

核心观点

本报告介绍了包含 213 个因子的高频因子库，从日内信息方面提供短周期量价维度因子定义逻辑，高频因子主要通过分钟、Level2 频率的基础数据进行资金流向、交易流动性、波动、买卖力量强弱等方面的统计和刻画，进而在日度频率上构建相应的指标和统计量作为因子值，从而捕捉交易信息中所包含的 Alpha。

主要结论

高频量价因子库

非线性建模在量化投研中的应用使得因子扮演的角色从 Alpha 信号转变为样本特征，从特征的角度出发，因子库的维护更加侧重于特征之间的信息差异性、以及特征库层面各项指标的统计分布情况。本报告介绍了基于分钟频率、Level2 数据构建的高频因子库，因子库包含 213 个日频量价因子，包含 12 大类，其中交易占比类因子反映了资金的交易意向和交易意图，上下行统计量类表示了价格涨跌过程中的强度、背后驱动力、以及交易行为等。

因子测试方法

本报告从 IC、ICIR、多空组合回测、多空净值等角度对整个因子库进行统计评估，在 1 日、5 日调仓频率上对因子进行分组、多空回测评估，在整个因子库、各个大类上综合分析因子的在各个维度上的表现。

因子库测试结果分析

所有因子 2018 年以来 1 日、5 日、10 日、20 日 rankIC 均值均在 2% 附近，六成因子的 1 日、5 日收益率 rankIC 在 1% 以上。1 日、5 日调仓频率下，全市场选股头部组合年化超额均值都为 4% 左右，年化波动率均匀分布在均值 20% 附近；多空年化收益均值分别为 7.42%、4.89%，前 10% 的因子多空日胜率在 55% 以上，八成因子多空、年化超为正。

风险提示：模型计算可能存在偏误，业绩不代表未来；因子测试、机器学习建模是对历史经验的总结，模型学习到的市场规律在未来存在失效的可能。

鲁植宸

luzhichen@csc.com.cn

15611537505

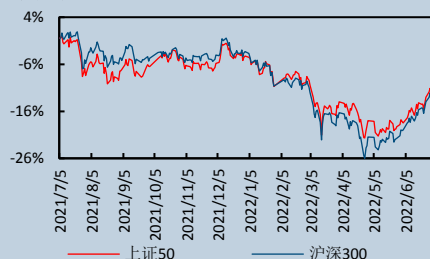
SAC 编号：S1440522080005

研究助理：徐建华

xujianhua@csc.com.cn

发布日期：2023 年 03 月 19 日

市场表现



相关研究报告

- 22.09.01 量价因子策略库（更新）
- 22.05.30 权益类公募基金量化标签库
- 22.05.12 CNE7 经典版多因子模型
- 22.03.18 FOF 投资量化策略
- 22.02.22 多因子拥挤度模型

目录

一、引言	1
二、高频因子	1
2.1 因子库概览	1
2.2 因子评估	3
三、因子库评估结果分析	3
3.1 各周期下 IC 情况	3
3.2 日频调仓测试	5
3.3 五日调仓测试	6
3.4 大类因子表现情况	8
四、总结	12

图表目录

图表 1：因子分类及说明	2
图表 2：IC 评估指标结果统计	4
图表 3：IC 评估指标分布情况	4
图表 4：回测结果统计（1 日调仓）	5
图表 5：回测指标分布情况（1 日调仓）	6
图表 6：回测结果统计（5 日调仓）	7
图表 7：回测指标分布情况（5 日调仓）	7
图表 8：各类别因子 IC 评估指标结果均值	8
图表 9：各类别因子回测指标均值（1 日调仓）	9
图表 10：各类别因子回测指标均值（5 日调仓）	9
图表 11：各类别因子多空净值（1 日调仓）	10
图表 12：各类别因子多空净值（5 日调仓）	11

一、引言

在 AlphaGo 问世后的上一波人工智能热潮以来，计算资源、数据挖掘手段日益丰富，从而使得对日内分钟数据、Level2 数据的因子挖掘在设备和技术门槛上大大降低。从算法、模型应用角度看，机器学习、人工智能算法的应用对于样本的采样频率、数量、以及应用场景的决策频率都有更高的要求。进而，周频、日频等中高频率调仓形式的指数增强、对冲策略在国内逐渐普及。^[1]

股票的基本面因子主要来源于公司的基本面状况，即公司财报以及其他报告，因此基于基本面因子的调仓频率与公司财务报表和其他报告的发布时间紧密相关。而量价因子主要来源于股票在市场上的交易，信息更迭更快，因此基于量价因子的调仓频率更高。相比于日频量价信息，日内分钟频率、Level2 层面的数据提供了更细粒度的交易层面的信息，为因子挖掘提供了更丰富的数据形式、以及更大的数据量。过往学术研究中，博弈论、微观市场结构等领域的研究成果在高频数据层面也有更多的指标化、因子化的途径和空间。

本篇报告中主要介绍包含 213 个高频因子的因子库，从因子定义逻辑上将因子库中因子分为 12 类，并基于因子 IC、日频调仓组合表现情况等角度对每个因子及整个因子库进行评估。

二、高频因子

2.1 因子库概览

本报告所介绍的因子库总共包含 213 个高频量价因子，因子库利用日内分钟频率、Level2 数据，从收益率相关的统计量、特定类型成交占比(主动买入、主动卖出等)、知情成交、流动性等维度进行因子构建，从而捕捉日内交易中所隐含的 Alpha。

