

使用多因子框架的沪深 300 指数增强模型

——多因子模型研究系列之七

分析师：宋旻

SAC NO: S1150517100002

2019 年 03 月 29 日

证券分析师

宋旻

022-28451131

18222076300

songyang@bhzq.com

助理分析师

张世良

022-23839061

zhangsl@bhzq.com

相关研究报告

《多因子模型研究之一：单因子测试》20171011

《多因子模型研究之二：收益预测模型》20171229

《多因子模型研究之三：风险模型与组合优化》20180416

《随机森林多因子模型与传统多因子模型的选股风格对比——多因子模型研究系列之四》20180726

《使用 bandit learning 算法的多因子模型——多因子模型研究系列之五》20180925

《使用 Thompson Sampling 算法的策略混合模型——多因子模型研究系列之六》20181228

核心观点：

- 本篇报告中，我们使用多因子框架构建了沪深 300 指数增强模型。多因子框架的建立主要包括四个步骤：数据预处理、单因子检测、收益模型和风险模型。原始数据构成的因子经过数据对齐、去极值、标准化、缺失值处理的预处理步骤，进入单因子检测环节。单因子检测分为统计检验和分层回测两种。通过单因子检测后，我们采用逐步回归法，为收益模型逐个引入新因子，根据拟合优度的统计量决定新引入的因子对整个模型的贡献程度。最后，使用 Barra 风险模型控制整个组合的风险，以及在特定风险因子上的暴露。
- 针对沪深 300 和中证 500 成分股，最终选取估值因子、盈利因子、成长因子、动量因子和流动性因子作为收益预测因子构建收益模型，选取波动率因子、市值因子、中性市值因子、Beta 因子和行业因子作为风险因子代入风险模型。
- 我们针对沪深 300 成分股，以月度调仓的频率构建投资组合。回测数据为 2011 年 2 月-2019 年 1 月。收益模型采用 12 个月移动平均的方式预测因子收益，风险模型中，在二次规划的条件部分对风险因子的暴露设定为 0。经过回测，模型年化超额收益 10.15%，并在 2011-2018 期间每一年都跑赢指数，平均跟踪误差 3.55%，平均换手率 4.09 倍。
- 未来，我们将持续关注该模型的运行情况，定期公布净值。除了沪深 300 指数增强模型，我们还将构建中证 500 指数增强模型，以及其他指数增强模型。我们还将构建使用机器学习方法预测股票收益的指数增强模型。
- 风险提示：随着市场环境变化，模型存在失效风险。

目 录

1. 理论简介.....	4
2. 多因子模型建立流程.....	5
2.1 数据预处理	5
2.2 单因子检测	7
2.3 收益预测模型	10
2.4 风险预测模型	14
3. 沪深 300 指数增强模型建立	15
4. 总结与未来研究方向展望	17

表目录

表 1: 因子定义	10
表 2: 最终入选因子	15
表 3: 模型分年度收益情况	16
表 4: 模型分年度业绩归因	16

图目录

图 1: 多因子模型建立流程图	5
图 2: 数据预处理结果展示	7
图 3: 单因子统计检测结果展示（以成交量 1 月波动率为例）	9
图 4: 单因子分层回测结果对比	10
图 5: 选股模型相对沪深 300 净值曲线	16

1. 理论简介

Barra 结构化风险模型是目前全球最知名的多因子模型之一。根据 Barra 手册的内容，多因子模型被分为两部分，收益模型与风险模型，收益模型的基本表达形式如下：

