

2021年06月01日

证券研究报告—金融工程专题报告

# 股票多因子研究：资金流因子分析及组合构建

## 摘要：

分析日期 2021年05月31日

证券分析师： 曾磊  
执业证书编号： S0630519070001  
电话： 021-20333347  
邮箱： zenglei@longone.com.cn

股票单因子是股票多因子研究的基础，基于单因子有效性、稳健性的识别及分析是开展多因子研究的前提条件。

在资金流因子设计上，我们使用个股过去20个交易日的累计资金净流入与当日股票总市值之比作为因子值，为了研究不同类型资金流入情况，设计了四个因子，分别为：资金净流入（特大单净流入+大单净流入）CFLOW\_LB\_20D、特大单资金净流入 CFLOW\_LG\_20D、大单资金净流入 CFLOW\_BG\_20D 和中单资金净流入 CFLOW\_MD\_20D。

资金净流入 CFLOW\_LB\_20D 因子在持仓周期为30天的情况下，IC均值为0.056，IR为0.48，综合、交通运输、电气设备、采掘、家用电器和农林牧渔行业的因子IC相对较高。

特大单资金净流入 CFLOW\_LG\_20D 因子在持仓周期为30天的情况下，IC均值为0.01，IR为0.11，钢铁、银行、综合、和家用电器行业的因子IC相对较高。

大单资金净流入 CFLOW\_BG\_20D 因子在持仓周期为30天的情况下，IC均值为0.087，IR为0.73，综合、汽车、采掘、电气设备及机械设备行业的因子IC相对较高。

中单资金净流入 CFLOW\_MD\_20D 因子在持仓周期为30天的情况下，IC均值为0.055，IR为0.65，公用事业、汽车、纺织服装、电气设备及机械设备行业的因子IC相对较高。

根据四个不同类型资金流因子在不同持仓周期下的IC、IR等统计指标的对比结果，总体上看，大单资金净流入 CFLOW\_BG\_20D 因子在持仓10天和30天天下的IC、IR表现均强于另外三个因子。

针对资金流入因子中大单资金净流入因子具有较好的IC表现，我们在中证500指数成分中构建相应的大单资金净流入因子组合，进而衡量该因子在中证500指数上的表现，纯资金流向因子（不考虑因子择时及其他类型因子收益）超额收益在组合历史回测期间内实现56.76%的超额收益，年化收益率为7.57%，年化波动率为5.76%，Sharpe比率为1.31，超额收益净值最大历史回撤为8.05%，Calmar比率为0.94。

## 相关研究报告

## 正文目录

1. 单因子研究流程.....	4
1.1. 样本空间.....	4
1.2. 参数设定.....	4
1.3. 研究对象.....	4
2. 资金净流入因子.....	5
2.1. 因子特征分析.....	5
2.2. 因子 IC 分析.....	6
2.3. 因子换手率分析.....	9
3. 特大单资金净流入因子.....	9
3.1. 因子特征分析.....	9
3.2. 因子 IC 分析.....	11
3.3. 因子换手率分析.....	13
4. 大单资金净流入因子.....	14
4.1. 因子特征分析.....	14
4.2. 因子 IC 分析.....	16
4.3. 因子换手率分析.....	18
5. 中单资金净流入因子.....	18
5.1. 因子特征分析.....	18
5.2. 因子 IC 分析.....	20
5.3. 因子换手率分析.....	22
6. 因子对比与模拟组合.....	23
6.1. 因子 IC 对比.....	23
6.2. 因子历史组合回测.....	24

## 图表目录

图 1 CFLOW_LB_20D 因子分位数对比图.....	5
图 2 CFLOW_LB_20D 因子分位数组合累积收益图.....	6
图 3 CFLOW_LB_20D 因子加权组合累计收益图.....	6
图 4 CFLOW_LB_20D 因子按行业分位数收益对比图（持仓 10 天）.....	6
图 5 CFLOW_LB_20D 因子 IC 分布直方图.....	7
图 6 CFLOW_LB_20D 因子 IC 月度平均（持仓 10 天）.....	8
图 7 CFLOW_LB_20D 因子 IC 月度均值热力图.....	8
图 8 CFLOW_LB_20D 因子 IC 衰减图.....	8
图 9 CFLOW_LB_20D 因子 IC 行业均值.....	9
图 10 CFLOW_LG_20D 因子分位数对比图.....	10
图 11 CFLOW_LG_20D 因子分位数组合累积收益图.....	10
图 12 CFLOW_LG_20D 因子加权组合累计收益图.....	11
图 13 CFLOW_LG_20D 因子按行业分位数收益对比图（持仓 10 天）.....	11
图 14 CFLOW_LG_20D 因子 IC 分布直方图.....	12
图 15 CFLOW_LG_20D 因子 IC 月度平均（持仓 10 天）.....	12
图 16 CFLOW_LG_20D 因子 IC 月度均值热力图.....	13
图 17 CFLOW_LG_20D 因子 IC 衰减图.....	13
图 18 CFLOW_LG_20D 因子 IC 行业均值.....	13
图 19 CFLOW_BG_20D 因子分位数对比图.....	15
图 20 CFLOW_BG_20D 因子分位数组合累积收益图.....	15

图 21 CFLOW_BG_20D 因子加权组合累计收益图.....	15
图 22 CFLOW_BG_20D 因子按行业分位数收益对比图（持仓 10 天） .....	16
图 23 CFLOW_BG_20D 因子 IC 分布直方图.....	16
图 24 CFLOW_BG_20D 因子 IC 月度平均（持仓 10 天） .....	17
图 25 CFLOW_BG_20D 因子 IC 月度均值热力图.....	17
图 26 CFLOW_BG_20D 因子 IC 衰减图.....	17
图 27 CFLOW_BG_20D 因子 IC 行业均值.....	18
图 28 CFLOW_MD_20D 因子分位数对比图 .....	19
图 29 CFLOW_MD_20D 因子分位数组合累积收益图 .....	19
图 30 CFLOW_MD_20D 因子加权组合累计收益图.....	20
图 31 CFLOW_MD_20D 因子按行业分位数收益对比图（持仓 10 天） .....	20
图 32 CFLOW_MD_20D 因子 IC 分布直方图.....	21
图 33 CFLOW_MD_20D 因子 IC 月度平均（持仓 10 天） .....	21
图 34 CFLOW_MD_20D 因子 IC 月度均值热力图.....	22
图 35 CFLOW_MD_20D 因子 IC 衰减图.....	22
图 36 CFLOW_MD_20D 因子 IC 行业均值.....	22
图 37 资金流因子选股历史回测走势对比图 .....	24
图 38 资金流因子选股历史回测超额净值走势图.....	24
表 1 资金流向类型定义.....	4
表 2 CFLOW_LB_20D 因子分位数统计表 .....	5
表 3 CFLOW_LB_20D 因子 IC 统计表 .....	7
表 4 CFLOW_LB_20D 因子分位数换手率对比表.....	9
表 5 CFLOW_LG_20D 因子分位数统计表 .....	10
表 6 CFLOW_LG_20D 因子 IC 统计表 .....	12
表 7 CFLOW_LG_20D 因子分位数换手率对比表.....	14
表 8 CFLOW_BG_20D 因子分位数统计表 .....	14
表 9 CFLOW_BG_20D 因子 IC 统计表 .....	16
表 10 CFLOW_BG_20D 因子分位数换手率对比表.....	18
表 11 CFLOW_MD_20D 因子分位数统计表 .....	19
表 12 CFLOW_MD_20D 因子 IC 统计表.....	21
表 13 CFLOW_MD_20D 因子分位数换手率对比表.....	23
表 14 资金流因子 IC 对比表（持仓 10 天） .....	23
表 15 资金流因子 IC 对比表（持仓 30 天） .....	23

## 1. 单因子研究流程

股票单因子是股票多因子研究的基础，基于单因子有效性、稳健性的识别及分析是开展多因子研究的前提条件。因此，多因子研究的基础便在于如何衡量因子的有效性，如何筛选出有用的因子。根据因子测试的基本框架，我们在单因子研究方面主要涉及以下几个方面。

### 1.1. 样本空间

样本范围：沪深 A 股

时间周期：2015 年 1 月 1 日至 2021 年 4 月 30 日

为了满足研究需求，并使得研究结果符合投资逻辑，我们在每月最后一个交易日按照设定以下几个规则建立样本空间：

- (1)、剔除截止选股日上市天数不满 60 个交易日的股票；
- (2)、剔除选股日为已退市、暂停上市、退市整理、风险警示这几种状态的股票；
- (3)、剔除截止选股日连续停牌超过 5 个交易日的股票；
- (4)、剔除截止选股日不存在申万一级行业分类标识的股票。

### 1.2. 参数设定

在因子研究过程中，涉及分组、处理方法、换仓周期等设定，具体如下：

- 因子分组数：5 组，行业内按照因子值进行分位数分组
- 标准化方法：因子行业内标准化处理
- 权重处理方法：对数流通市值加权
- 因子换仓周期：10、30 天

### 1.3. 研究对象

本文研究所用个股资金流向数据均来源于同花顺，成交单类型定义如下表所示：

表 1 资金流向类型定义

类型	标准
特大单	成交量在 20 万股及以上，或成交金额在 100 万元及以上的成交单。
大单	成交量在 6 万股到 20 万股之间，或成交金额在 30 万元到 100 万元之间，或成交量占流通盘 0.1% 的成交单。
中单	成交量在 1 万股到 6 万股之间，或成交金额在 5 万元到 30 万元之间的成交单。
小单	成交量在 1 万股及以下，或成交金额在 5 万元及以下的成交单。

