

因子拥挤度专题

因子择时的三个标尺： 因子动量、因子离散 度与因子拥挤度

科大财经 2023年10月8日

大纲：

1. 因子动量
 1. 横截面动量
 2. 时间序列动量
2. 因子离散度
3. 因子拥挤度
 1. MSCI因子拥挤度
 2. 国盛证券因子拥挤度

因子动量（截面）

因子预测能力的持续性，根据往期因子的IC表现，外推因子未来的表现。

因子截面配置策略：

- A) 因子等权：即七个大类因子等权配置；
- B) IC 加权：即按照过去 N 期的 IC 均值进行因子配权；
- C) ICIR 加权：即按照过去 N 期的 IC 均值/IC 标准差进行因子配权；
- D) ICWR 加权：即按照过去 N 期的 IC 大于 0 的概率进行因子配权；

从结果来看，横截面因子动量加权对多因子策略的影响并非是完全线性的，而是呈现出明显的 U 型结构。如果我们按照滚动窗口的长度来分析的话，可以得到一些洞察：

短期（6-12）横截面动量加权可以控制因子回撤，提升多因子策略的超额收益。滚动窗口并非越短越好，N=9 以及 N=12 为较优的参数；

中期（12-24）横截面动量加权既不能通过快速的动态因子调整来适应市场，也不能稳定地持有高风险溢价地因子来获取收益，因此无论在回撤控制以及超额收益提升上均显著跑输因子等权策略；

长期（24-36）横截面因子动量加权方案的特征是典型的高收益高风险策略，其本质上长期超配高风险溢价的因子（如规模因子），通过忍受策略的高波动和高回撤从而获取长期来看更高的收益。

图表 12: 横截面因子动量对超额收益的影响



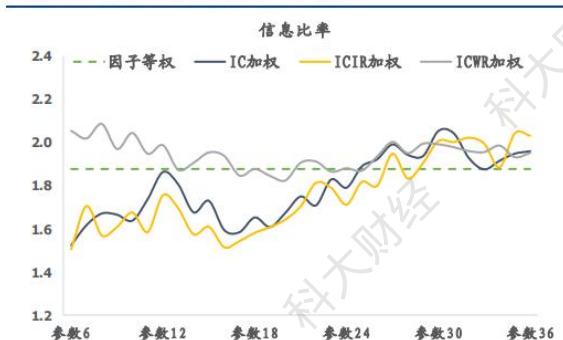
资料来源: Wind, 国盛证券研究所

图表 13: 横截面因子动量对最大回撤的影响



资料来源: Wind, 国盛证券研究所

图表 14: 横截面因子动量对信息比率的影响



资料来源: Wind, 国盛证券研究所

图表 15: 横截面因子动量对 Calmar 比率的影响



资料来源: Wind, 国盛证券研究所

因子动量（时序）

因子时序配置策略：

A) IC 加权：剔除过去 N 期 IC 均值 <0 的因子，剩余因子以过去 N 期的 IC 均值进行配权；

B) ICIR 加权：剔除过去 N 期 ICIR <0 的因子，剩余因子以过去 N 期的 ICIR 进行配权；

C) ICWR 加权：剔除过去 N 期 IC 胜率 <0.5 的因子，剩余因子以过去 N 期的 IC 胜率配权；

混合因子动量策略本质上就是用时序动量来筛选因子，用横截面动量进行加权，是时序

动量和横截面动量策略的有机结合，对比单纯的横截面动量，我们有两个结论：

短期（6-12）混合动量可以比较有效地提高因子配路策略的收益以及降低多因子策略的回撤。在因子选择上，混合动量效应比单一的横截面动量效应要好；

相比于 IC 加权和 ICWR 加权，ICIR 加权既能提高策略超额收益，又能很好地控制策略回撤，同时对参数 N 的敏感性较低。因此 ICIR 是较好的因子动量代理指标。