基于高频因子所刻画量价信息的不同维度，本报告将所介绍因子库中的因子根据其计算逻辑分为 12 类，因子分类及说明见图表 1。

图表1： 因子分类及说明

因子类别	因子数量	类别说明
收益统计量	47	收益率本身刻画了股票价格趋势的动量反转效应，进一步在收益率的基础上，统计高频特定时间段收益、收益率峰度、偏度等统计量作为衍生因子
交易占比类	36	某些特定条件下的成交一定程度反映了资金的交易意向、交易意图，如早盘成交反映了基于盘前信息的决策带来的交易；尾盘成交反映了基于日内走势、交易情况所做出的决策带来的交易；根据成交额排序得到的大单、小单、中单刻画了买单的分散程度。统计不同条件下的成交在某段区间内的占比，反映了该时间区间内的成交量属性，这些属性背后存在着不同程度的 Alpha
知情成交类	21	将在买单价格上成交的订单定义为主动卖出，在卖单价格上成交的订单定义为主动买入。在日内高频维度上将收益率对其滞后项进行回归，结合残差项的正负和主买主卖情况，识别知情主动买入、知情主动卖出订单，并进一步在不同区间内、不同频率下统计知情成交占比作为知情成交类因子
幅度刻画类	20	波动率、涨跌情况都是刻画振幅的方式，在日内的不同区间、不同频率下对波动率和涨跌幅进行统计和衍生，作为振幅刻画类因子
量价相关性	20	量价相关性类因子的刻画角度主要是刻画成交量走势和价格走势的背离、一致程度。在此思路的基础上进一步进行衍生，在价格方面通过时间加权价格、成交量加权价格等变种，在量方面通过大单小单区分、或者对某些时间段进行区分，从而对某些子区间进行量价背离与同向程度
成交统计量类	13	成交量在日内的分布情况，总额、峰度、偏度等
资金流向类	13	股价处于上升状态时产生的成交额是推动股价上涨的力量，这部分成交额被定义为资金流入。股价处于下跌状态时产生的成交额是推动股价下跌的力量，这部分成交额被定义为资金流出。基于资金流入流出的定义，在日内不同频率、不同区间下统计资金流入流出的量、占比等指标，作为资金流向因子
价格类	12	日内价格根据成交量、形态等、某些特定条件信息进行调整之后的结果
经典理论衍生	9	基于动量反转、帕累托最优、黄金分割等经典理论，在高频数据维度构建相应思路的指标作为因子
上下行统计量	9	上行和下行过程中的某些统计量反映了价格涨跌过程中的强度、背后驱动力、以及交易行为。如上行波动率刻画了股票上涨过程中收益率的稳定性，衡量了上涨的买入力量的稳定性
波动率类	8	收益率标准差用于衡量收益波动性，在日内层面上，根据不同的时间窗口、不同的时间区间进行波动率的衍生统计，构建波动率类因子
流动相关类	5	将股票涨跌幅与成交金额联合考虑，作为刻画流动性的维度，在此思路上进行不同采样频率上的因子衍生和构建，作为流动相关类因子

数据来源：Wind，中信建投

注：数据截至 2023/02/28

2.2 因子评估

报告进一步通过分组、多空测试进行回测，对所有因子库中的因子进行评估。因子净值回测按照以下步骤进行：

1. 从全市场股票中剔除停牌、涨跌停、ST、次日一字板的股票。然后对股票进行 5 组分组，参考基准为中证全指。
2. 根据 T 日因子值、市值计算的权重作为 T+1 日收盘的目标权重，将 T 日的成交量加权平均价格记为 $VWAP_T$ ，则 T 日收盘后选出的股票在 T+1 日按照 VWAP 调仓，在 T+2 日获得的收益率为 $\frac{VWAP_{t+2}}{VWAP_{t+1}} - 1$ ，根据每组权重将收益加权求和，即得到每个等分组合的当日收益，进而可计算出分组净值。
3. 每日根据因子值排序选取头部组合和尾部组合分别构建多头组合和空头组合，记 T 日的头部组合收益为 r_1^T ，尾部组合收益为 r_5^T ，则当日的多空收益为 $r^T = r_1^T - r_5^T$ ，多空净值为 $\prod_t (1 + r^t)$ 。采用多空组合交易模型易于跟踪和资金的分配。可以每天都产生数量固定的定数目的股票，组合持有固定的期限到期后换仓。

本文所使用的多空净值评价指标有年化夏普比率、年化收益率、年化波动率、日收益率胜率以及最大回撤等。其中，夏普比率描述了投资组合收益对风险的补偿程度，年化收益率衡量了投资组合的收益情况，年化波动率刻画了投资组合的风险，收益率日胜率衡量了投资组合盈利的稳定性，最大回撤代表了收益率回撤幅度的最大值，衡量了投资组合控制风险的能力。

三、因子库评估结果分析

本节将根据上文所述方法对因子进行逐个评测，从各个评估指标的维度对整个因子库进行评估，在 1 日调仓、5 日调仓两种频率下进行回测，对于 5 日调仓场景，将资金分为 5 份，分别以 T~T+4 日为起始日回测构建 5 个子组合，将 5 个子组合合并后作为最终净值。

3.1 各周期下 IC 情况

本报告在 2018 年-2023 年 2 月历史区间上对每个因子计算日均 IC、年化 ICIR，对每个因子在测试区间上进行 IC 和 IR 分析、多空组合回测，并计算各个评估指标进行分析。

图表 2 和图表 3 展示了高频因子库的 IC 和 IR 分布情况。因子库的 1 日、5 日、10 日、20 日 IC 的均值都在 1.5% 左右，最大值 6% 左右，rankIC 的均值在 2% 以上，ICIR 的均值在 0.2 以上，最大超过 0.8。