资料来源：东海证券研究所，同花顺财经

上述四种类型的成交单组成了市场交易的全部成交单，同时考虑到实际的交易情况，我们在对资金流向成交单类型的研究上暂不考虑“小单”因子。

在因子设计上，我们使用个股过去 20 个交易日的累计资金净流入与当日股票总市值之比作为因子值，为了研究不同类型资金流入情况，设计了四个因子，分别为：资金净流入（特大单净流入+大单净流入）CFLOW\_LB\_20D、特大单资金净流入 CFLOW\_LG\_20D、大单资金净流入 CFLOW\_BG\_20D 和中单资金净流入 CFLOW\_MD\_20D。

下文我们分别对四个资金流因子从因子特征、因子 IC 和因子换手率这三个方面对因子进行分析研究，最终择优选择一个因子进行组合回测。

## 2. 资金净流入因子

### 2.1. 因子特征分析

通过对 CFLOW\_LB\_20D 因子数据在行业内按照因子值的大小进行百分位分组，各分位组因子统计数据如下表所示，由于因子数据在进行数据清洗时采用了 MAD 方法将异常值向第 1、5 两组收缩，进而导致第 1、5 两组的组内标准差相对较大外，其余各组组内分布相对较为均匀，采用分位数分组后样本个数在各组分布相对较为平均。

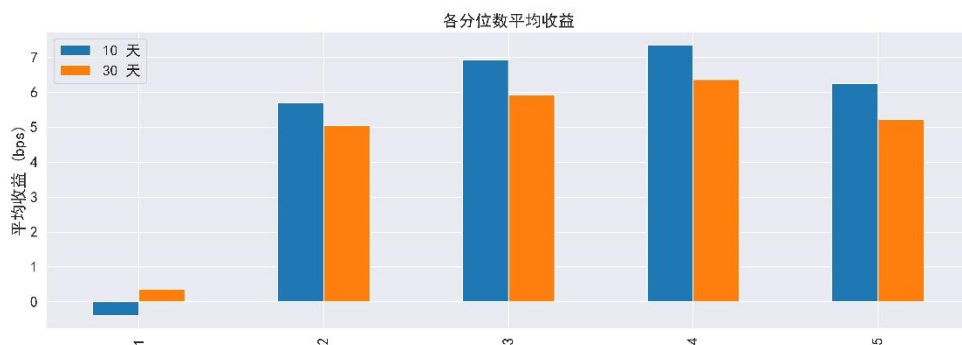
表 2 CFLOW\_LB\_20D 因子分位数统计表

	最小值	最大值	平均值	标准差	样本数占比%
分位数第 1 组	-89.3458	0.2735	-3.4277	3.0512	20.38
分位数第 2 组	-13.5339	1.0476	-1.2745	1.0462	19.82
分位数第 3 组	-12.0630	1.3389	-0.6722	0.6629	19.82
分位数第 4 组	-8.9131	1.8635	-0.2679	0.4092	19.82
分位数第 5 组	-4.4230	47.6494	0.4919	1.1554	20.16

资料来源：东海证券研究所，iFinD

因子各分位数在不同周期下的表现如下图所示，从各分位数对比上看，各分位数的平均收益呈现一定的弱单调性，这说明经过行业处理后的 CFLOW\_LB\_20D 因子越大，其因子未来平均收益相对较高，而且随着持仓周期的增大，除第一分位数外，其他分位数水平下的平均收益也呈现一定的衰减。

图 1 CFLOW\_LB\_20D 因子分位数对比图



资料来源：东海证券研究所，iFinD

从各分组未来 10 天日均累积收益（如下图所示）在整个测试周期上看，高分位组（第 2-5）的累积收益明显高于低分位组，且 2017 和 2018 两年，低分位组的累积收益表现尤为不佳。

图 2 CFLOW\_LB\_20D 因子分位数组组合累积收益图



资料来源：东海证券研究所，iFinD

下图所示为因子值加权组合累积收益走势图，其中每日权重等于因子值与因子值绝对值之和的比，权重值的正负分布表示买入、卖出，根据图形走势，可以观察到做多高分位因子同时做空低分位因子在回测周期内能够取得较为稳定的组合收益。

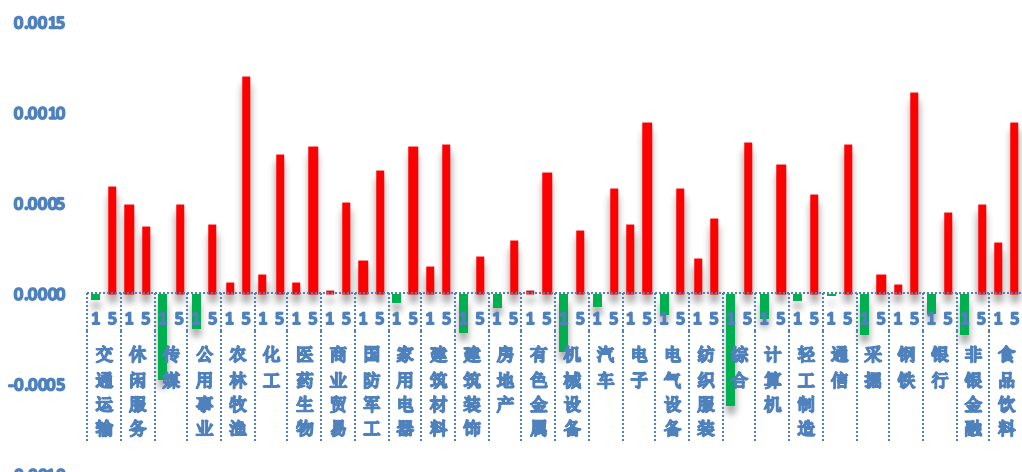
图 3 CFLOW\_LB\_20D 因子加权组合累计收益图



资料来源：东海证券研究所，iFinD

从行业角度来看高低分位数的差异，在 10 天的换仓期下各行业的分位数平均收益如下图所示，除休闲服务行业外，其他行业高分位数的平均收益均高于低分位数收益，传媒、农林牧渔、医药生物、家用电器、综合、通信、钢铁和非银金融这几个行业的分位数平均收益的差异性相对较大。

图 4 CFLOW\_LB\_20D 因子按行业分位数收益对比图（持仓 10 天）



资料来源：东海证券研究所，iFinD

## 2.2. 因子 IC 分析



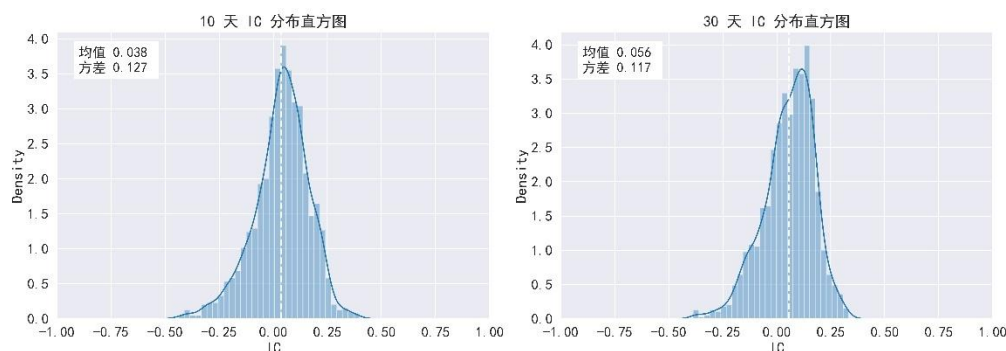
如下 CFLOW\_LB\_20D 因子 IC 统计表所示，随着持仓周期的增加，因子 IC 增大，在持仓周期为 30 天的情况下，IC 均值为 0.056，IR 为 0.48，从 IC 的 T 检验以及 IC 分布直方图上看，IC 的分布相对集中，IC 统计的稳定性较强。

表 3 CFLOW\_LB\_20D 因子 IC 统计表

	10 天	30 天
IC Mean	0.0378	0.0559
IC Std.	0.1268	0.1167
IR	0.2977	0.4785
t-stat(IC)	11.5704	18.6014
p-value(IC)	0.0000	0.0000
IC Skew	-0.5307	-0.6200
IC Kurtosis	0.6806	0.4887