图表 16：混合因子动量对超额收益的影响



资料来源：Wind，国盛证券研究所

图表 17：混合因子动量对最大回撤的影响



资料来源：Wind，国盛证券研究所

图表 18：混合因子动量对信息比率的影响



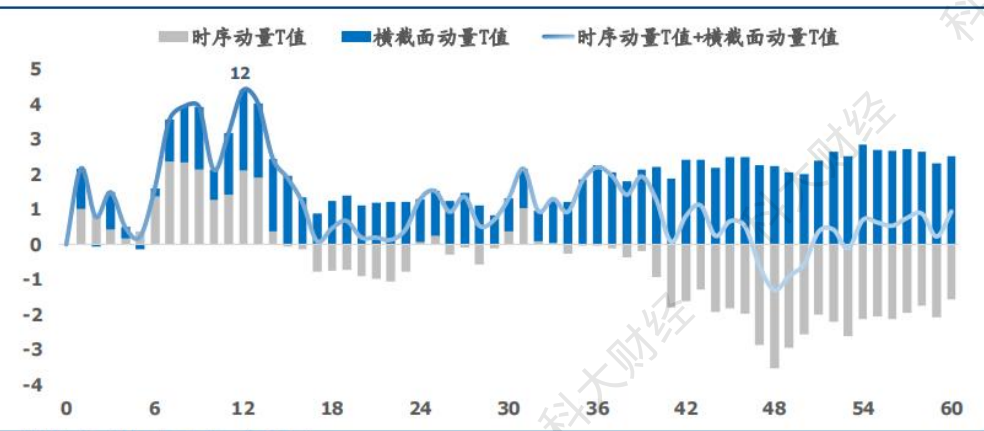
资料来源：Wind，国盛证券研究所

图表 19：混合因子动量对 Calmar 比率的影响



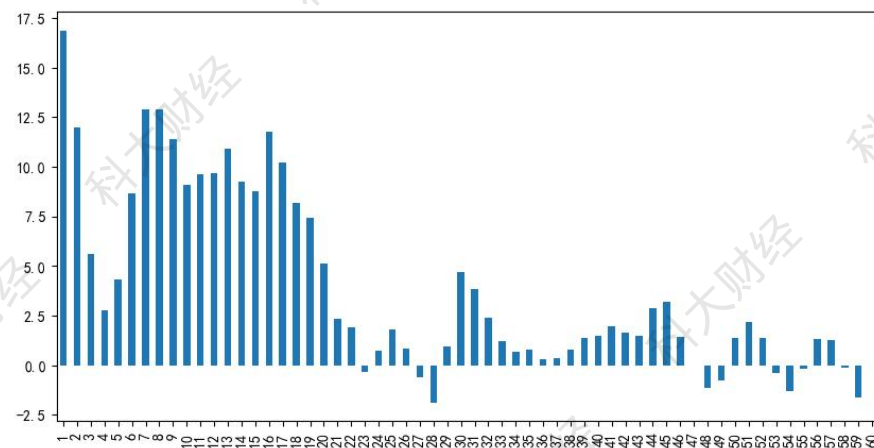
资料来源：Wind，国盛证券研究所

图表 22: 混合因子动量效应与参数设置

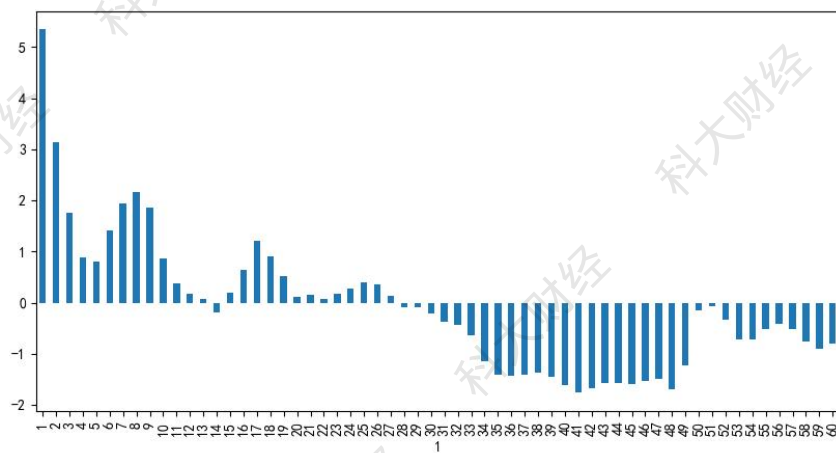


资料来源: Wind, 国盛证券研究所

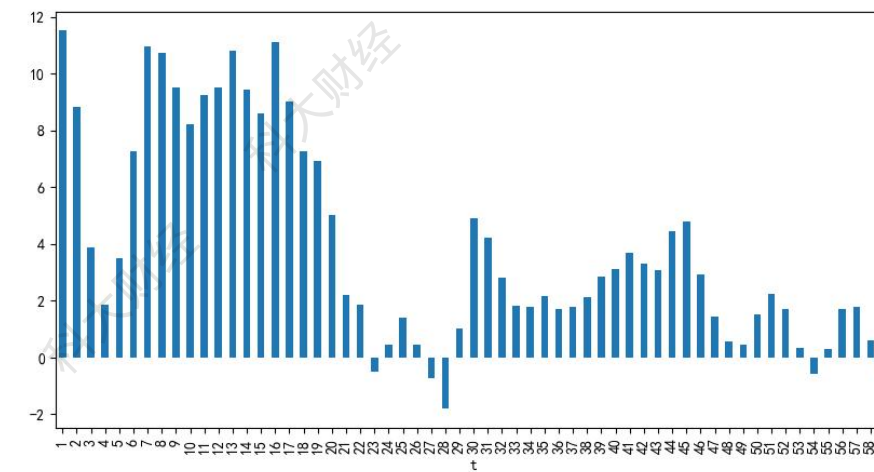
因子收益的
时序+截面
特征



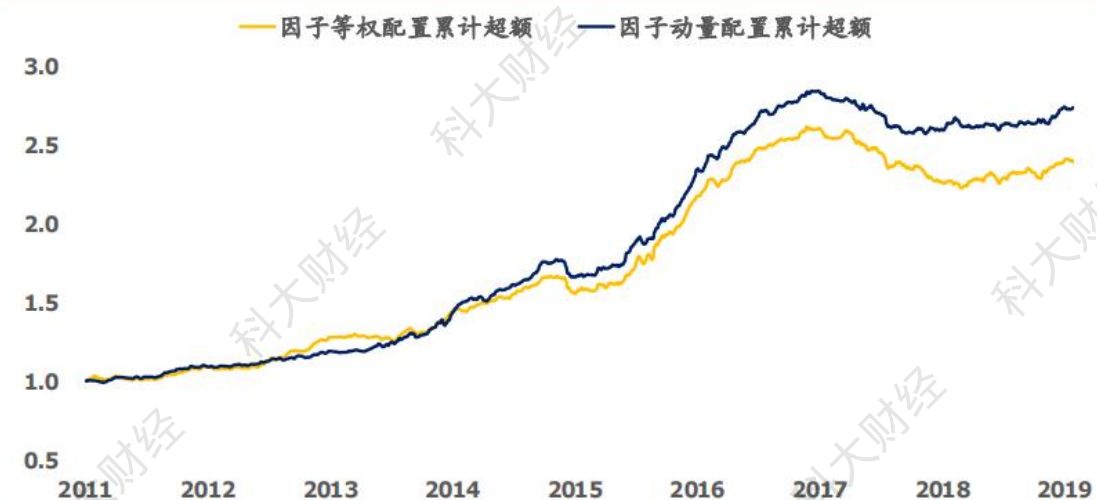
因子收益的
时序特征



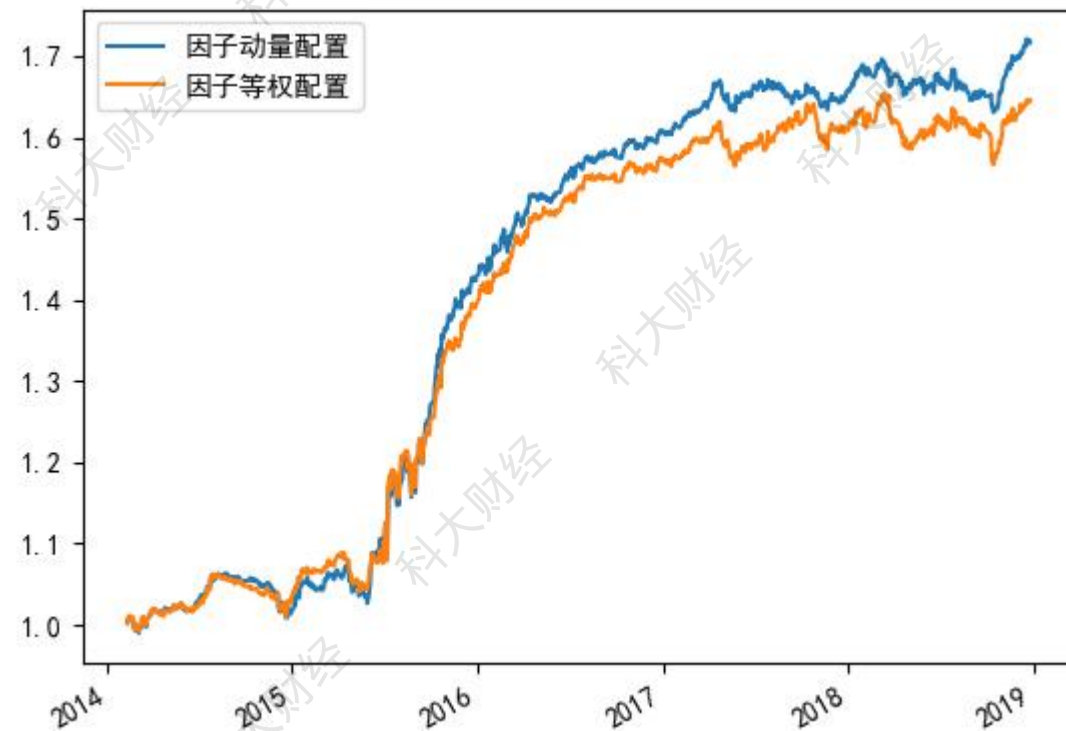
因子收益的
截面特征



图表 24: 基于因子动量的因子配置策略净值表现



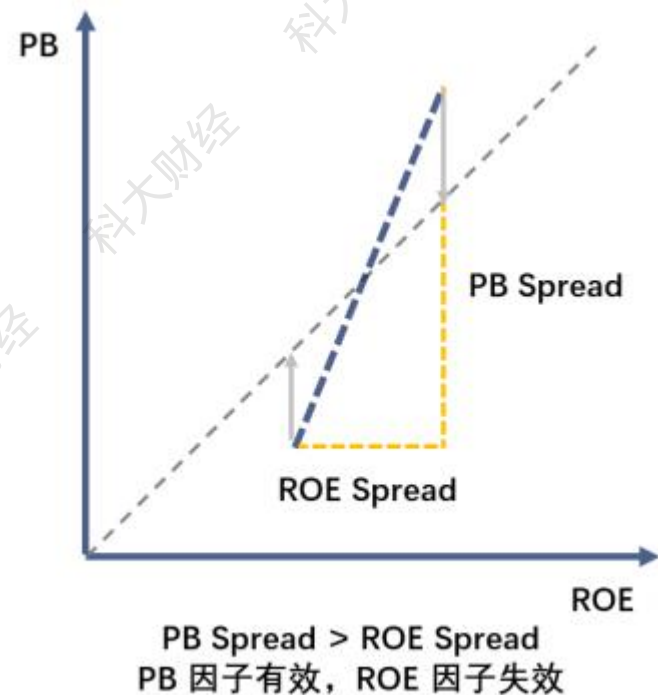