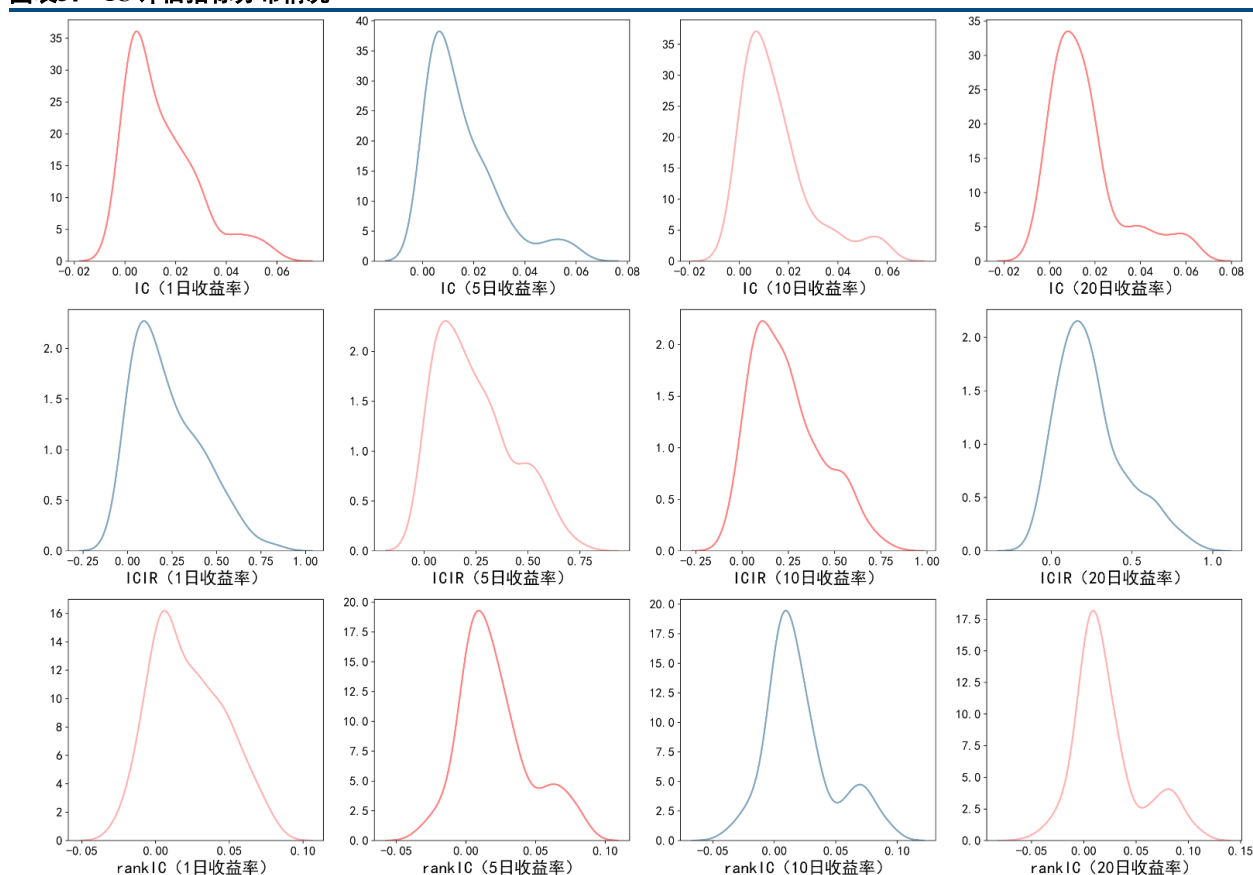
图表2： IC 评估指标结果统计

	1 日收益率			5 日收益率			10 日收益率			20 日收益率		
	IC	ICIR	rankIC	IC	ICIR	rankIC	IC	ICIR	rankIC	IC	ICIR	rankIC
均值	1.48%	0.22	2.21%	1.54%	0.24	2.09%	1.51%	0.24	2.09%	1.58%	0.25	2.25%
标准差	1.42%	0.19	2.43%	1.39%	0.18	2.49%	1.43%	0.19	2.76%	1.59%	0.21	3.14%
最小值	-0.32%	-0.07	-2.64%	0.01%	0.00	-3.18%	-0.43%	-0.08	-3.84%	-0.62%	-0.11	-4.96%
10%分位数	0.10%	0.02	-0.68%	0.18%	0.04	-0.24%	0.13%	0.04	-0.49%	0.06%	0.01	-0.44%
20%分位数	0.31%	0.05	0.08%	0.39%	0.07	0.18%	0.42%	0.07	0.22%	0.41%	0.08	0.17%
30%分位数	0.45%	0.09	0.52%	0.65%	0.11	0.53%	0.57%	0.11	0.56%	0.61%	0.12	0.52%
40%分位数	0.70%	0.13	1.05%	0.80%	0.15	1.07%	0.82%	0.16	0.87%	0.82%	0.16	0.81%
50%分位数	1.02%	0.18	1.82%	1.10%	0.20	1.49%	1.07%	0.21	1.46%	1.17%	0.21	1.40%
60%分位数	1.53%	0.23	2.68%	1.50%	0.26	2.09%	1.43%	0.26	1.86%	1.55%	0.26	1.69%
70%分位数	1.90%	0.31	3.57%	1.76%	0.33	2.81%	1.79%	0.31	2.72%	1.78%	0.30	2.89%
80%分位数	2.51%	0.39	4.51%	2.42%	0.39	3.81%	2.30%	0.40	3.71%	2.14%	0.42	3.94%
90%分位数	3.18%	0.50	5.79%	3.36%	0.52	6.29%	3.63%	0.54	7.10%	4.08%	0.59	8.01%
最大值	5.92%	0.84	8.10%	6.20%	0.75	8.35%	6.02%	0.79	9.34%	6.31%	0.89	10.78%

数据来源：Wind，中信建投

注：数据截至 2023/02/28

图表3： IC 评估指标分布情况



数据来源：Wind，中信建投

注：数据截至 2023/02/28

3.2 日频调仓测试

图表 4 和图表 5 展示了因子库日频调仓回测指标分布情况：

- 1) 多空夏普比率的密度中心约为 0.7，最大值 3.5，超额夏普比率均值接近 0.2。
- 2) 多空年化收益率以 7% 为中心分布，且分布较为集中，年化超额最高 17%。
- 3) 多空年化波动率均在 20% 以下，整体分布较为均匀，超额年化波动率呈现双峰值现象。
- 4) 近 80% 因子的多空日胜率超过 50%，最高超过 60%，超额日胜率最高达 52%。
- 5) 各因子多空最大回撤分布集中，以 24% 为中心，超额最大回撤均在 45% 以内。