资料来源：东海证券研究所，iFinD

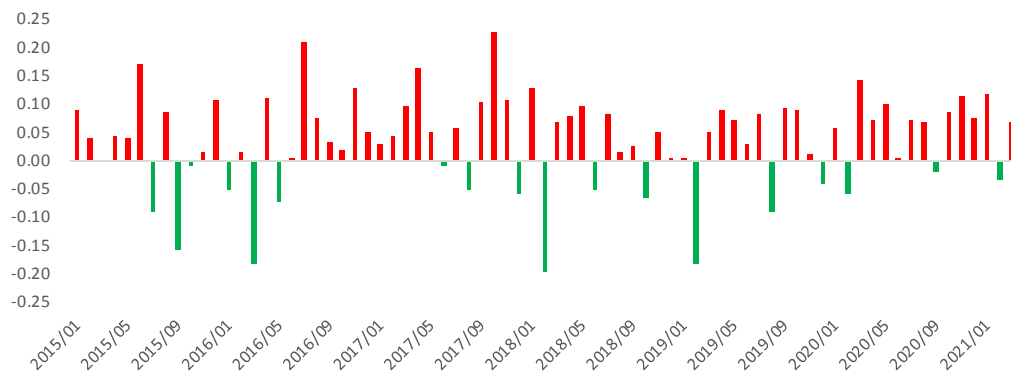
图 5 CFLOW\_LB\_20D 因子 IC 分布直方图



资料来源：东海证券研究所，iFinD

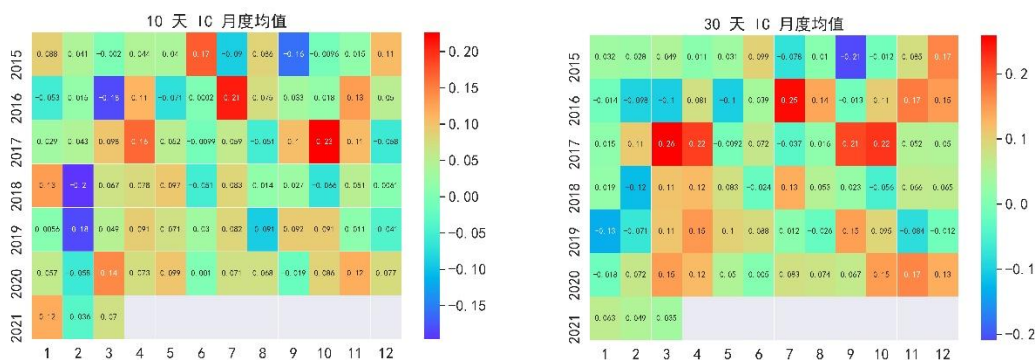
根据因子 IC 的计算公式，日频率因子的 IC 即为其每天的因子值与未来一段时间收益率的截面相关系数，测试周期的 IC 均值即为其每日 IC 的平均值，因此可以进一步的从月度或者年度的时间视角考察 CFLOW\_LB\_20D 因子的区间表现。如下两图分别为 10 天持仓周期的因子月度平均 IC 及因子月历图（月度均值热力图），结合两图可以看出，全市场 CFLOW\_LB\_20D 因子 IC 在测试周期内，IC 为正数的频率相对较高，IC 为正则表示当期因子与未来收益呈现一定的正相关性，IC 越大则其相关性越强，持有较高正 IC 组合的未来获得较高收益的可能性越大。从因子月度 IC 均值热力图上看，CFLOW\_LB\_20D 因子并无明显月历效应。

图 6 CFLOW\_LB\_20D 因子 IC 月度平均 (持仓 10 天)



资料来源：东海证券研究所，iFinD

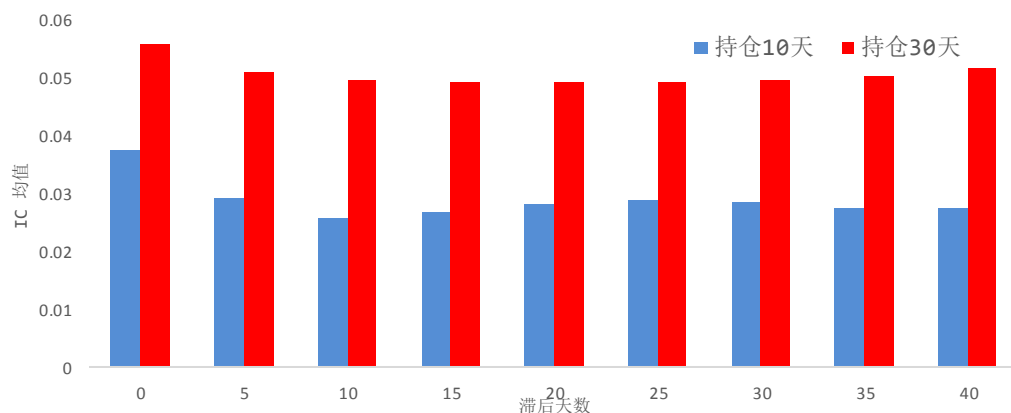
图 7 CFLOW\_LB\_20D 因子 IC 月度均值热力图



资料来源：东海证券研究所，iFinD

因子作为反映股票某一特征的信息，同样具有时效性，在一般情况下，随着时间的滞后，因子所反映信息的价值也将逐步下降。因子 IC 的时间衰减可以衡量一个因子对未来预测能力的持续性，衰减度越高，说明预测稳定性越差。下图所示为 10 和 30 天的预测(持仓)周期下，CFLOW\_LB\_20D 因子的平均衰减情况，可以看出，长周期的衰减幅度慢于短周期的衰减。

图 8 CFLOW\_LB\_20D 因子 IC 衰减图



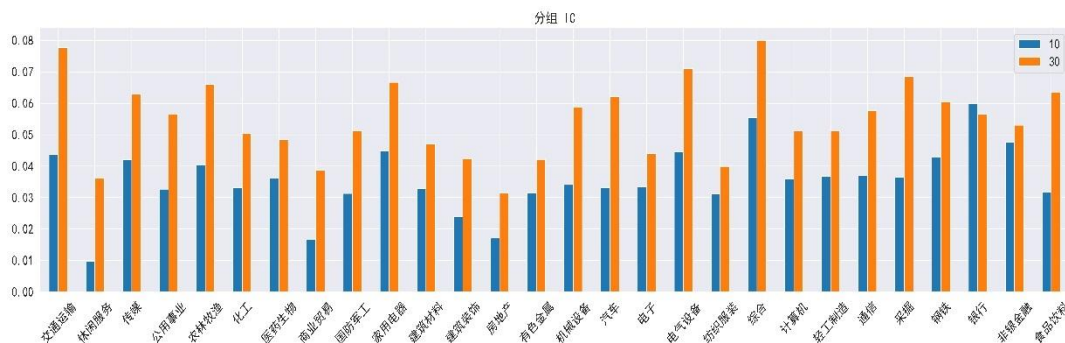
资料来源：东海证券研究所，iFinD

此外，从 CFLOW\_LB\_20D 因子行业的角度来看行业 IC，可以反映该因子在各个行业的预测效果，如下图所示各个行业在不同预测周期下的平均 IC，总体来看，综合、交通运输、



电气设备、采掘、家用电器和农林牧渔行业的因子 IC 相对较高，在这些行业里持有高 CFLOW\_LB\_20D 因子的股票组合收益相对较高。

图 9 CFLOW\_LB\_20D 因子 IC 行业均值



资料来源：东海证券研究所，iFinD

## 2.3. 因子换手率分析

因子换手率是因子稳定性和持续性的一种反映，如果因子表现不稳定或者随时间衰减过快时，就可能导致组合出现较高的换手率，由换手率引致的交易成本提升会较大幅度的侵蚀模型的盈利能力，因此需要从换手率的角度对因子进行考察。

CFLOW\_LB\_20D 因子在不同因子分位数及持仓周期下的平均换手率如下表所示，此处定义换手率为当期在指定分位数中的因子不在前 N 期分位数中的比例，从组别上看，高低分位组的换手率相对较低，中间分位组或因数据集中度较高而致在组间相对频繁变化。从时间跨度上看，时间跨度越长，换手率相对越高，且因子的自相关性越低。

表 4 CFLOW\_LB\_20D 因子分位数换手率对比表

	10 天	30 天
第 1 组平均换手率	0.3844	0.5949
第 2 组平均换手率	0.6388	0.7598
第 3 组平均换手率	0.6741	0.7660
第 4 组平均换手率	0.6270	0.7184
第 5 组平均换手率	0.4963	0.7288
因子自相关系数	0.6401	0.3047

资料来源：东海证券研究所，iFinD

## 3. 特大单资金净流入因子

### 3.1. 因子特征分析

通过对 CFLOW\_LG\_20D 因子数据在行业内按照因子值的大小进行百分位分组，各分位组因子统计数据如下表所示，由于因子数据在进行数据清洗时采用了 MAD 方法将异常值向第 1、5 两组收缩，进而导致第 1、5 两组的组内标准差相对较大外，其余各组组内分布相对较为均匀，采用分位数分组后样本个数在各组分布相对较为平均。

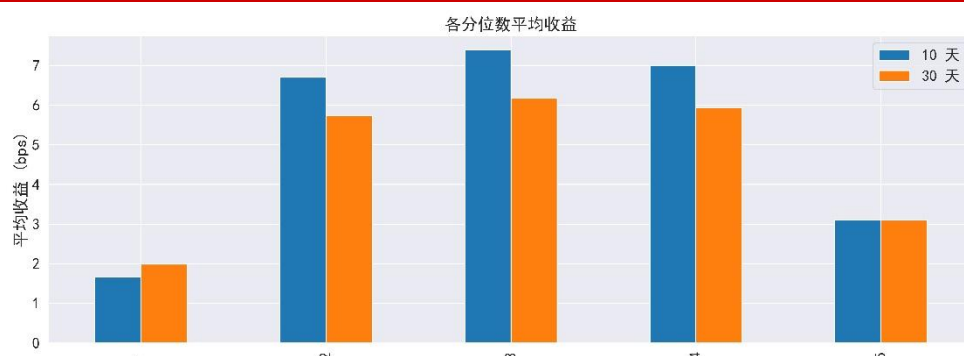
表 5 CFLOW\_LG\_20D 因子分位数统计表

	最小值	最大值	平均值	标准差	样本数占比%
分位数第 1 组	-71.1211	0.7214	-1.9771	2.2699	20.38
分位数第 2 组	-10.7566	1.5059	-0.5968	0.6757	19.82
分位数第 3 组	-6.8455	2.7570	-0.2454	0.3850	19.87
分位数第 4 组	-4.0569	3.3738	-0.0058	0.2317	19.76
分位数第 5 组	-1.8092	48.3720	0.7662	1.2694	20.16

资料来源：东海证券研究所，iFinD

因子各分位数在不同测试周期下的表现如下图所示，从各分位数对比上看，各分位数的平均收益不存在明显的单调性，分位数与其平均收益呈现中间高、两端低的特征，但是随着持仓周期的增大，其平均收益也呈现一定的衰减。