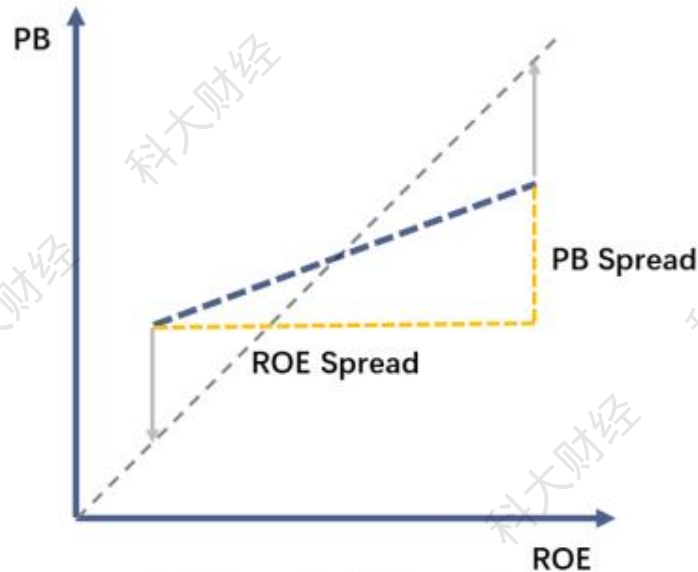
资料来源: Wind, 国盛证券研究所



因子离散度

首先我们该如何理解因子离散度为什么会影响因子的未来收益呢？可以假设一种极端场景，在某一个时间点上，所有股票的估值都相等，那么对于投资者来讲，估值这个指标就失去了选股的能力，所以说因子存在差异是因子未来能够区分股票收益的必要条件。其次，我们以 PB-ROE 模型为例来作进一步的解释，假设市场完全由 ROE 进行定价，那么 PB 和 ROE 之间将存在稳定的线性关系，如图表 26 中的灰色虚线。而如果市场并非完全有效的话，那么将出现两种情况：“欠定价市场”：高 ROE 股票的估值过低，低 ROE 股票的估值过高，ROE 的离散度明显高于 PB 的离散度，此时市场存在明显的套利空间。因此，未来一段时间大概率出现的现象是 ROE 因子有效而 PB 因子失效；

“过度定价市场”：低 ROE 股票的估值过低，高 ROE 股票的估值过高，PB 的离散度高于 ROE 的离散度，未来大概率出现的现象则是 PB 因子有效而 ROE 因子失效；因子离散度有效的核心逻辑在于资产定价在极度无效后会出现均值回复。进一步引申到多个因子的情况，如果某一个因子的离散度比其他因子的离散度都要大，那么这个因子“欠定价”的程度较深，未来这个因子更有可能产生超额收益。