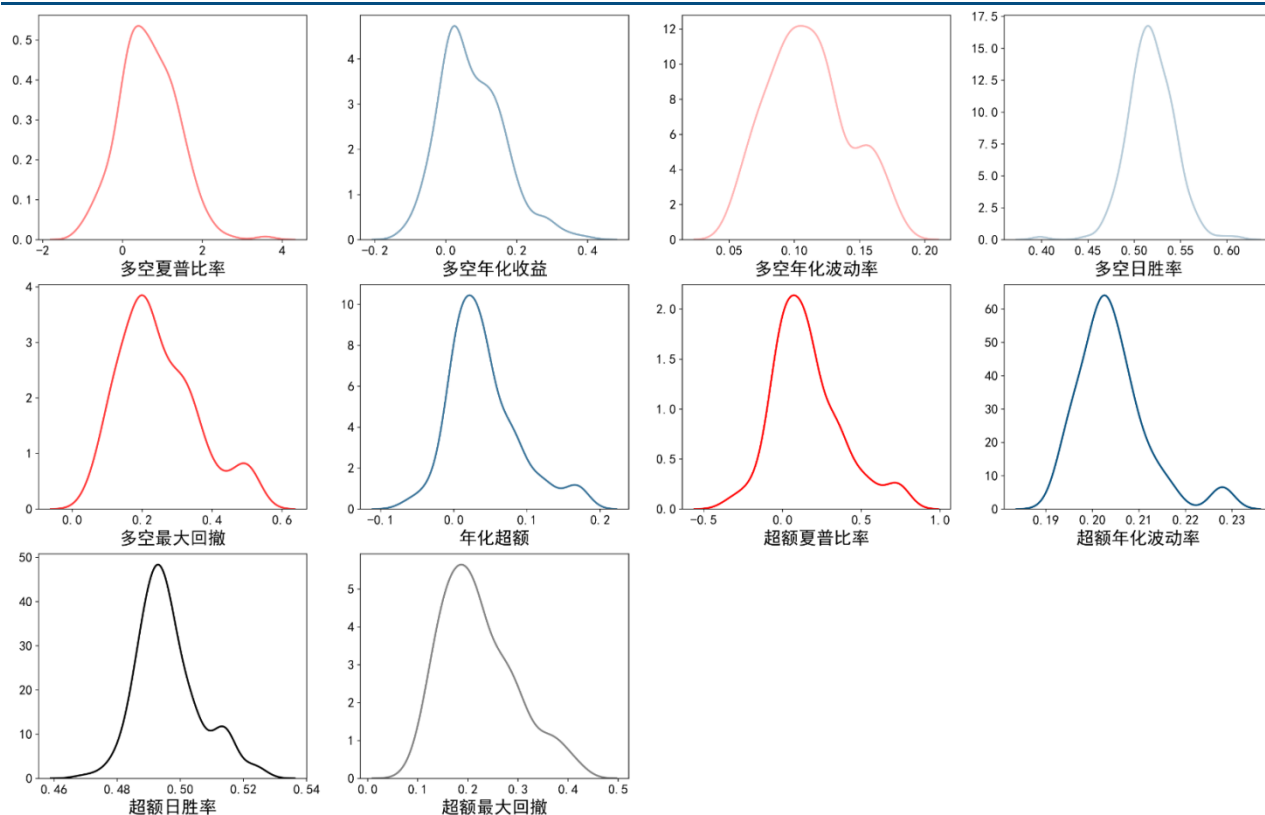
图表4： 回测结果统计（1日调仓）

统计量	多空夏普比率	多空年化收益	多空年化波动率	多空日胜率	多空最大回撤	年化超额	超额夏普比率	超额年化波动率	超额日胜率	超额最大回撤
均值	0.66	7.42%	11.07%	51.82%	24.81%	4.07%	0.16	20.42%	49.63%	22.09%
标准差	0.71	8.90%	3.05%	2.50%	11.10%	4.68%	0.22	0.75%	0.99%	7.33%
最小值	-1.07	-11.53%	5.47%	39.84%	5.28%	-6.29%	-0.33	19.13%	46.92%	8.50%
10%分位数	-0.16	-2.49%	7.28%	48.90%	11.81%	-0.24%	-0.04	19.57%	48.61%	13.85%
20%分位数	0.11	0.60%	8.28%	50.00%	14.96%	0.44%	-0.02	19.85%	48.84%	15.76%
30%分位数	0.25	1.76%	9.35%	50.48%	18.55%	1.24%	0.03	20.03%	49.08%	17.36%
40%分位数	0.41	3.32%	10.05%	51.26%	20.24%	2.31%	0.08	20.19%	49.24%	18.77%
50%分位数	0.61	5.68%	10.95%	51.76%	22.05%	3.14%	0.12	20.30%	49.48%	21.01%
60%分位数	0.79	9.14%	11.58%	52.26%	25.56%	4.10%	0.17	20.43%	49.64%	22.25%
70%分位数	1.06	11.91%	12.65%	53.04%	30.17%	5.03%	0.21	20.63%	49.88%	25.57%
80%分位数	1.24	14.37%	13.46%	53.76%	33.43%	7.71%	0.34	20.82%	50.36%	28.83%
90%分位数	1.62	17.75%	15.68%	54.72%	40.22%	10.89%	0.47	21.26%	51.31%	32.49%
最大值	3.58	39.03%	17.91%	61.04%	52.17%	17.42%	0.77	22.84%	52.60%	42.14%

数据来源：Wind，中信建投

注：数据截至 2023/02/28

图表5： 回测指标分布情况（1日调仓）



数据来源：Wind，中信建投

注：数据截至 2023/02/28

3.3 五日调仓测试

图表 6 和图表 7 展示了因子库 5 日调仓回测指标分布情况：

- 1) 多空夏普比率的密度中心约为 0.7，80%以上因子的夏普在 1 以上，最大值 2.6；超额夏普比率均值接近 0.2。
- 2) 多空年化收益率以 5%为中心分布，且分布较为集中；年化超额最高 17%。
- 3) 多空年化波动率以 7%左右为均值；超额年化波动率呈现双峰值现象。
- 4) 近 70%因子的多空日胜率超过 50%，最高超过 57%；超额日胜率最高达 52%。
- 5) 各因子多空最大回撤分布集中，以 19%为中心，最低达 3%；超额最大回撤均在 40%以内。