图 10 CFLOW\_LG\_20D 因子分位数对比图



资料来源：东海证券研究所，iFinD

从各分组未来 10 天日均累积收益(如下图所示)在整个测试周期上看，中间分位组(第 2、3、4 组)的累积收益明显高于两端分位组(第 1、5 组)。

图 11 CFLOW\_LG\_20D 因子分位数组组合累积收益图



资料来源：东海证券研究所，iFinD

下图所示为因子值加权组合累积收益走势图，其中每日权重等于因子值与因子值绝对值之和的比，权重值的正负分布表示买入、卖出，根据图形走势，可以观察到多空组合累积收益走势图波动较大。

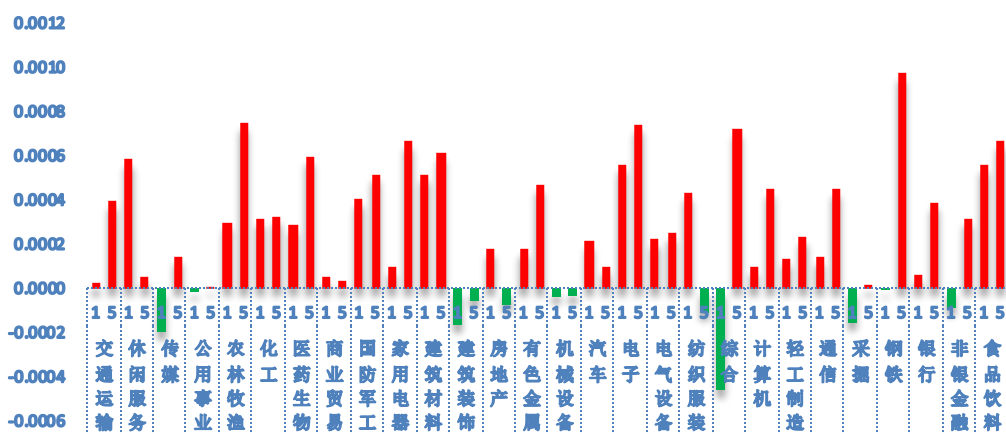
图 12 CFLOW\_LG\_20D 因子加权组合累计收益图



资料来源：东海证券研究所，iFinD

从行业角度来看高低分位数的差异,在 10 天的换仓期下各行业的分位数平均收益如下图所示,除休闲服务、公用事业、化工、商业贸易、机械设备、电气设备和纺织服装行业外其他行业高分位数的平均收益均明显高于低分位数收益,综合、钢铁、银行、非银金融、家用电器和传媒这几个行业的分位数平均收益的差异性相对较大。

图 13 CFLOW\_LG\_20D 因子按行业分位数收益对比图（持仓 10 天）



资料来源：东海证券研究所，iFinD

### 3.2. 因子 IC 分析

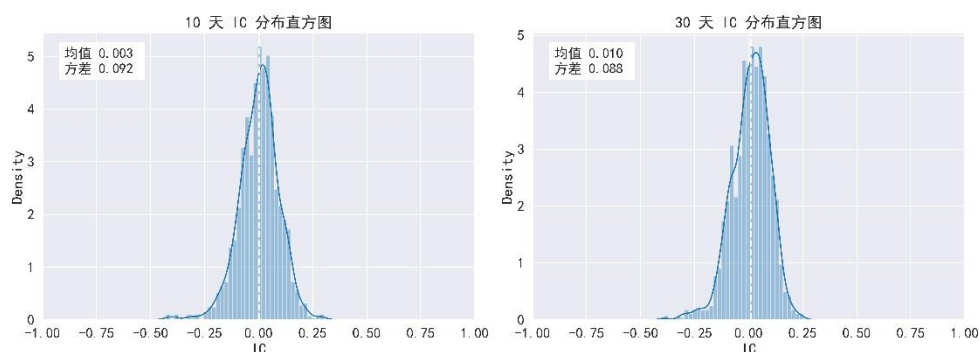
如下 CFLOW\_LG\_20D 因子 IC 统计表所示,随着持仓周期的增加,因子 IC 增大,在持仓周期为 30 天的情况下,IC 均值为 0.01,IR 为 0.11,在测试时间区间整体上看,因子 IC 均值不高。同时,从 IC 的 T 检验以及 IC 分布直方图上看,IC 的分布相对集中。

表 6 CFLOW\_LG\_20D 因子 IC 统计表

	10 天	30 天
IC Mean	0.0026	0.0098
IC Std.	0.0917	0.0876
IR	0.0284	0.1116
t-stat(IC)	1.1031	4.3365
p-value(IC)	0.2701	0.0000
IC Skew	-0.4625	-0.5755
IC Kurtosis	1.5928	0.9552

资料来源：东海证券研究所，iFinD

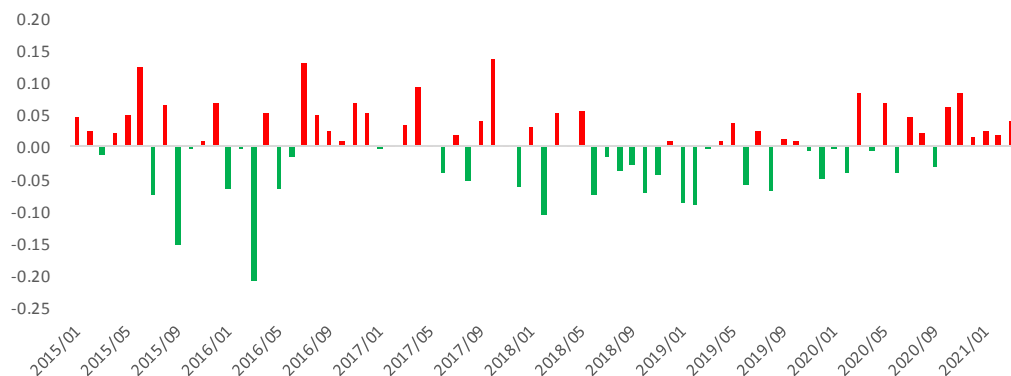
图 14 CFLOW\_LG\_20D 因子 IC 分布直方图



资料来源：东海证券研究所，iFinD

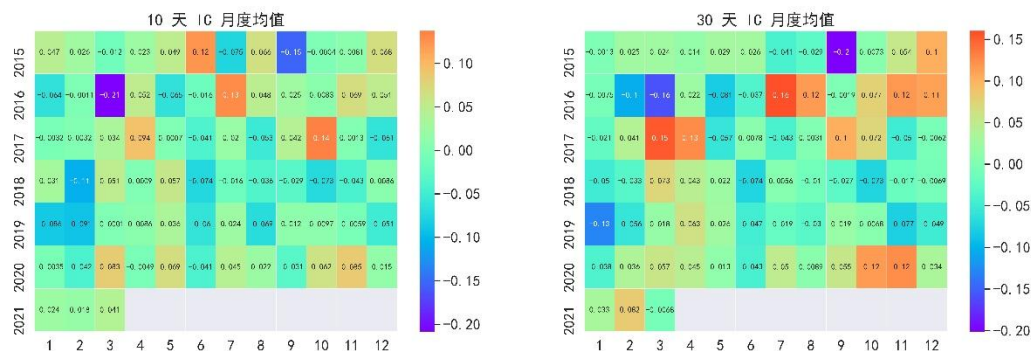
如下两图分别为 10 天持仓周期的 CFLOW\_LG\_20D 因子月度平均 IC 及因子月历图(月度均值热力图)，结合两图可以看出，全市场 CFLOW\_LG\_20D 因子 IC 在测试周期内，IC 为正数的频率略多，IC 为正则表示当期因子与未来收益呈现一定的正相关性，IC 越大则其相关性越强，持有较高正 IC 组合的未来获得较高收益的可能性越大。从因子月度 IC 均值热力图上看，CFLOW\_LG\_20D 因子无明显月历效应。

图 15 CFLOW\_LG\_20D 因子 IC 月度平均（持仓 10 天）



资料来源：东海证券研究所，iFinD

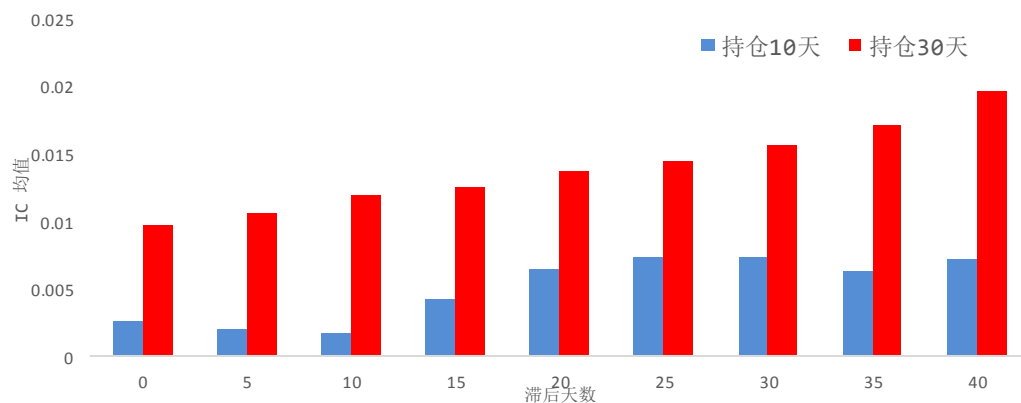
图 16 CFLOW\_LG\_20D 因子 IC 月度均值热力图



资料来源：东海证券研究所，iFinD

下图所示为 10 和 30 天的预测(持仓)周期下，CFLOW\_LG\_20D 因子的平均衰减情况，该因子未能随时间的滞后而呈现衰减的态势，在一定程度上源于持仓周期的设置。