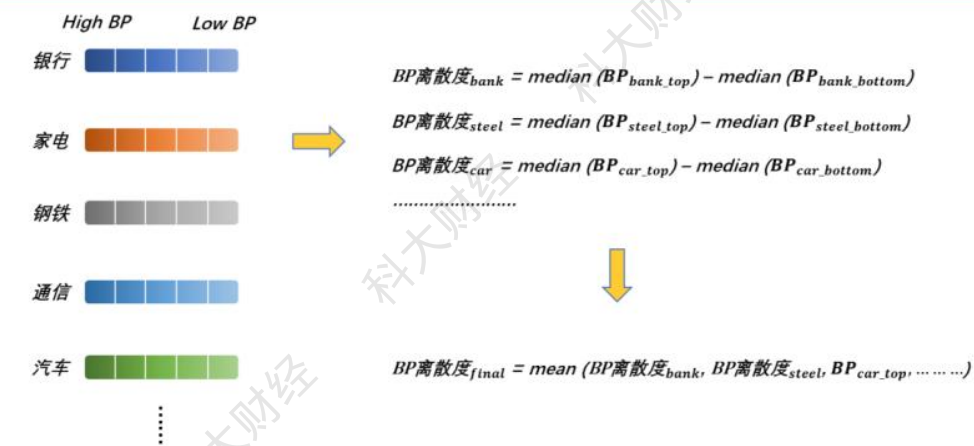
因子离散度计算流程

我们以 BP 因子离散度计算过程为例，简单说明一下因子离散度的计算方法：

- 1) 在每一个行业内，按照 BP 的大小将行业内股票分为五组，计算 Top 组和 Bottom 组的股票 BP 的中位数，记为 High BP 和 Low BP；
- 2) 将每个行业的 High BP 和 Low BP 的绝对值差异作为行业内的 BP 因子离散度；
- 3) 全市场的 BP 因子离散度为所有行业 BP 因子离散度的平均值；
- 4) 由于不同因子的量纲不完全可比，我们以过去 72 个月的数据对因子离散度进行标准化，标准化后的因子离散度具有横向可比性：

$$\text{标准化因子离散度}_t = \frac{(\text{因子离散度}_t - \text{avg}(\text{因子离散度}_{t-72,t}))}{\text{stdev}(\text{因子离散度}_{t-72,t})}$$

