860.64
600694.SH	大商股份	商业贸易	8.30	60.98
600132.SH	重庆啤酒	食品饮料	71.60	958.02
000799.SZ	酒鬼酒	食品饮料	125.18	830.52
002732.SZ	燕塘乳业	食品饮料	26.57	36.33
300394.SZ	天孚通信	通信	32.66	81.01
688016.SH	心脉医疗	医药生物	127.41	167.73
002540.SZ	亚太科技	有色金属	15.49	43.60
600490.SH	鹏欣资源	有色金属	15.27	76.72
688122.SH	西部超导	有色金属	61.47	201.35

数据来源：Wind、开源证券研究所

## 7、附录：分析师一致预期数据调整方式对比

前文我们提到分析师一致预期数据是按照全年的形式给出的预测值，而财报数据是按照逐季累加的形式提供的。在不做调整的前提下，分析师一致预期数据仅能在年报季才可直接使用。为了避免潜在有效信息的浪费，我们对分析师一致预期数据进行季度化调整。常用的调整方法有两种，一种是将年度分析师一致预期数据转化为和财报发布格式相同的逐季累加的数据，一种是将分析师预测数据和财报数据同时转化为单季度数据。以下就两种计算方法分别进行说明。



假设股票 A 在 Q3 季度分析师一致预期净利润为  $E_t$ ，其去年实际净利润水平为  $R_{t-1}^4$ ，去年和今年前三季度实际累计净利润水平分别为  $R_{t-1}^3$  和  $R_t^3$ 。

### 7.1、年度数据转成逐季累加模式

第一步，将今年 Q3 季度的分析师一致预期数据  $E_t$  除以去年年报披露的真实净利润数据  $R_{t-1}^4$ ，得到 Q3 季度时分析师预期的全年净利润增长率  $G_t = \frac{E_t}{R_{t-1}^4} - 1$ ，需要注意的是，若去年实际净利润水平  $R_{t-1}^4$  为负数时，则预期净利润增长率计算方式调整为  $G_t = 1 - \frac{E_t}{R_{t-1}^4}$ 。

第二步，将去年前三季度累计实现净利润  $R_{t-1}^3$  乘以预期的增长率水平  $G_t$ ，得到今年前三季度分析师预期累计净利润  $R_{t-1}^3(1 + G_t)$ 。当  $R_t^3 > R_{t-1}^3(1 + G_t)$  时，则认为业绩超预期，反之，业绩不达预期。

其他季度预期数据处理方式同理，年报分析师一致预期数据无需调整，可直接与实际披露数据进行比较。

年度数据转成逐季累加模式的缺陷在于分析师会根据本年度已发布财报的披露数据对全年的预期净利润水平进行调整，即其预测增速对应的是全年的增速水平，但我们在计算分析师预测的逐季累加净利润时，我们将全年的增速水平套用到特定季度，存在一定的偏差。根据上面的计算方法，当四季度公司业绩超速超全年平均水平时，该方法会高估前三季度的累计净利润水平，导致业绩超预期可能性降低。反之，则会低估前三季度的累计净利润水平，导致业绩超预期可能性上升。

### 7.2、年度数据转成单季度模式

第一步，将分析师一致预期净利润  $E_t$  减去今年前三季度已实现净利润  $R_t^3$  水平，得到四季度分析师一致预期净利润值  $E_t^{4s} = E_t - R_t^3$ 。

第二步，将去年年报净利润  $R_{t-1}^4$  减去去年前三季度累计净利润  $R_{t-1}^3$ ，得到去年第四季度的净利润  $R_{t-1}^{4s} = R_{t-1}^4 - R_{t-1}^3$ 。

第三步，将分析师一致预期的四季度净利润  $E_t^{4s}$  除以去年第四季度的净利润  $R_{t-1}^{4s}$ ，得到预期净利润增速  $G_t = \frac{E_t^{4s}}{R_{t-1}^{4s}} - 1$ ，若  $R_{t-1}^{4s}$  为负，则  $G_t = 1 - \frac{E_t^{4s}}{R_{t-1}^{4s}}$ 。

第四步，将去年三季度的实际净利润  $R_{t-1}^{3s}$  乘以净利润增速  $G_t$  得到今年三季度预期净利润水平  $E_t^{3s} = R_{t-1}^{3s}(1 + G_t)$ 。当  $E_t^{3s} > R_t^{3s}$  时，业绩超预期；反之，业绩不达预期。

年度数据转成单季度数据的模式，整体的逻辑在于未来的增速是过去增速的复现，其依然存在第一种转换方式的问题。

### 7.3、数据转换实例

为了更清楚地比较两种转换方式的差异，我们以一个实际案例来进行说明。沪电股份 (002463.SZ) 作为 5G 行业核心 PCB 厂商，在 2019 年由于 5G 的推出迎来了业绩爆发，业绩持续超预期。我们对比了两种方式下年化的分析师一致预期数据的转化结果，可以看到虽然转化后的具体数值有所差异，但是方向保持整体保持一致。

表4: 分析师一致预期数据的不同转换方式结果比较

沪电股份	一致预期	定期报告 (累计值)	定期报告 (单季度)	一致预期 (累计值)	一致预期 (单季度)	是否超预期 (累积值)	是否超预期 (单季度)
2018/3/31	3.39	0.70	0.70				
2018/6/30	3.13	1.97	1.26				
2018/9/30	4.61	3.83	1.87				
2018/12/31	5.46	5.70	1.87				
2019/3/31	7.14	1.62	1.62	0.88	0.88	Y	Y
2019/6/30	7.61	4.78	3.16	2.62	1.51	Y	Y
2019/9/30	10.76	8.51	3.73	7.23	2.98	Y	Y
2019/12/31	11.56	12.06	3.55	11.56	3.05	Y	Y
2020/3/31	15.64	2.14	2.14	2.11	2.11	Y	Y
2020/6/30	15.63	5.84	3.70	6.20	4.09	N	N
2020/9/30	15.04	9.55	3.14	10.62	4.72	N	N
2020/12/31	14.23	13.43	3.88	14.23	4.68	N	N
2021/3/31	15.86	2.21	2.21	2.53	2.53	N	N

数据来源: Wind、开源证券研究所

## 8、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的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	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	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

### 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

### 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

### 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