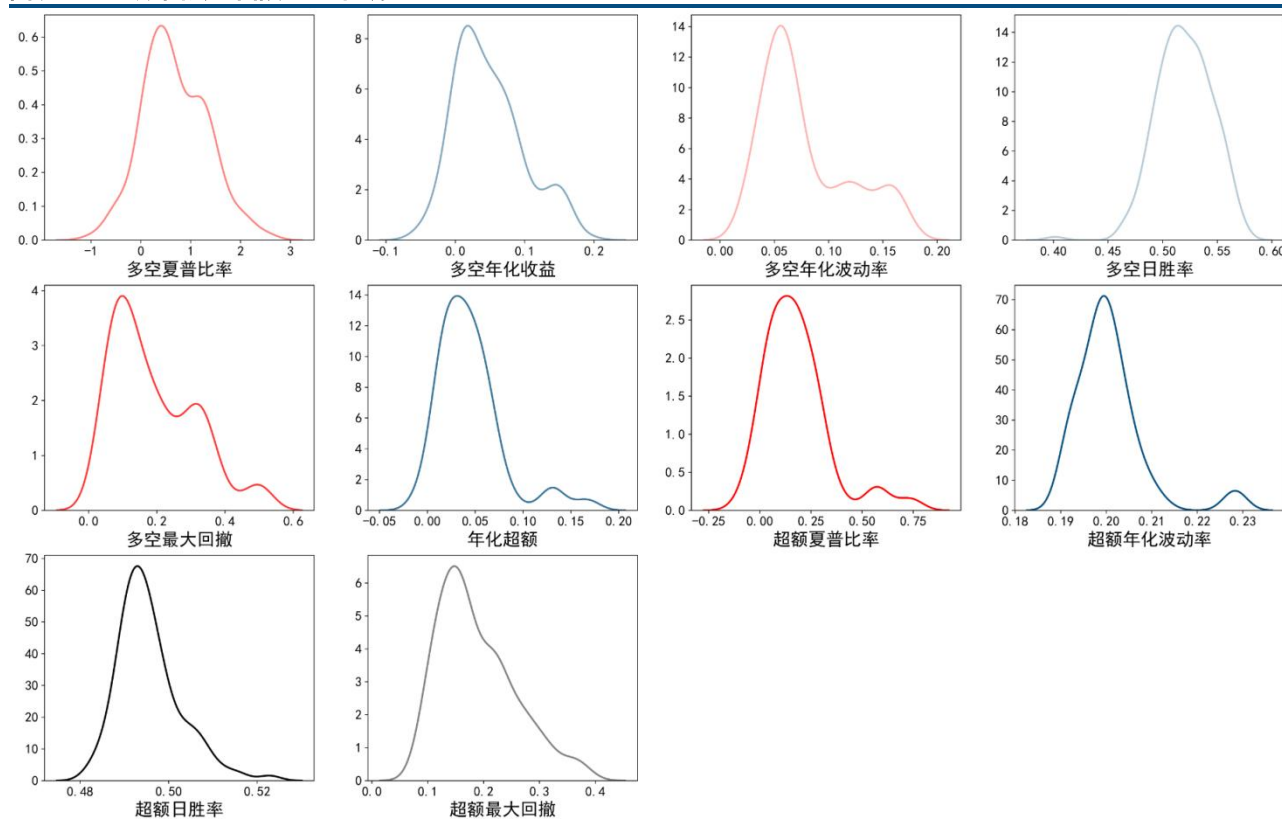
图表6：回测结果统计（5日调仓）

统计量	多空夏普比率	多空年化收益	多空年化波动率	多空日胜率	多空最大回撤	年化超额	超额夏普比率	超额年化波动率	超额日胜率	超额最大回撤
均值	0.68	4.89%	7.83%	52.00%	19.08%	4.44%	0.18	20.06%	49.56%	18.82%
标准差	0.65	5.00%	4.03%	2.51%	12.15%	3.43%	0.16	0.75%	0.71%	6.78%
最小值	-1.02	-5.70%	2.63%	40.16%	3.15%	-1.41%	-0.11	19.04%	48.20%	8.46%
10%分位数	0.00	-0.40%	3.68%	48.98%	6.18%	1.05%	0.00	19.23%	48.84%	11.24%
20%分位数	0.12	0.66%	4.75%	49.95%	7.73%	1.78%	0.05	19.53%	49.00%	12.98%
30%分位数	0.32	1.63%	5.39%	50.64%	10.52%	2.36%	0.08	19.73%	49.16%	14.33%
40%分位数	0.47	2.53%	5.79%	51.34%	12.25%	3.19%	0.12	19.84%	49.32%	15.76%
50%分位数	0.59	4.08%	6.17%	51.92%	15.41%	3.87%	0.16	19.97%	49.40%	16.93%
60%分位数	0.78	5.37%	6.92%	52.82%	19.31%	4.54%	0.19	20.05%	49.56%	19.40%
70%分位数	1.06	6.99%	9.01%	53.39%	24.62%	5.19%	0.22	20.20%	49.72%	21.86%
80%分位数	1.26	8.79%	11.81%	54.24%	30.63%	6.21%	0.27	20.43%	50.01%	24.21%
90%分位数	1.52	12.81%	15.71%	55.28%	34.34%	7.79%	0.34	20.69%	50.58%	28.54%
最大值	2.56	19.36%	16.92%	57.44%	50.03%	17.16%	0.76	22.87%	52.28%	38.40%

数据来源：Wind，中信建投

注：数据截至 2023/02/28

图表7：回测指标分布情况（5日调仓）



数据来源：Wind，中信建投

注：数据截至 2023/02/28

3.4 大类因子表现情况

图表 8 和图表 9 展示了各类别因子的 IC 评估指标和回测评估指标的均值。在 IC 评估指标上，从图表 8 可以看出：

- 1) IC 方面，大部分类别因子的 IC 均值都在 1% 以上，最大为波动率类，超过了 4%。
- 2) ICIR 方面，超半数类别因子的 ICIR 都在 0.2 以上，最大为波动率类，超过了 0.4。
- 3) rankIC 方面，半数类别因子的 rankIC 均值都在 2% 以上，最大为波动率类，超过了 6%。

图表8： 各类别因子 IC 评估指标结果均值

因子类别	因子数量	1 日收益率			5 日收益率			10 日收益率			20 日收益率		
		IC	ICIR	rankIC	IC	ICIR	rankIC	IC	ICIR	rankIC	IC	ICIR	rankIC
收益统计量	47	1.89%	0.30	2.94%	1.81%	0.30	2.59%	1.66%	0.29	2.51%	1.66%	0.29	2.63%
交易占比类	36	0.78%	0.12	1.31%	0.93%	0.16	1.40%	0.98%	0.17	1.48%	1.01%	0.17	1.55%
知情成交类	21	0.84%	0.16	1.59%	1.19%	0.22	2.15%	1.40%	0.25	2.53%	1.56%	0.27	2.87%
幅度刻画类	20	1.10%	0.15	0.77%	1.00%	0.14	0.15%	0.84%	0.12	-0.16%	0.80%	0.11	-0.29%
量价相关性	20	1.40%	0.25	2.14%	1.08%	0.21	1.37%	0.83%	0.18	1.04%	0.72%	0.15	0.89%
成交统计量类	13	2.21%	0.37	2.97%	2.23%	0.38	3.03%	2.23%	0.38	3.12%	2.30%	0.38	3.43%
资金流向类	13	0.57%	0.14	1.32%	0.54%	0.13	1.15%	0.50%	0.12	1.11%	0.50%	0.12	1.15%
价格类	12	0.71%	0.11	0.15%	0.95%	0.14	0.00%	1.14%	0.16	-0.07%	1.51%	0.18	0.04%
经典理论衍生	9	1.45%	0.20	1.65%	1.27%	0.20	1.07%	1.12%	0.19	0.87%	1.08%	0.20	0.79%
上下行统计量	9	3.53%	0.38	5.66%	3.82%	0.44	5.70%	3.80%	0.45	5.88%	4.08%	0.51	6.43%
波动率类	8	4.17%	0.45	6.09%	4.64%	0.56	6.56%	4.65%	0.58	7.01%	4.91%	0.65	7.73%
流动相关类	5	1.39%	0.22	4.35%	2.32%	0.34	5.72%	2.96%	0.40	6.56%	3.91%	0.49	7.97%