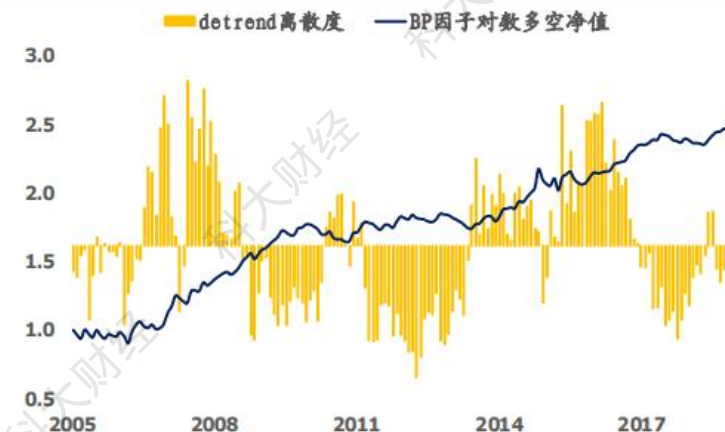
图表 27: 因子离散度计算流程: 以 BP 离散度计算为例



资料来源: 国盛证券研究所

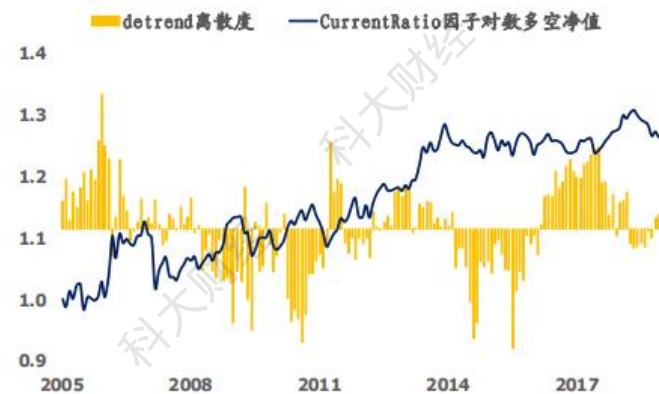


图表 28: BP 因子表现与因子离散度



资料来源: Wind, 国盛证券研究所

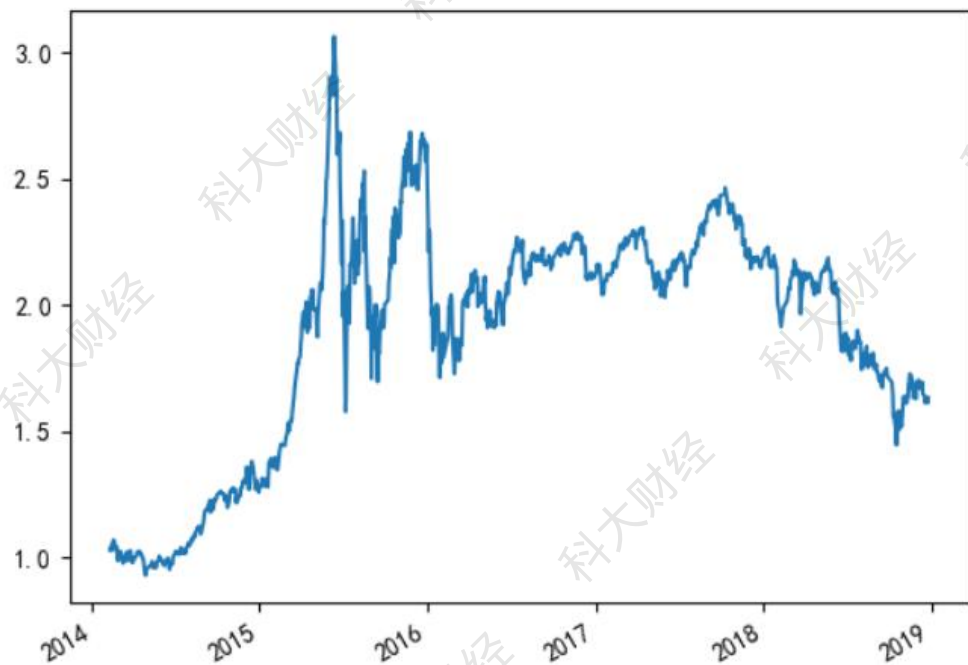
图表 36: CurrentRatio 因子表现与因子离散度



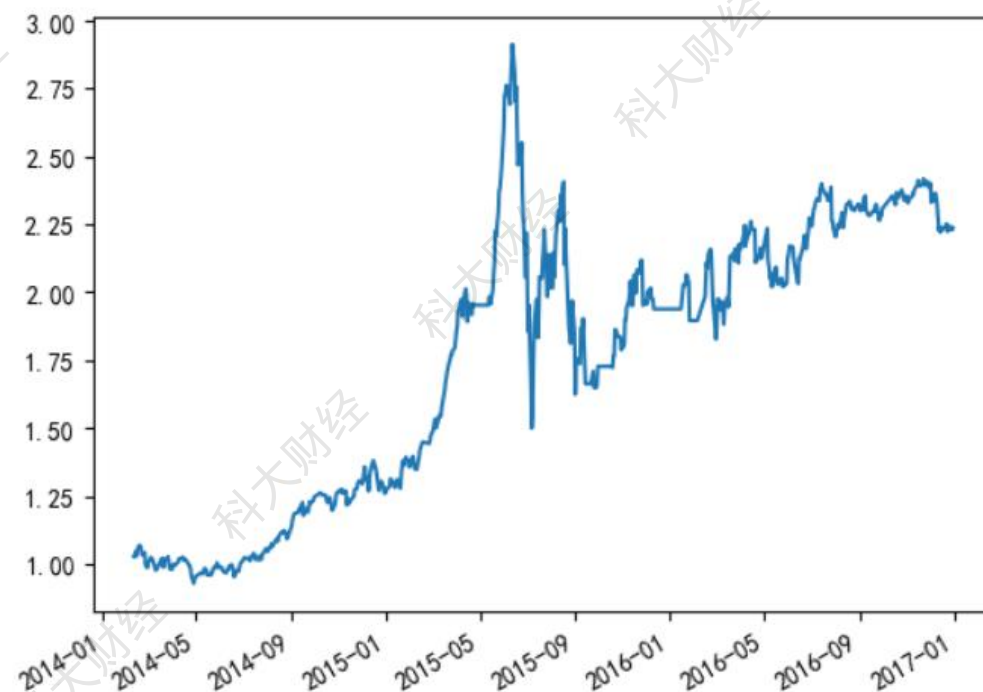
资料来源: Wind, 国盛证券研究所

离散度合成因子表现

	net_values	year_ret_sqrt	sharpe	downlow	volitiy
raw_net	1.6331	11.48	0.23	52.85	36.66



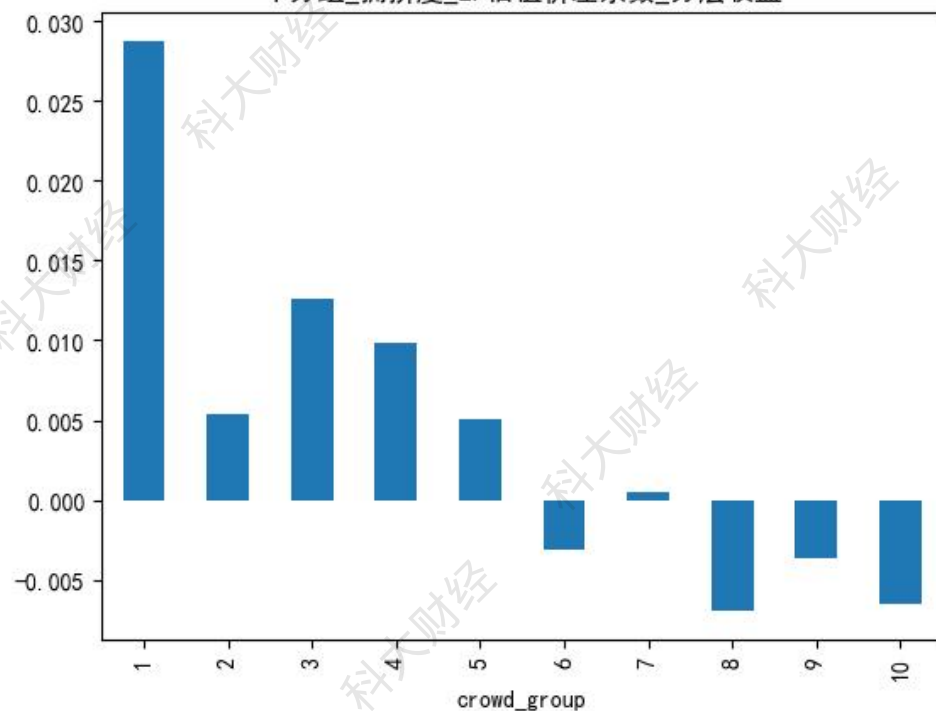
	net_values	year_ret_sqrt	sharpe	downlow	volitiy
fix_net	2.2367	35.0	0.8	48.53	39.85



因子拥挤度（十分组》五分组）

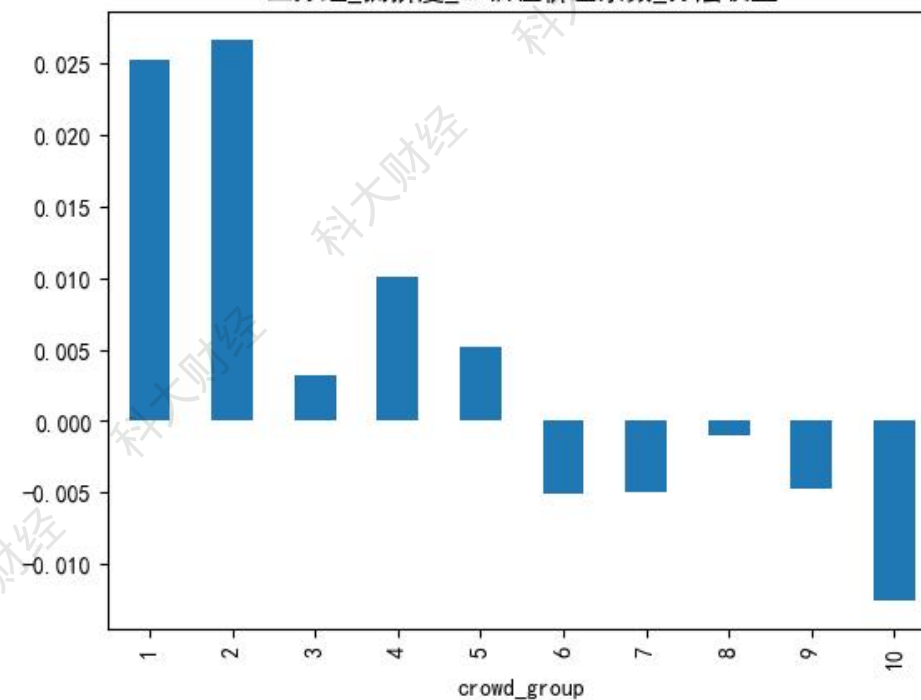
	crowd	crowd_z60	crowd_z120	crowd_z240	mon_cumret	mom_ret20	mom_ret60
crowd	1.000000	0.554262	0.665610	0.711329	-0.268516	-0.241213	-0.238559
crowd_z60	0.554262	1.000000	0.858981	0.750867	0.126445	-0.089819	-0.102933
crowd_z120	0.665610	0.858981	1.000000	0.923309	0.229199	-0.132840	-0.155653
crowd_z240	0.711329	0.750867	0.923309	1.000000	0.374200	-0.212825	-0.326482
mon_cumret	-0.268516	0.126445	0.229199	0.374200	1.000000	-0.220319	-0.443381
mom_ret20	-0.241213	-0.089819	-0.132840	-0.212825	-0.220319	1.000000	0.586361
mom_ret60	-0.238559	-0.102933	-0.155653	-0.326482	-0.443381	0.586361	1.000000