图 17 CFLOW\_LG\_20D 因子 IC 衰减图



资料来源：东海证券研究所，iFinD

此外，从 CFLOW\_LG\_20D 因子行业的角度来看行业 IC，可以反映该因子在各个行业的预测效果，如下图所示各个行业在不同预测周期下的平均 IC，总体来看，钢铁、银行、综合、和家用电器行业的因子 IC 相对较高，在这些行业里持有高 CFLOW\_LG\_20D 因子的股票组合收益相对较高。

图 18 CFLOW\_LG\_20D 因子 IC 行业均值



资料来源：东海证券研究所，iFinD

### 3.3. 因子换手率分析

CFLOW\_LG\_20D 因子在不同因子分位数及持仓周期下的平均换手率如下表所示，此处定义换手率为当期在指定分位数中的因子不在前 N 期分位数中的比例，从组别上看，高低分位组的换手率相对较低，中间分位组或因数据集中度较高而致在组间相对频繁变化。从时间跨度上看，时间跨度越长，换手率相对越高，且因子的自相关性越低。

表 7 CFLOW\_LG\_20D 因子分位数换手率对比表

	10 天	30 天
第 1 组平均换手率	0.4164	0.6252
第 2 组平均换手率	0.6534	0.7710
第 3 组平均换手率	0.6685	0.7600
第 4 组平均换手率	0.6295	0.7266
第 5 组平均换手率	0.5089	0.7530
因子自相关系数	0.5595	0.1932

资料来源：东海证券研究所，iFinD

## 4. 大单资金净流入因子

### 4.1. 因子特征分析

通过对 CFLOW\_BG\_20D 因子数据在行业内按照因子值的大小进行百分位分组，各分位组因子统计数据如下表所示，由于因子数据在进行数据清洗时采用了 MAD 方法将异常值向第 1、5 两组收缩，进而导致第 1、5 两组的组内标准差相对较大外，其余各组组内分布相对较为均匀，采用分位数分组后样本个数在各组分布相对较为平均。

表 8 CFLOW\_BG\_20D 因子分位数统计表

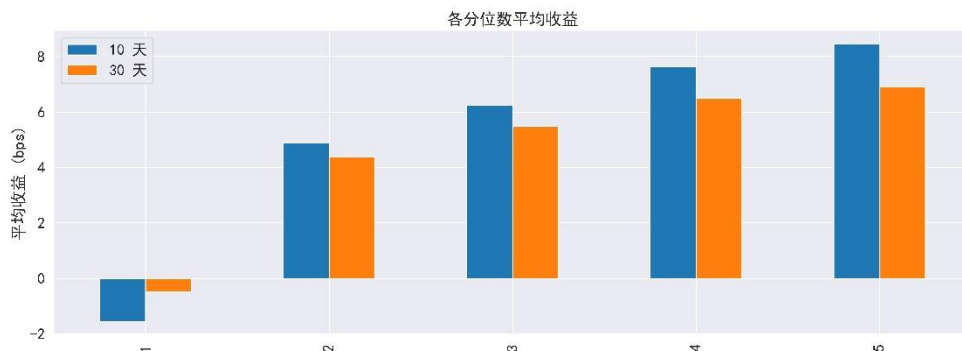
	最小值	最大值	平均值	标准差	样本数占比%
分位数第 1 组	-23.9036	0.0086	-1.9479	1.4309	20.38
分位数第 2 组	-6.7336	0.0418	-0.7547	0.4571	19.82
分位数第 3 组	-5.5712	0.2924	-0.4078	0.2948	19.82
分位数第 4 组	-4.0402	0.9847	-0.1739	0.1879	19.82
分位数第 5 组	-2.3415	17.0248	0.1996	0.4280	20.16

资料来源：东海证券研究所，iFinD

因子各分位数在不同测试周期下的表现如下图所示，从各分位数对比上看，各分位数的平均收益呈现单调性，这说明经过行业处理后的 CFLOW\_BG\_20D 因子越大，其因子未来平均收益相对较高，而且随着持仓周期的增大，其平均收益也呈现一定的衰减。



图 19 CFLOW\_BG\_20D 因子分位数对比图



资料来源：东海证券研究所，iFind

从各分组未来 10 天日均累积收益（如下图所示）在整个测试周期上看，高分位组（第 2-5）的累积收益明显高于低分位组。

图 20 CFLOW\_BG\_20D 因子分位数组合累积收益图



资料来源：东海证券研究所，iFind

下图所示为因子值加权组合累积收益走势图，其中每日权重等于因子值与因子值绝对值之和的比，权重值的正负分布表示买入、卖出，根据图形走势，可以观察到做多高分位因子同时做空低分位因子存在较为稳定和持续的组合收益。

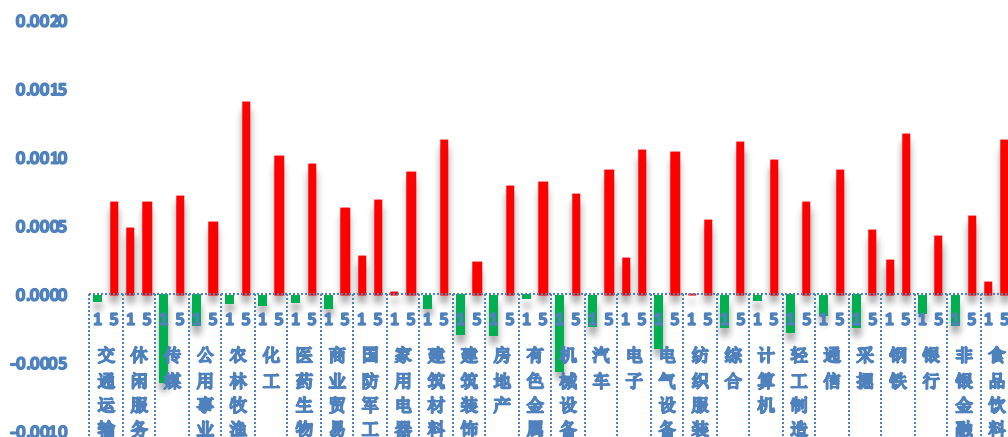
图 21 CFLOW\_BG\_20D 因子加权组合累计收益图



资料来源：东海证券研究所，iFind

从行业角度来看高低分位数的差异，在 10 天的换仓期下各行业的分位数平均收益如下图所示，行业高分位数的平均收益均高于低分位数收益，农林牧渔、电气设备、机械设备、建筑材料和传媒这几个行业的分位数平均收益的差异性相对较大。

图 22 CFLOW\_BG\_20D 因子按行业分位数收益对比图（持仓 10 天）



资料来源：东海证券研究所，iFinD

## 4.2. 因子 IC 分析

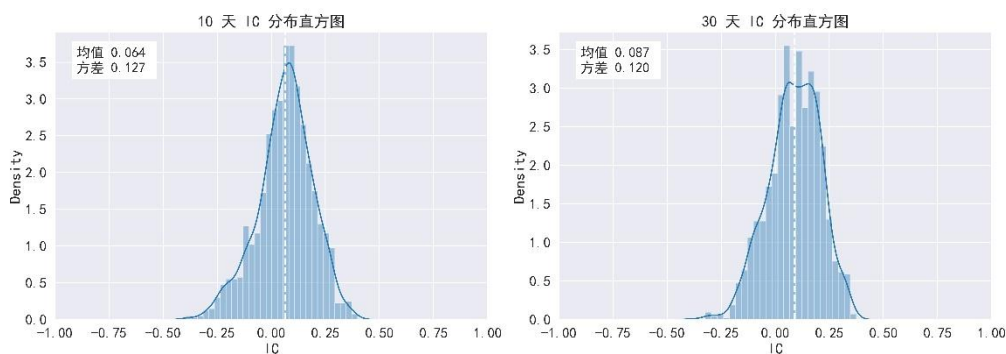
如下 CFLOW\_BG\_20D 因子 IC 统计表所示，随着持仓周期的增加，因子 IC 增大，在持仓周期为 30 天的情况下，IC 均值为 0.087，IR 为 0.73，在测试时间区间整体上看，因子 IC 均值较高。同时，从 IC 的 T 检验以及 IC 分布直方图上看，IC 的分布相对集中，IC 统计的稳定性较强。

表 9 CFLOW\_BG\_20D 因子 IC 统计表

	10 天	30 天
IC Mean	0.0638	0.0874
IC Std.	0.1273	0.1199
IR	0.5011	0.7294
t-stat(IC)	19.4788	28.3519
p-value(IC)	0.0000	0.0000
IC Skew	-0.3930	-0.3363
IC Kurtosis	0.2475	-0.0432