数据来源：Wind，中信建投

注：数据截至 2023/02/28

各大类因子在 1 日调仓频率下回测结果均值如图表 9 所示：

- 1) 半数类别因子的多空夏普比率都在 0.8 以上，最高为流动相关类，均值 1.45；超额夏普最高的类别为流动相关类，超过 0.6。
- 2) 半数类别因子的多空年化收益在 10% 以上，最高为流动相关类，均值近 20%；年化超额最高的类别为流动相关类，类内均值 15% 左右。
- 3) 各类别因子的多空年化波动率均在 20% 以下，最低为资金流向类，低于 10%；超额年化波动率均在 22% 以下。

图表9： 各类别因子回测指标均值（1日调仓）

因子类别	因子数量	多空夏普比率	多空年化收益	多空年化波动率	多空日胜率	多空最大回撤	年化超额	超额夏普比率	超额年化波动率	超额日胜率	超额最大回撤
收益统计量	47	0.58	7.04%	10.93%	51.47%	24.46%	2.96%	0.10	20.43%	49.56%	22.70%
交易占比类	36	0.81	8.68%	10.85%	51.82%	24.20%	4.38%	0.17	20.17%	49.56%	21.68%
知情成交类	21	0.77	7.35%	9.72%	51.11%	25.25%	5.55%	0.22	19.66%	49.38%	23.99%
幅度刻画类	20	0.44	5.05%	12.34%	51.32%	28.62%	2.83%	0.11	21.06%	49.76%	22.88%
量价相关性	20	0.62	5.46%	9.11%	51.99%	18.67%	2.46%	0.09	20.32%	49.20%	18.60%
成交统计量类	13	0.88	9.29%	9.76%	53.10%	24.57%	6.29%	0.26	20.42%	50.02%	19.71%
资金流向类	13	0.36	3.50%	8.11%	50.63%	17.58%	5.49%	0.23	19.85%	50.32%	21.20%
价格类	12	0.19	1.63%	13.58%	52.73%	41.39%	3.44%	0.13	21.79%	49.19%	27.83%
经典理论衍生	9	0.75	10.44%	11.25%	51.84%	20.56%	4.23%	0.17	20.69%	50.08%	18.72%
上下行统计量	9	0.86	12.43%	15.01%	52.80%	27.42%	3.05%	0.11	20.64%	49.47%	20.54%
波动率类	8	0.73	10.89%	15.90%	51.24%	26.92%	1.76%	0.03	20.23%	49.18%	25.95%
流动相关类	5	1.45	18.53%	12.28%	56.14%	19.15%	14.88%	0.64	20.84%	51.37%	19.40%

数据来源：Wind，中信建投

注：数据截至 2023/02/28

各大类因子在 5 日调仓频率下回测结果均值如图表 9 所示：

- 1) 半数类别因子的多空夏普比率都在 0.7 以上，最高为流动相关类，均值 1.2，超额夏普最高的类别为流动相关类。
- 2) 半数类别因子的多空年化收益在 5% 以上，最高为流动相关类，近 15%，最低为价格类，仅为 2.3%；年化超额最高的类别为流动相关类，近 13%。
- 3) 各类别因子的多空年化波动率均在 15% 以下，最低为量价相关性，低于 5%；超额年化波动率均在 22% 以下。

图表10： 各类别因子回测指标均值（5日调仓）

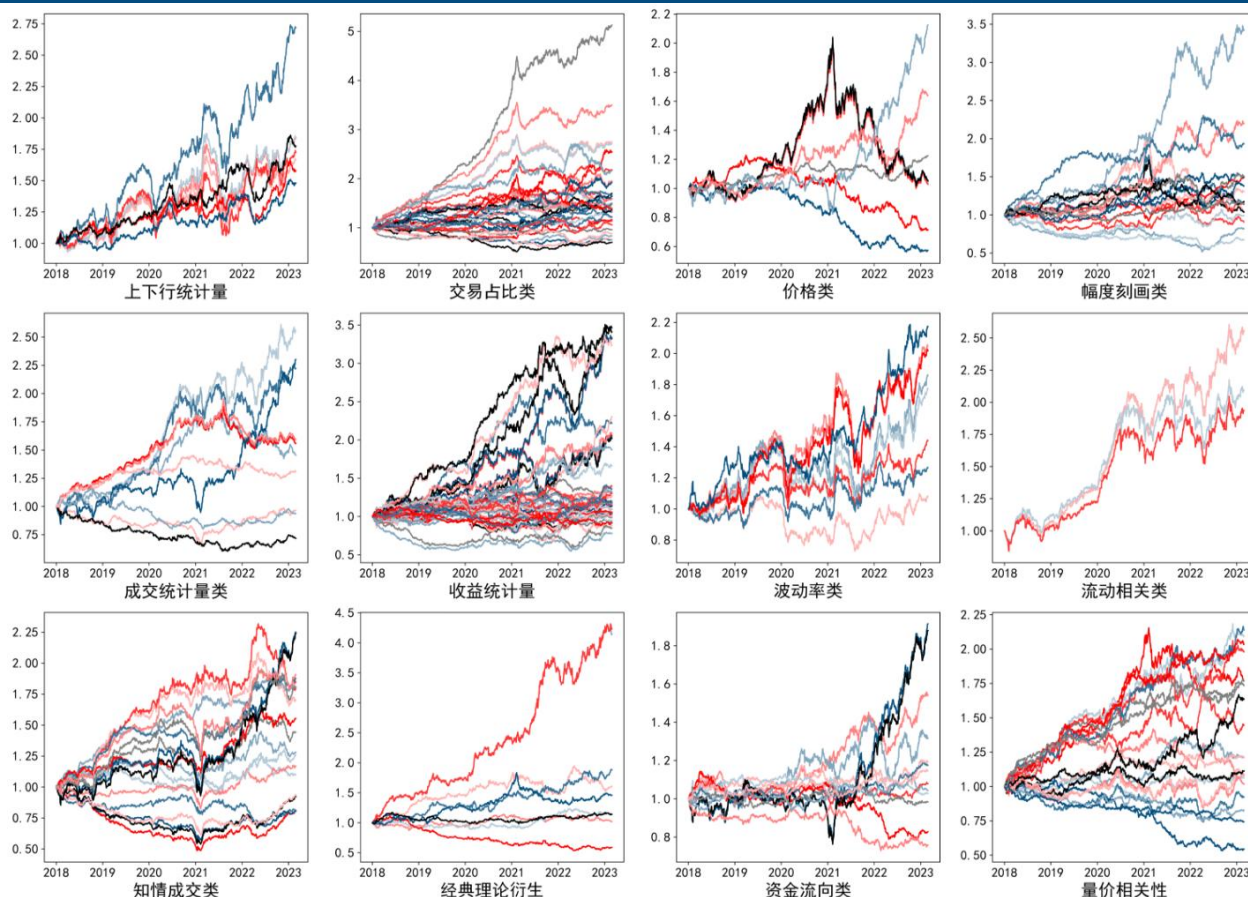
因子类别	因子数量	多空夏普比率	多空年化收益	多空年化波动率	多空日胜率	多空最大回撤	年化超额	超额夏普比率	超额年化波动率	超额日胜率	超额最大回撤
收益统计量	47	0.68	4.39%	7.42%	51.55%	18.14%	3.40%	0.13	20.01%	49.37%	18.82%
交易占比类	36	0.76	5.29%	7.33%	51.82%	18.20%	4.49%	0.18	19.80%	49.62%	17.69%
知情成交类	21	0.78	5.21%	6.80%	51.87%	19.05%	5.85%	0.24	19.31%	50.00%	22.32%
幅度刻画类	20	0.46	2.91%	7.68%	51.92%	20.84%	3.26%	0.13	20.46%	49.55%	18.35%
量价相关性	20	0.76	3.64%	4.74%	52.12%	9.68%	3.19%	0.13	19.95%	49.07%	14.82%
成交统计量类	13	0.95	8.24%	8.19%	53.35%	19.84%	7.34%	0.31	20.22%	49.95%	16.18%
资金流向类	13	0.46	2.99%	5.19%	50.73%	12.16%	6.32%	0.27	19.55%	50.00%	19.62%
价格类	12	0.36	2.29%	11.93%	52.87%	35.57%	4.99%	0.20	21.63%	49.41%	22.77%
经典理论衍生	9	0.60	4.51%	6.15%	51.95%	14.66%	3.06%	0.12	20.09%	49.22%	15.62%
上下行统计量	9	0.87	8.04%	12.52%	52.53%	24.80%	3.39%	0.12	20.39%	49.33%	18.47%
波动率类	8	0.43	5.58%	14.91%	50.94%	31.00%	1.59%	0.03	20.12%	49.21%	27.58%
流动相关类	5	1.20	14.77%	12.07%	56.29%	19.74%	12.66%	0.55	20.69%	50.81%	17.83%