十分组_拥挤度_EP估值价差系数_分层收益



	crowd	crowd_z60	crowd_z120	crowd_z240	mon_cumret	mom_ret20	mom_ret60
crowd	1.000000	0.532455	0.654486	0.777719	-0.050366	-0.216683	-0.287491
crowd_z60	0.532455	1.000000	0.841712	0.719292	-0.007034	0.003379	0.008928
crowd_z120	0.654486	0.841712	1.000000	0.917237	0.109446	-0.022572	-0.089304
crowd_z240	0.777719	0.719292	0.917237	1.000000	0.208644	-0.097574	-0.233649
mon_cumret	-0.050366	-0.007034	0.109446	0.208644	1.000000	-0.220319	-0.443381
mom_ret20	-0.216683	0.003379	-0.022572	-0.097574	-0.220319	1.000000	0.586361
mom_ret60	-0.287491	0.008928	-0.089304	-0.233649	-0.443381	0.586361	1.000000

五分组_拥挤度_EP估值价差系数_分层收益



MSCI因子拥挤度（改进一）

forecast E/P among both the top and bottom quintile stocks of a factor, and then calculate the ratio or difference between the two. For B/P, we calculate the natural log of the ratio of

$$(\text{median B/P in bottom quintile of factor}) / (\text{median B/P in top quintile of factor})$$

In this formulation, large positive values are associated with relatively high degrees of crowding in the factor, as the denominator will be small relative to the numerator. We use the ratio of bottom quintile B/P to top quintile B/P, rather than the arithmetic difference, because the difference will be sensitive to overall market valuations, but the ratio will not. We use the same formulation with the S/P valuation ratio. However, because earnings yield can theoretically become negative or near zero, we use the difference for the forecast E/P valuation ratio. Specifically, we calculate

$$(\text{median E/P in bottom quintile of factor}) - (\text{median E/P in top quintile of factor})$$

We standardize each valuation spread using its own cumulative time series up to the given time, and then take the equal-weighted average of the three as the valuation spread metric

改进后

	crowd	crowd_z60	crowd_z120	crowd_z240	mon_cumret	mom_ret20	mom_ret60
crowd	1.000000	0.384245	0.439753	0.596257	0.478668	0.199716	0.174463
crowd_z60	0.384245	1.000000	0.853559	0.708344	-0.166105	0.173488	0.100208
crowd_z120	0.439753	0.853559	1.000000	0.887624	-0.260987	0.155638	0.106418
crowd_z240	0.596257	0.708344	0.887624	1.000000	-0.200170	0.172188	0.207499
mon_cumret	0.478668	-0.166105	-0.260987	-0.200170	1.000000	-0.220319	-0.443381
mom_ret20	0.199716	0.173488	0.155638	0.172188	-0.220319	1.000000	0.586361
mom_ret60	0.174463	0.100208	0.106418	0.207499	-0.443381	0.586361	1.000000