资料来源：东海证券研究所，iFinD

图 23 CFLOW\_BG\_20D 因子 IC 分布直方图

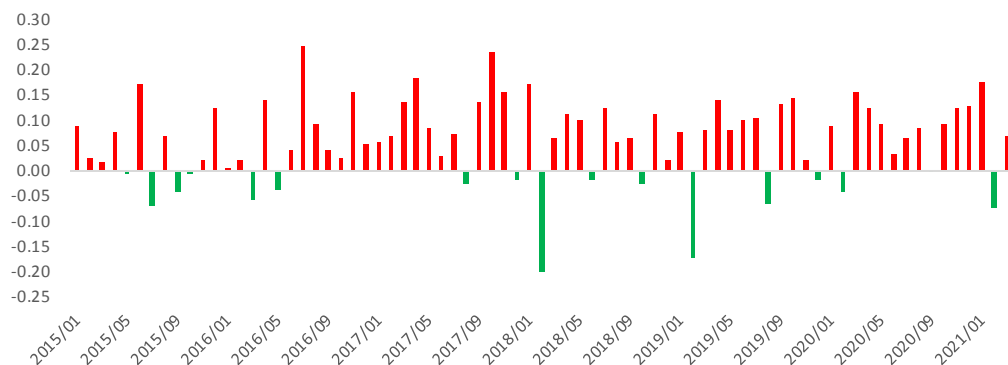


资料来源：东海证券研究所，iFinD

如下两图分别为 10 天持仓周期的 CFLOW\_BG\_20D 因子月度平均 IC 及因子月历图(月度均值热力图)，结合两图可以看出，全市场 CFLOW\_BG\_20D 因子 IC 在测试周期内，IC 为正

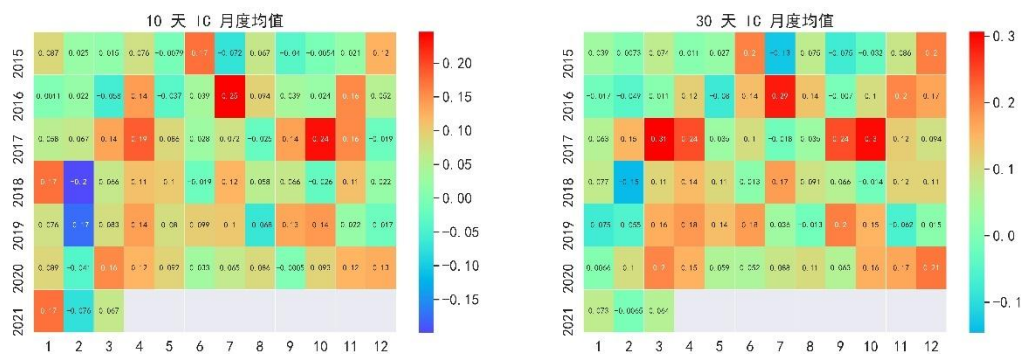
数的频率相对较高, IC 为正则表示当期因子与未来收益呈现一定的正相关性, IC 越大则其相关性越强, 持有较高正 IC 组合的未来获得较高收益的可能性越大。从因子月度 IC 均值热力图上看, CFLOW\_BG\_20D 因子无明显月历效应。

图 24 CFLOW\_BG\_20D 因子 IC 月度平均 (持仓 10 天)



资料来源: 东海证券研究所, iFind

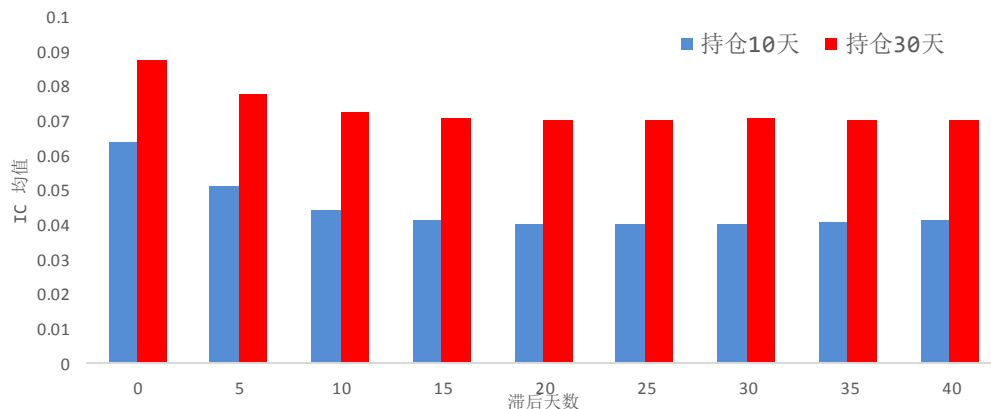
图 25 CFLOW\_BG\_20D 因子 IC 月度均值热力图



资料来源: 东海证券研究所, iFind

下图所示为 10 和 30 天的预测(持仓)周期下, CFLOW\_BG\_20D 因子的平均衰减情况, 可以看出, 因子衰减速度相对较慢, 而且长周期的衰减幅度快于短周期的衰减。

图 26 CFLOW\_BG\_20D 因子 IC 衰减图

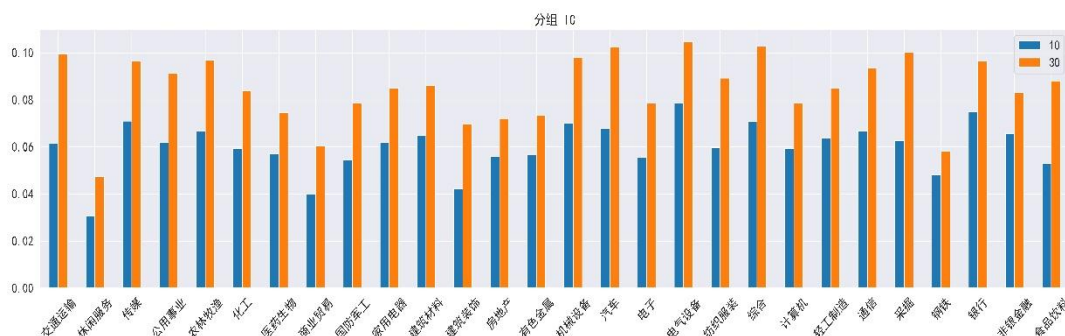


资料来源: 东海证券研究所, iFind

此外, 从 CFLOW\_BG\_20D 因子行业的角度来看行业 IC, 可以反映该因子在各个行业的预测效果, 如下图所示各个行业在不同预测周期下的平均 IC, 总体来看, 综合、汽车、采

掘、电气设备及机械设备行业的因子 IC 相对较高，在这些行业里持有高 CFLOW\_BG\_20D 因子的股票组合收益相对较高。

图 27 CFLOW\_BG\_20D 因子 IC 行业均值



资料来源：东海证券研究所，iFind

### 4.3. 因子换手率分析

CFLOW\_BG\_20D 因子在不同因子分位数及持仓周期下的平均换手率如下表所示，此处定义换手率为当期在指定分位数中的因子不在前 N 期分位数中的比例，从组别上看，高低分位组的换手率相对较低，中间分位组或因数据集中度较高而致在组间相对频繁变化。从时间跨度上看，时间跨度越长，换手率相对越高，且因子的自相关性越低。

表 10 CFLOW\_BG\_20D 因子分位数换手率对比表

	10 天	30 天
第 1 组平均换手率	0.3588	0.5715
第 2 组平均换手率	0.6295	0.7549
第 3 组平均换手率	0.6704	0.7665
第 4 组平均换手率	0.6286	0.7197
第 5 组平均换手率	0.4696	0.7015
因子自相关系数	0.6941	0.3686

资料来源：东海证券研究所，iFind

## 5. 中单资金净流入因子

### 5.1. 因子特征分析

通过对 CFLOW\_MD\_20D 因子数据在行业内按照因子值的大小进行百分位分组，各分位组因子统计数据如下表所示，由于因子数据在进行数据清洗时采用了 MAD 方法将异常值向第 1、5 两组收缩，进而导致第 1、5 两组的组内标准差相对较大外，其余各组组内分布相对较为均匀，采用分位数分组后样本个数在各组分布相对较为平均。

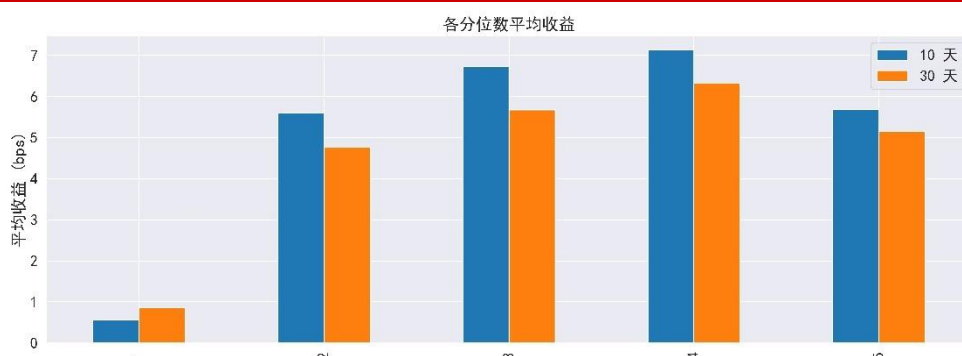
表 11 CFLOW\_MD\_20D 因子分位数统计表

	最小值	最大值	平均值	标准差	样本数占比%
分位数第 1 组	-41.0650	0.7932	-1.4063	1.2516	20.38
分位数第 2 组	-2.4796	1.2383	-0.4557	0.2337	19.82
分位数第 3 组	-1.5081	3.0599	-0.1706	0.1806	19.82
分位数第 4 组	-0.7315	4.9038	0.0606	0.2479	19.82
分位数第 5 组	-0.4887	26.0037	0.6206	0.7931	20.16

资料来源：东海证券研究所，iFinD

因子各分位数在不同测试周期下的表现如下图所示，从各分位数对比上看，各分位数的平均收益呈现一定的弱单调性，这说明经过行业处理后的 CFLOW\_MD\_20D 因子越大，其因子未来平均收益相对较高，而且随着换仓周期的增大，其平均收益也呈现一定的衰减。