数据来源：Wind，中信建投

注：数据截至 2023/02/28

进一步，图表 11 和图表 12 展示了 12 类因子的多空净值，从类内平均水平来看，流动相关类的表现最佳，净值最终超过了 2.3，其次，上下行统计量类因子普遍多空净值更为稳定；知情交易、交易占比、收益统计量类因子多空净值分化较大。日频调仓、5 日调仓下各类因子表现区别不大。

图表11： 各类别因子多空净值（1 日调仓）

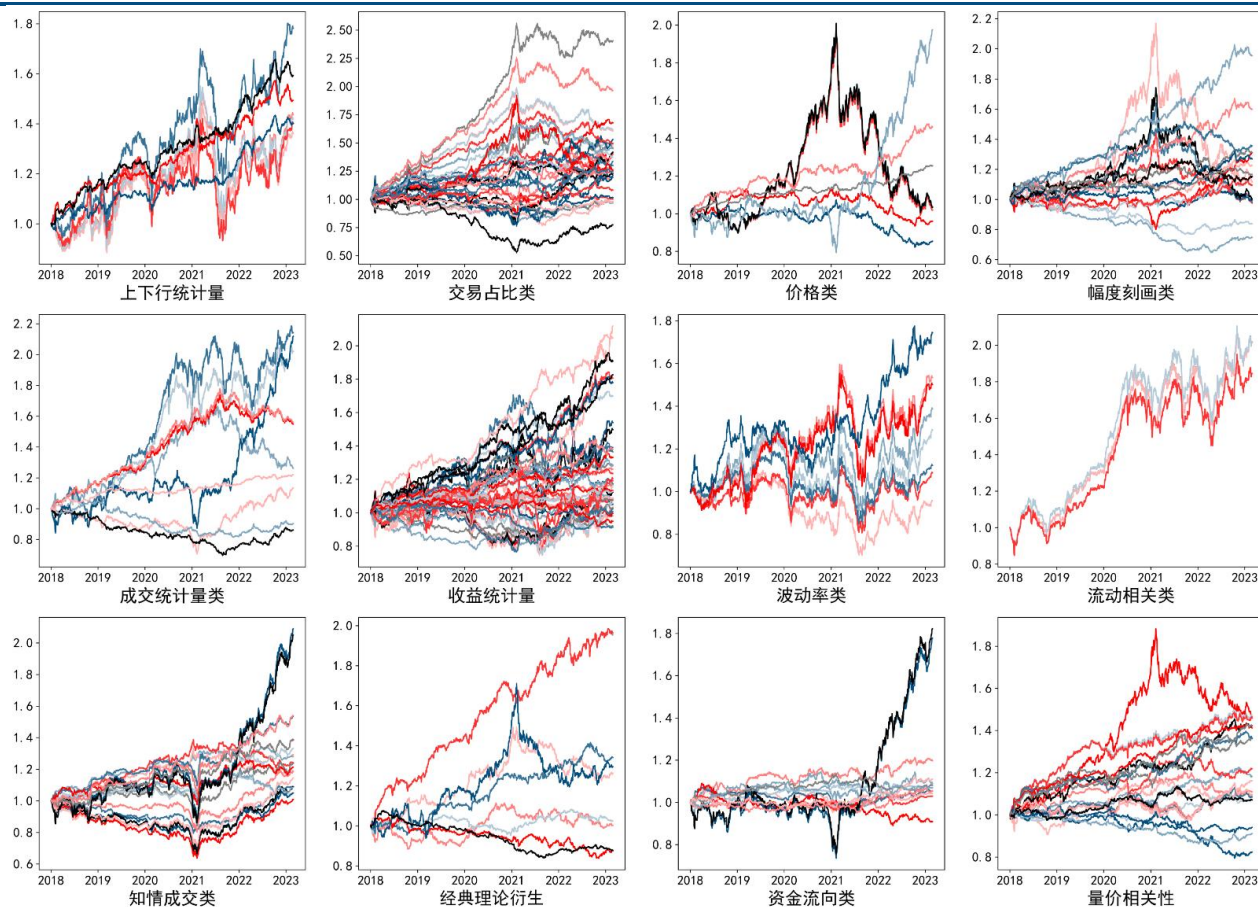


数据来源：Wind，中信建投

注：数据截至 2023/02/28

各类因子内部在多空净值方面存在不同程度的分化情况，其中流动相关性、上下行统计量两类因子整体分化情况较小；部分类别因子在 IC 指标、回测指标及净值等各方面均展现出相对更加良好的选股能力，如流动相关类、波动率类。