改进前

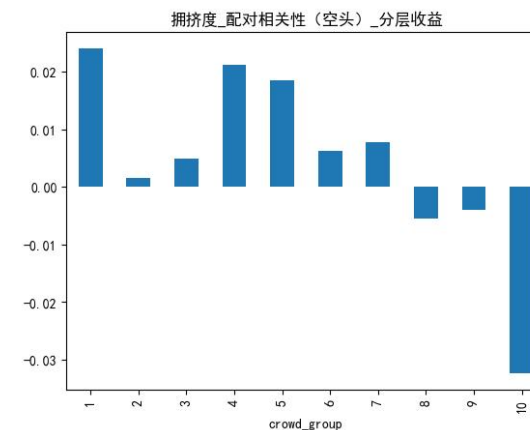
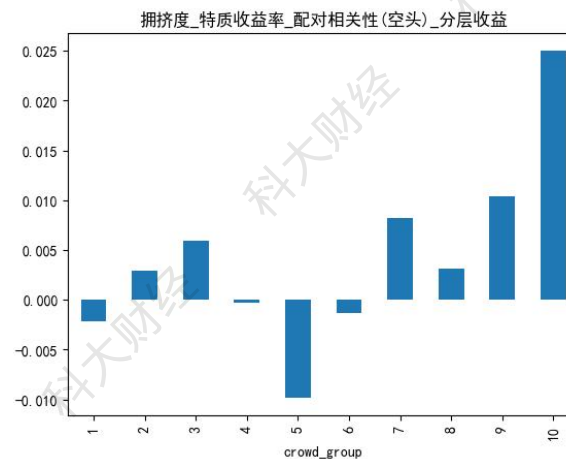
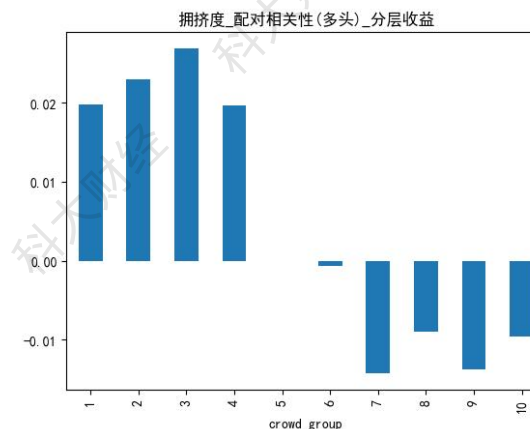
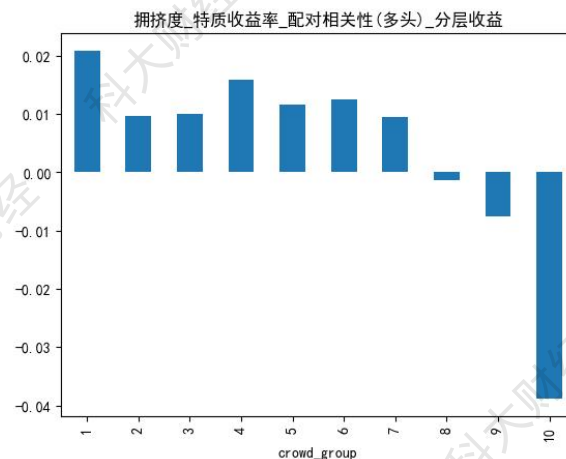
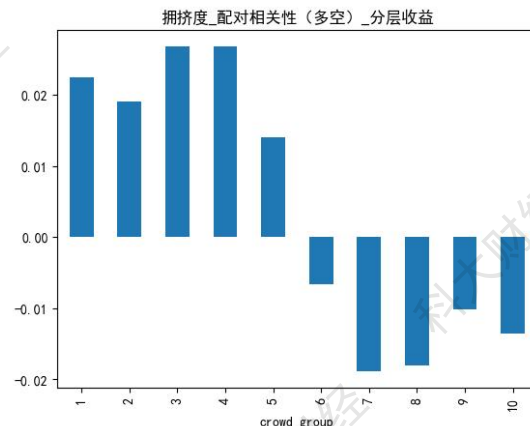
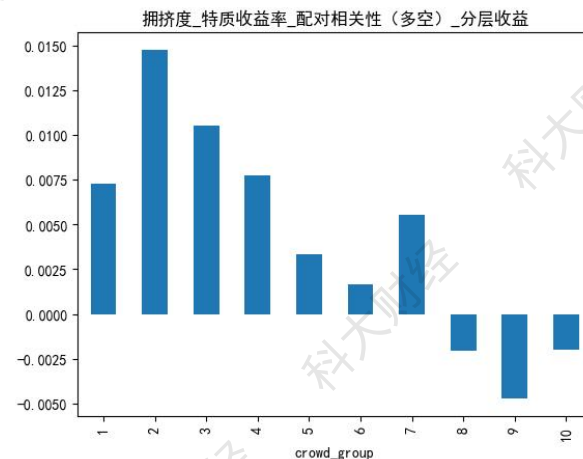
	crowd	crowd_z60	crowd_z120	crowd_z240	mon_cumret	mom_ret20	mom_ret60
crowd	1.000000	0.401085	0.526985	0.661688	-0.317405	-0.270371	-0.303478
crowd_z60	0.401085	1.000000	0.872685	0.713482	0.231977	-0.183735	-0.128303
crowd_z120	0.526985	0.872685	1.000000	0.902966	0.297554	-0.239067	-0.183389
crowd_z240	0.661688	0.713482	0.902966	1.000000	0.235079	-0.187604	-0.205465
mon_cumret	-0.317405	0.231977	0.297554	0.235079	1.000000	-0.220319	-0.443381
mom_ret20	-0.270371	-0.183735	-0.239067	-0.187604	-0.220319	1.000000	0.586361
mom_ret60	-0.303478	-0.128303	-0.183389	-0.205465	-0.443381	0.586361	1.000000