图 28 CFLOW\_MD\_20D 因子分位数对比图



资料来源：东海证券研究所，iFinD

从各分组未来 10 天日均累积收益（如下图所示）在整个测试周期上看，高分位组（第 2-5）的累积收益明显高于低分位组。

图 29 CFLOW\_MD\_20D 因子分位数组合累积收益图



资料来源：东海证券研究所，iFinD

下图所示为因子值加权组合累积收益走势图，其中每日权重等于因子值与因子值绝对值之和的比，权重值的正负分布表示买入、卖出，根据图形走势，可以观察到做多高分位因子同时做空低分位因子在 2016 年以来有较稳定的组合收益。

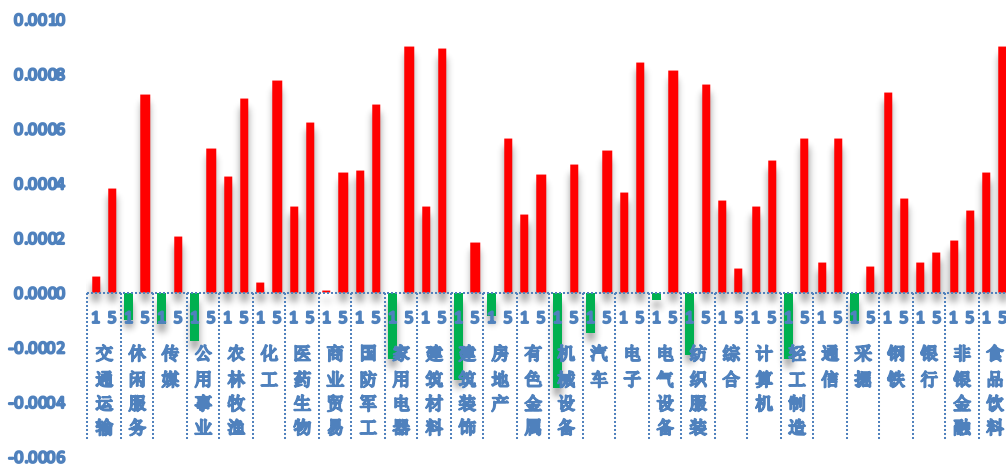
图 30 CFLOW\_MD\_20D 因子加权组合累计收益图



资料来源：东海证券研究所，iFinD

从行业角度来看高低分位数的差异,在 10 天的换仓期下各行业的分位数平均收益如下图所示,除综合和钢铁这两个行业外,其他行业高分位数的平均收益均高于低分位数收益,休闲服务、公用事业、家用电器、机械设备、纺织服装和轻工制造这几个行业的分位数平均收益的差异性相对较大。

图 31 CFLOW\_MD\_20D 因子按行业分位数收益对比图 (持仓 10 天)



资料来源：东海证券研究所，iFinD

## 5.2. 因子 IC 分析

如下 CFLOW\_MD\_20D 因子 IC 统计表所示,随着持仓周期的增加,因子 IC 增大,在持仓周期为 30 天的情况下,IC 均值为 0.055,IR 为 0.65,在测试时间区间整体上看,因子 IC 均值较高。同时,从 IC 的 T 检验以及 IC 分布直方图上看,IC 的分布相对集中,IC 统计的稳定性较强。

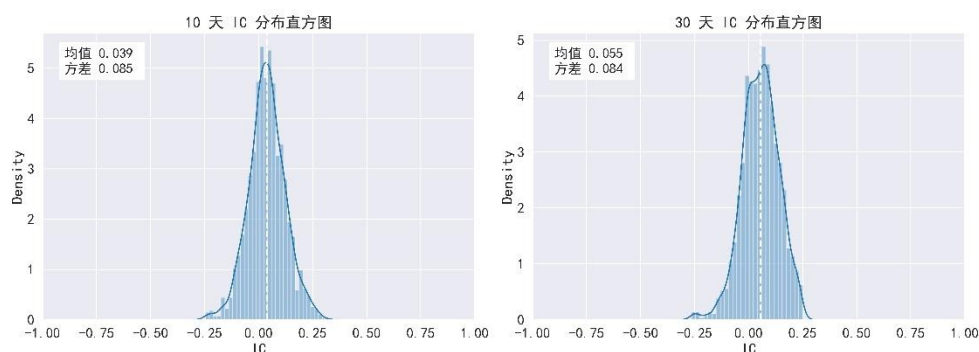


表 12 CFLOW\_MD\_20D 因子 IC 统计表

	10 天	30 天
IC Mean	0.0388	0.0549
IC Std.	0.0845	0.0844
IR	0.4588	0.6508
t-stat(IC)	17.8329	25.2964
p-value(IC)	0.0000	0.0000
IC Skew	-0.0154	-0.2941
IC Kurtosis	0.4165	0.3905

资料来源：东海证券研究所，iFinD

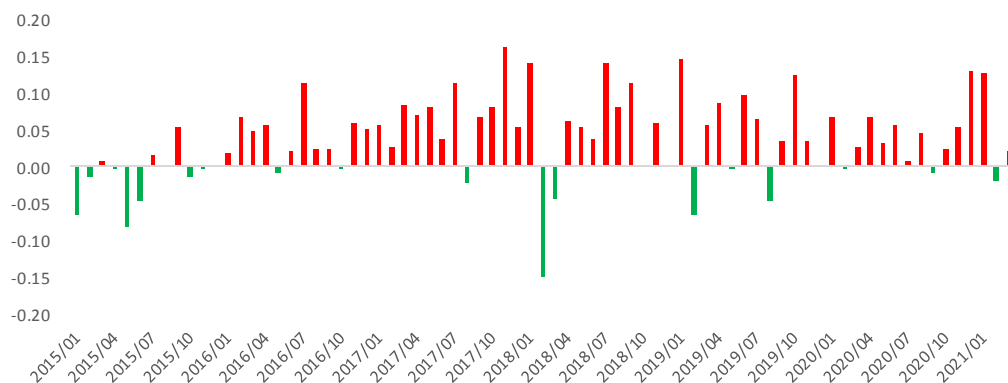
图 32 CFLOW\_MD\_20D 因子 IC 分布直方图



资料来源：东海证券研究所，iFinD

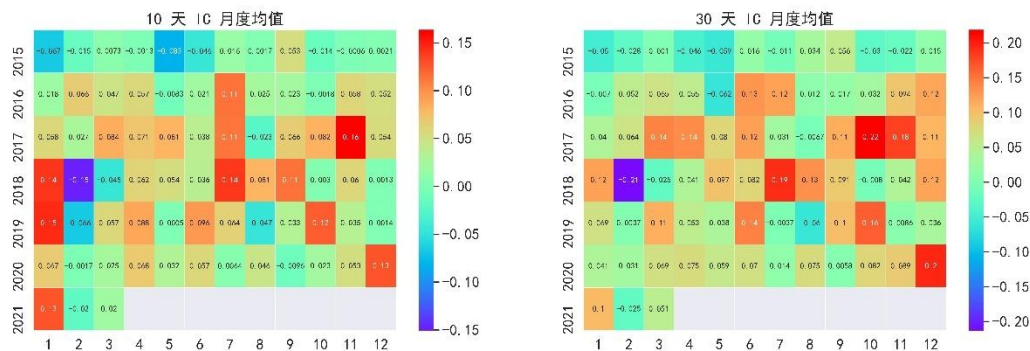
如下两图分别为 10 天持仓周期的 CFLOW\_MD\_20D 因子月度平均 IC 及因子月历图(月度均值热力图)，结合两图可以看出，全市场 CFLOW\_MD\_20D 因子 IC 在测试周期内，IC 为正数的频率相对较高，IC 为正则表示当期因子与未来收益呈现一定的正相关性，IC 越大则其相关性越强，持有较高正 IC 组合的未来获得较高收益的可能性越大。从因子月度 IC 均值热力图上看，CFLOW\_MD\_20D 因子无明显月历效应。

图 33 CFLOW\_MD\_20D 因子 IC 月度平均（持仓 10 天）



资料来源：东海证券研究所，iFinD

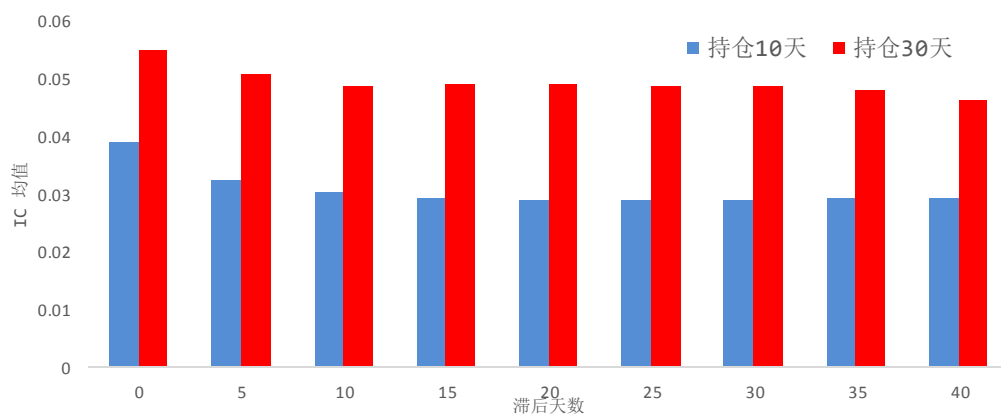
图 34 CFLOW\_MD\_20D 因子 IC 月度均值热力图



资料来源：东海证券研究所，iFinD

下图所示为 10 和 30 天的预测(持仓)周期下，CFLOW\_MD\_20D 因子的平均衰减情况，可以看出，因子衰减速度相对较慢，而且长周期的衰减幅度快于短周期的衰减。