图表12： 各类别因子多空净值（5日调仓）



数据来源: Wind, 中信建投

注: 数据截至 2023/02/28

四、总结

随着机器学习这类非线性模型广泛应用于量化投研，从建模的角度来看，因子扮演的角色从 Alpha 信号转变为样本特征，从特征的角度出发，更加侧重于特征之间的信息差异性、以及特征库层面各项指标的统计分布情况。本文介绍了基于分钟频率、Level2 数据构建的高频因子库。相比于基于日频率量价数据构建的量价因子而言，高频数据包含更加细粒度的交易行为信息，基于高频基础数据进行因子挖掘，可以捕捉到日内交易行为背后所包含的 Alpha，作为基本面、日频量价信息之外的一个补充。

本报告所介绍的因子库共包含 213 个因子，所有因子 2018 年以来 1 日、5 日、10 日、20 日 rankIC 均值均在 2% 附近，六成因子的 1 日、5 日收益率 rankIC 在 1% 以上。1 日、5 日调仓频率下，全市场选股头部组合年化超额均值都为 4% 左右，年化波动率均匀分布在均值 20% 附近；多空年化收益均值分别为 7.42%、4.89%，前 10% 的因子多空日胜率在 55% 以上，八成因子多空、年化超为正。

分析师介绍

鲁植宸，多因子与 ESG 策略组分析师，中国科学院硕士，曾在头部券商自营从事量化投研和实盘管理工作，在人工智能、机器学习算法的量化应用方面有从学术研究到策略设计、实盘落地及策略管理的全周期经验，研究方向覆盖机器学习算法、多因子模型-策略设计、建模训练过程优化、样本特征处理技术、ESG 策略和指数编制等。

研究助理

徐建华，总监，多因子与 ESG 策略组负责人，先后就职于 MSCI Barra 公司和标普公司等专业研究机构，具有十多年金融市场研究、统计建模等专业经验。为多家银行、基金和保险等资管客户提供因子模型、投资组合优化管理、基金评价、指数研究以及风险管理等咨询。

评级说明

投资评级标准		评级	说明
报告中投资建议涉及的评级标准为报告发布日后 6 个月内的相对市场表现，也即报告发布日后的 6 个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。A 股市场以沪深 300 指数作为基准；新三板市场以三板成指为基准；香港市场以恒生指数作为基准；美国市场以标普 500 指数为基准。	股票评级	买入	相对涨幅 15% 以上
		增持	相对涨幅 5%—15%
		中性	相对涨幅 -5%—5% 之间
		减持	相对跌幅 5%—15%
		卖出	相对跌幅 15% 以上
	行业评级	强于大市	相对涨幅 10% 以上
		中性	相对涨幅 -10—10% 之间
		弱于大市	相对跌幅 10% 以上

分析师声明

本报告署名分析师在此声明：(i) 以勤勉的职业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，结论不受任何第三方的授意或影响。(ii) 本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

法律主体说明

本报告由中信建投证券股份有限公司及/或其附属机构（以下合称“中信建投”）制作，由中信建投证券股份有限公司在中华人民共和国（仅为本报告目的，不包括香港、澳门、台湾）提供。中信建投证券股份有限公司具有中国证监会许可的投资咨询业务资格，本报告署名分析师所持中国证券业协会授予的证券投资咨询执业资格证书编号已披露在报告首页。

在遵守适用的法律法规情况下，本报告亦可能由中信建投（国际）证券有限公司在香港提供。本报告作者所持香港证监会牌照的中央编号已披露在报告首页。

一般性声明

本报告由中信建投制作。发送本报告不构成任何合同或承诺的基础，不因接收者收到本报告而视其为中信建投客户。

本报告的信息均来源于中信建投认为可靠的公开资料，但中信建投对这些信息的准确性及完整性不作任何保证。本报告所载观点、评估和预测仅反映本报告出具日该分析师的判断，该等观点、评估和预测可能在不发出通知的情况下有所变更，亦有可能因使用不同假设和标准或者采用不同分析方法而与中信建投其他部门、人员口头或书面表达的意见不同或相反。本报告所引证券或其他金融工具的过往业绩不代表其未来表现。报告中所含任何具有预测性质的内容皆基于相应的假设条件，而任何假设条件都可能随时发生变化并影响实际投资收益。中信建投不承诺、不保证本报告所含具有预测性质的内容必然得以实现。

本报告内容的全部或部分均不构成投资建议。本报告所包含的观点、建议并未考虑报告接收人在财务状况、投资目的、风险偏好等方的具体情况，报告接收者应当独立评估本报告所含信息，基于自身投资目标、需求、市场机会、风险及其他因素自主做出决策并自行承担投资风险。中信建投建议所有投资者应就任何潜在投资向其税务、会计或法律顾问咨询。不论报告接收者是否根据本报告做出投资决策，中信建投都不对该等投资决策提供任何形式的担保，亦不以任何形式分享投资收益或者分担投资损失。中信建投不对使用本报告所产生的任何直接或间接损失承担责任。

在法律法规及监管规定允许的范围内，中信建投可能持有并交易本报告中所述公司的股份或其他财产权益，也可能在过去 12 个月、目前或者将来为本报告中所提公司提供或者争取为其提供投资银行、做市交易、财务顾问或其他金融服务。本报告内容真实、准确、完整地反映了署名分析师的观点，分析师的薪酬无论过去、现在或未来都不会直接或间接与其所撰写报告中的具体观点相联系，分析师亦不会因撰写本报告而获取不当利益。

本报告为中信建投所有。未经中信建投事先书面许可，任何机构和/或个人不得以任何形式转发、翻版、复制、发布或引用本报告全部或部分内容，亦不得从未经中信建投书面授权的任何机构、个人或其运营的媒体平台接收、翻版、复制或引用本报告全部或部分内容。版权所有，违者必究。

中信建投证券研究发展部

北京
东城区朝内大街 2 号凯恒中心 B 座 12 层
电话：(8610) 8513-0588
联系人：李祉瑶
邮箱：lizhiyao@csc.com.cn

上海
上海浦东新区浦东南路 528 号南塔 2106 室
电话：(8621) 6882-1600
联系人：翁起帆
邮箱：wengqifan@csc.com.cn

深圳
福田区益田路 6003 号荣超商务中心 B 座 22 层
电话：(86755) 8252-1369
联系人：曹莹
邮箱：caoying@csc.com.cn

中信建投（国际）

香港
中环交易广场 2 期 18 楼
电话：(852) 3465-5600
联系人：刘泓麟
邮箱：charleneliu@csci.hk