$$\tilde{r}_i = \sum_j X_{i,j} \cdot \tilde{f}_j + \tilde{u}_i$$

其中：

$\tilde{r}_i$ ：股票 i 下一期的预期收益率；

$X_{i,j}$ ：股票 i 在因子 j 上的因子暴露；

$\tilde{f}_j$ ：因子 j 的因子收益率（通过回归模型估计）；

$\tilde{u}_i$ ：股票 i 的残差收益率；

已知股票在每个因子上的因子暴露，通过多因子的收益预测模型，估计各个因子的因子收益率，从而得到股票的预期收益率，这就是多因子收益模型的主要思路。

多因子风险模型的基本思路为，通过估计因子的协方差矩阵，刻画股票池未来的波动风险。而后对选股结果以及股票配置仓位进行二次优化：

$$\max \quad \alpha'w - \frac{1}{2}\lambda w'\Sigma w \quad (1)$$

$$s.t. \quad f_l \leq X_f \cdot (w - w_b) \leq f_u \quad (2)$$

$$h_l \leq H \cdot (w - w_b) \leq h_u \quad (3)$$

$$0 \leq w_i \leq k_i \quad (4)$$

$$\mathbf{1}'w = 1 \quad (5)$$

其中 w 为待求解的组合权重，(1) 为待优化的目标函数，(2) - (6) 需满足的条件限制。

1) $\alpha'w$ 为收益预测模型给出的组合收益预测， $w'\Sigma w$ 为组合风险预测，其中 Σ 为

风险预测模型给出的个股协方差预测矩阵， λ 为风险厌恶系数，取值越大，组合越偏向于保守风格；

- 2) X_f 为组合中股票的因子暴露矩阵， w_b 为基准指数的股票权重，该条件限制了组合对于每个因子相对于基准指数的风格偏离范围，防止组合选出的股票风格过于雷同（例如有的收益预测模型给出的组合在市值因子上暴露过高，导致组合大中小盘股配比失衡）；
- 3) H 为组合的行业暴露矩阵，当投资目标为行业中性时需要引入该限制条件，以限制组合相对于基准指数行业配置的偏离范围，使组合的行业配比与基准指数一致；
- 4) 第四个条件限制了每只股票权重的上下限，因为A股不允许做空，一般个股权重下限设为0，为防止收益过于集中，有时也会对个股权重设定上限；
- 5) 第五个条件限定权重之和为1。

2. 多因子模型建立流程

多因子模型的建立主要包括四个步骤：数据预处理、单因子检测、收益模型和风险模型。

图 1：多因子模型建立流程图



资料来源：渤海证券研究所

2.1 数据预处理

提取的因子数据需经过数据对齐、去极值、标准化、缺失值处理等步骤，才可进入下一步的选股模型。

数据对齐：上市公司财报的报告期和报告发布日期之间有一定延迟，为避免未来

信息，在提取数据的时候，需要对日期进行修正，保证因子数据为当时能获取的最新财报数据。

去极值：为避免数据中的极端值对回归结果产生过多影响，我们使用“中位数去极值法”，将超过上下限的极端值用上下限值代替。

$$\tilde{x}_i = \begin{cases} x_M + 5 \times x_{MAD}, & x_i > x_M + 5 \times x_{MAD} \\ x_i, & x_M - 5 \times x_{MAD} \leq x_i \leq x_M + 5 \times x_{MAD} \\ x_M - 5 \times x_{MAD}, & x_i < x_M - 5 \times x_{MAD} \end{cases}$$