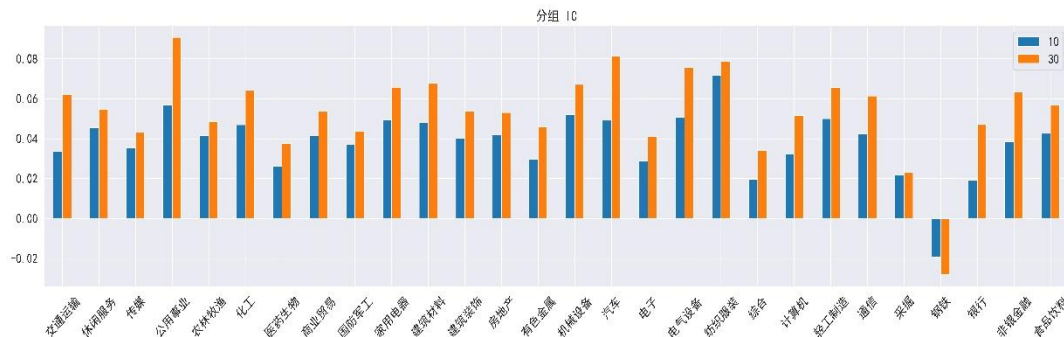
图 35 CFLOW\_MD\_20D 因子 IC 衰减图



资料来源：东海证券研究所，iFinD

此外，从 CFLOW\_MD\_20D 因子行业的角度来看行业 IC，可以反映该因子在各个行业的预测效果，如下图所示各个行业在不同预测周期下的平均 IC，总体来看，公用事业、汽车、纺织服装、电气设备及机械设备行业的因子 IC 相对较高，在这些行业里持有高 CFLOW\_MD\_20D 因子的股票组合收益相对较高。

图 36 CFLOW\_MD\_20D 因子 IC 行业均值



资料来源：东海证券研究所，iFinD

### 5.3. 因子换手率分析

CFLOW\_MD\_20D 因子在不同因子分位数及持仓周期下的平均换手率如下表所示，此处定义换手率为当期在指定分位数中的因子不在前 N 期分位数中的比例，从组别上看，高低分位组的换手率相对较低，中间分位组或因数据集中度较高而致在组间相对频繁变化。从时间跨度上看，时间跨度越长，换手率相对越高，且因子的自相关性越低。

表 13 CFLOW\_MD\_20D 因子分位数换手率对比表

	10 天	30 天
第 1 组平均换手率	0.3824	0.5865
第 2 组平均换手率	0.6362	0.7561
第 3 组平均换手率	0.6660	0.7559
第 4 组平均换手率	0.6332	0.7295
第 5 组平均换手率	0.4482	0.6700
因子自相关系数	0.6661	0.3353

资料来源：东海证券研究所，iFinD

## 6. 因子对比与模拟组合

### 6.1. 因子 IC 对比

如下表所示，分别为四个资金流因子（资金净流入、特大单资金净流入、大单资金净流入和中单资金净流入）在持仓 10 天和 30 天情况下的因子 IC 数据对比。

表 14 资金流因子 IC 对比表（持仓 10 天）

	CFLOW_LB_20D	CFLOW_LG_20D	CFLOW_BG_20D	CFLOW_MD_20D
IC Mean	0.0378	0.0026	0.0638	0.0388
IC Std.	0.1268	0.0917	0.1273	0.0845
IR	0.2977	0.0284	0.5011	0.4588
t-stat(IC)	11.5704	1.1031	19.4788	17.8329
p-value(IC)	0	0.2701	0	0
IC Skew	-0.5307	-0.4625	-0.393	-0.0154
IC Kurtosis	0.6806	1.5928	0.2475	0.4165

资料来源：东海证券研究所，iFinD

表 15 资金流因子 IC 对比表（持仓 30 天）

	CFLOW_LB_20D	CFLOW_LG_20D	CFLOW_BG_20D	CFLOW_MD_20D
IC Mean	0.0559	0.0098	0.0874	0.0549
IC Std.	0.1167	0.0876	0.1199	0.0844
IR	0.4785	0.1116	0.7294	0.6508
t-stat(IC)	18.6014	4.3365	28.3519	25.2964
p-value(IC)	0	0	0	0
IC Skew	-0.62	-0.5755	-0.3363	-0.2941
IC Kurtosis	0.4887	0.9552	-0.0432	0.3905

资料来源：东海证券研究所，iFinD

根据上述因子在不同持仓周期下的 IC、IR 等统计指标的对比结果，总体上看，CFLOW\_BG\_20D 因子在持仓 10 天和 30 天天下的 IC、IR 表现均强于另外三个因子。

## 6.2. 因子历史组合回测

针对资金流入因子中大单资金净流入因子具有较好的 IC 表现，我们在中证 500 指数成分中构建相应的大单资金净流入因子组合，进而衡量该因子在中证 500 指数上的表现。

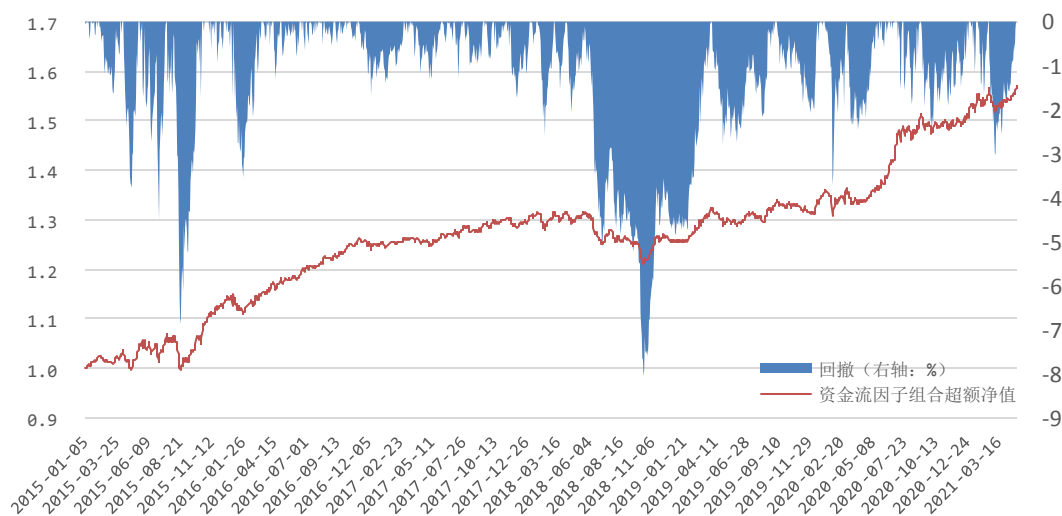
其中，历史模拟组合周期为 2015 年 1 月 5 日至 2021 年 4 月 30 日，换仓频率为 30 个交易日，交易手续费为 0.1%（单边计算），每期组合包含 150 只股票，组合中行业权重与中证 500 指数行业权重相匹配，组合内行业的个股按照对数流通市值赋权。纯因子组合表现如下图所示：

图 37 资金流因子选股历史回测走势对比图



资料来源：东海证券研究所，iFinD

图 38 资金流因子选股历史回测超额净值走势图



资料来源：东海证券研究所，iFinD

纯资金流向因子（不考虑因子择时及其他类型因子收益）超额收益在组合历史回测期间内实现 56.76% 的超额收益，年化收益率为 7.57%，年化波动率为 5.76%，Sharpe 比率为 1.31，超额收益净值最大历史回撤为 8.05%，Calmar 比率为 0.94。

## 分析师简介:

曾磊，数量经济学硕士，东海证券研究所研究员，证券、期货从业10年，具有丰富的量化研究、金融建模经验。

## 附注:

### 一、市场指数评级

看多——未来6个月内上证综指上升幅度达到或超过20%

看平——未来6个月内上证综指波动幅度在-20%—20%之间

看空——未来6个月内上证综指下跌幅度达到或超过20%

### 二、行业指数评级

超配——未来6个月内行业指数相对强于上证指数达到或超过10%

标配——未来6个月内行业指数相对上证指数在-10%—10%之间

低配——未来6个月内行业指数相对弱于上证指数达到或超过10%

### 三、公司股票评级

买入——未来6个月内股价相对强于上证指数达到或超过15%

增持——未来6个月内股价相对强于上证指数在5%—15%之间

中性——未来6个月内股价相对上证指数在-5%—5%之间

减持——未来6个月内股价相对弱于上证指数5%—15%之间

卖出——未来6个月内股价相对弱于上证指数达到或超过15%

### 四、风险提示

本报告所载的全部内容只提供给客户做参考之用，并不构成对客户的投资建议，并非作为买卖、认购证券或其它金融工具的邀请或保证，建议客户如有任何疑问应当咨询独立财务顾问并独自进行投资判断。

### 五、免责声明

本报告基于本公司研究所及研究人员认为可信的公开资料或实地调研的资料，但对这些信息的真实性、准确性和完整性不做任何保证。本报告仅反映研究员个人出具本报告当时的分析和判断，并不代表东海证券股份有限公司，或任何其附属或联营公司的立场，本公司可能发表其他与本报告所载资料不一致及有不同结论的报告。本报告可能因时间等因素的变化而变化从而导致与事实不完全一致，敬请关注本公司就同一主题所出具的相关后续研究报告及评论文章。在法律允许的情况下，本公司的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能为这些公司正在提供或争取提供多种金融服务，本公司的关联机构或个人可能在本报告公开发布之间已经了解或使用其中的信息。

分析师承诺“本人及直系亲属与本报告中涉及的内容不存在利益关系”。