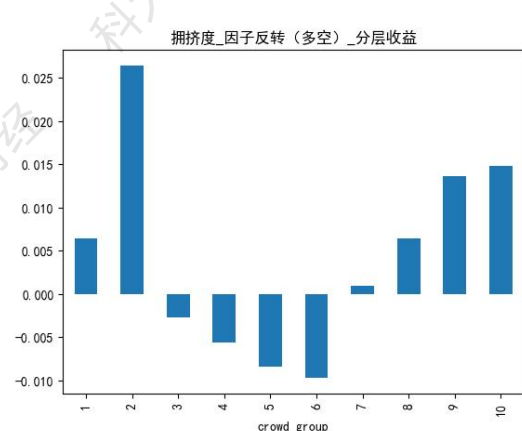
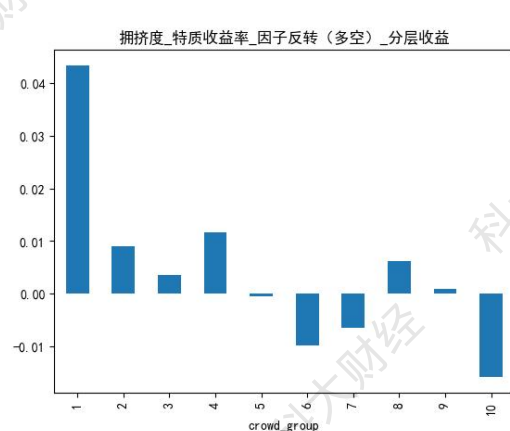
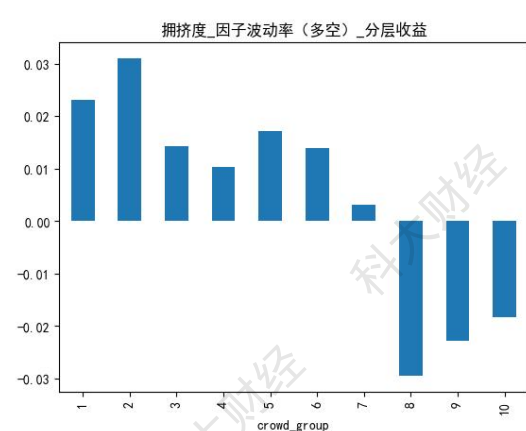
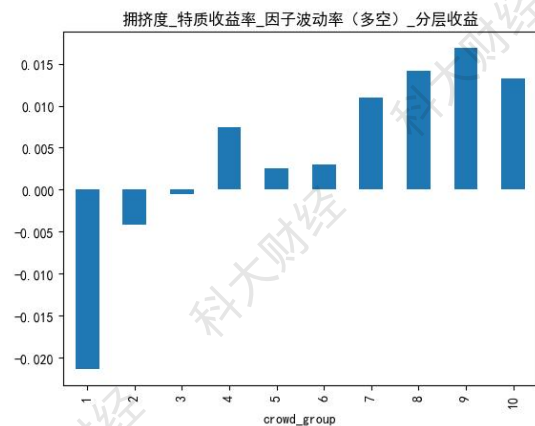
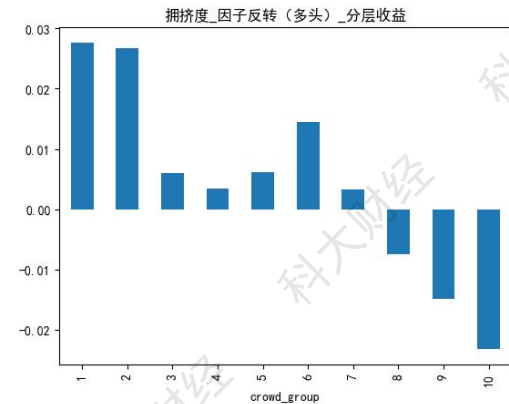
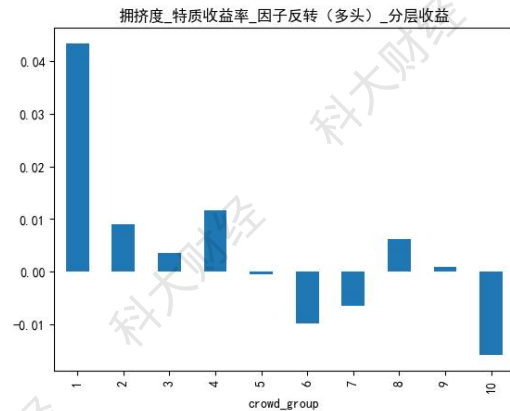
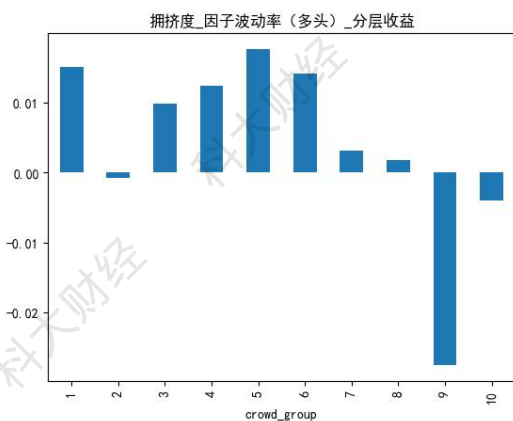
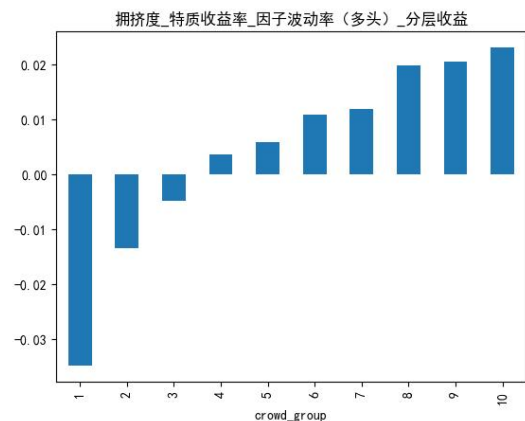
MSCI因子拥挤度（改进二）

PAIRWISE CORRELATION

This metric measures the degree to which stocks of the top or bottom quintile of a factor move together, after accounting for movements due to the market, size, beta and residual volatility factors. If a factor is being heavily followed by investors, stocks with very high exposures to the factor will, in theory, **tend to move together**, as will stocks with very low exposure to the factor. This will cause high average pairwise correlation. For this metric, we implemented the same formulation as described in Bayraktar et al. (2015a), which is based on the framework of Lou and Polk (2013).

In this metric, we select the top quintile securities of a factor and for each security we measure its correlation with the top quintile average return, excluding the individual stock, using the past 63 trading days of daily returns. We do the same for the bottom quintile. We use **specific returns** accounting for the standard risk factors of market, size, beta and residual volatility. We calculate the average pairwise correlation of the top and bottom quintiles separately, and then **take the average of the two quintiles to create the raw value**





国盛因子拥挤度

图表 42: 海外因子拥挤模型的思考

拥挤度定义	可能存在的问题
机构持仓强度差异 Fund's holdings spread	<ul style="list-style-type: none">❖ 机构持仓数据公布频率过低, 明显滞后;❖ 随着市场投资者结构的改变, 因子多空两端的机构持仓强度差异将出现中枢式的改变。若以此作为拥挤度定义, 实际操作中难以确定当前是否拥挤。
估值价差 Valuation spread	<ul style="list-style-type: none">❖ 高度依赖于价值因子的有效性, 在以价值风格为主导的市场中比较有效;❖ 存在明显的“名不副实”情况, 无论是近年来海外的实证结果还是笔者以 A 股数据的测算结果均显示估值价差的有效性并不强。
卖空强度差异 Short interest spread	<ul style="list-style-type: none">❖ 在融券做空比较自由、成本比较低的市场适用;❖ A 股的融资融券标的覆盖度有限。
配对相关性 Pairwise correlation	<ul style="list-style-type: none">❖ 国内一部分重要因子(反转、规模等)的未来收益与配对相关性为正;❖ 信号噪音太多, 实际操作中难以把握。
因子长期反转 Factor reversal	<ul style="list-style-type: none">❖ 美股因子存在 36 个月反转周期(图表 6), 而 A 股因子则存在 48 个月反转周期(图表 20), 由此来看反转效应在国别上并不稳定; 而且不同大类因子的反转周期也存在明显区别, 因此若以此作为拥挤度定义, 则存在很大的数据挖掘风险。

资料来源: 国盛证券研究所

图表 43: 国盛金工因子拥挤度模型

拥挤度指标	因子拥挤度定义
多空换手率比率	因子分组下, 多空两组的平均个股换手率比值; 换手率为个股过去 3 个月的日均成交换手率;
多空波动率比率	因子分组下, 多空两组的平均个股波动率比值; 波动率为个股过去 3 个月的收益率标准差;
多空 beta 比率	因子分组下, 多空两组的平均个股 beta 比值; beta 为过去 3 个月个股收益与万得全 A 指数收益回归的 beta;

资料来源: 国盛证券研究所

	净值	年化收益率	超额年化收益率	夏普比率	跟踪误差	信息比率	最大回撤	卡玛比率	日平均换手率
net	1.0638	0.85	5.75	-0.08	0.083916	-0.26	50.99	-0.04	0.007486