x_i : 原始序列

x_M : 序列 x_i 的中位数

x_{MAD} : 序列 $|x_i - x_M|$ 的中位数

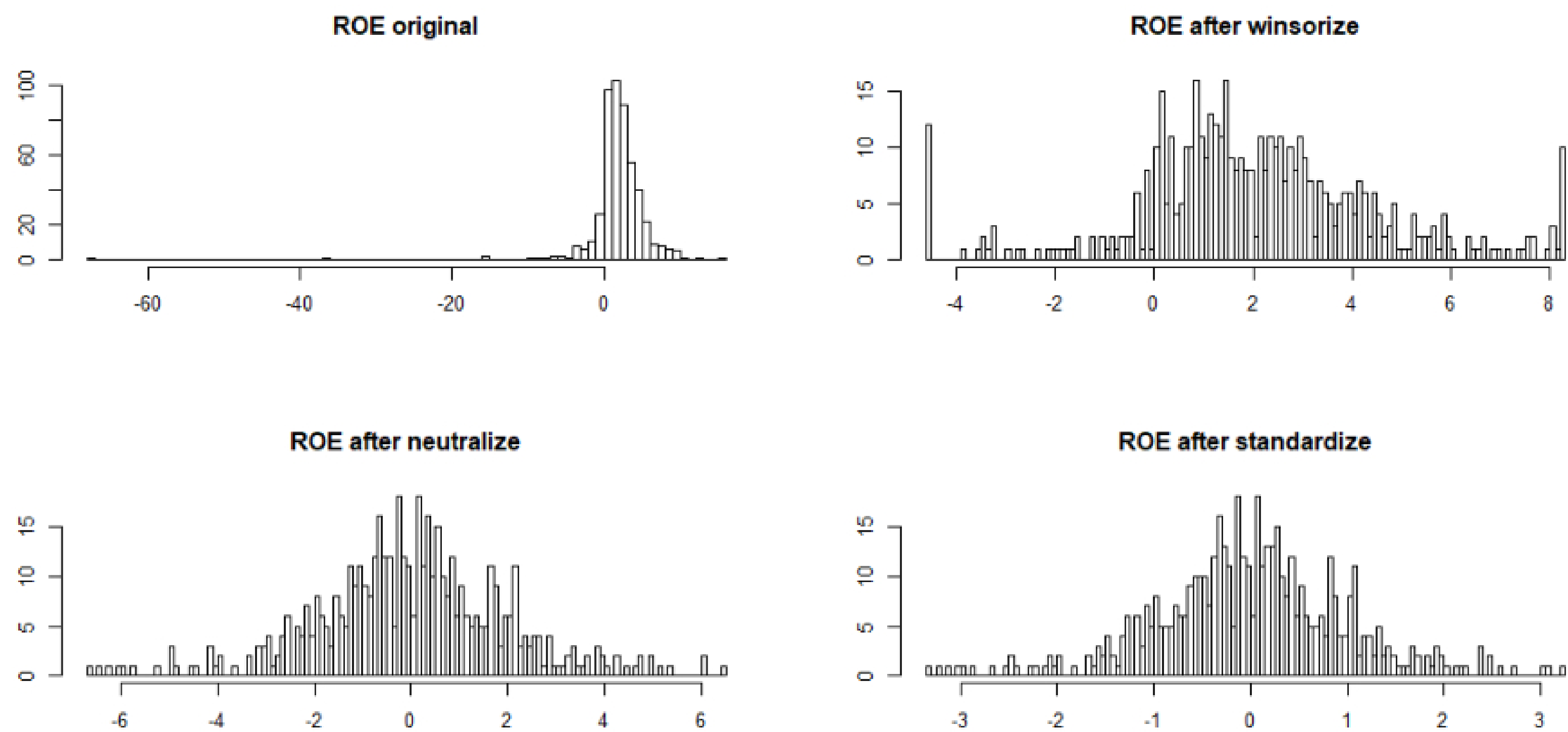
$\tilde{x}_i$: 去极值处理后的新序列

缺失值处理：提取出的因子可能会因为技术原因等情况出现缺失值，在缺失值少于 10%的情况下，因子缺失值使用行业中位数代替，当缺失值过多时，考虑更换数据源或使用其他因子。

标准化：由于各个因子的单位不同，为了使其具有可比性，需要对其进行 ZScore 标准化处理，即减去序列均值除以序列标准差，使因子序列近似成为一个符合 $N(0,1)$ 正态分布的序列。

行业市值中性化：将最后得到的因子序列对流动市值与行业哑变量做线性回归，取残差作为新的因子值。

图 2：数据预处理结果展示



资料来源：Wind，渤海证券研究所

2.2 单因子检测

单因子检测分为统计检验和分层回测两种。在统计检验中，根据 Barra 手册中关于因子显著性测试的内容，对因子进行横截面回归，同时考虑行业与市值因素的影响：

$$r_i^{T+1} = \sum_j X_j^T f_{i,j}^T + x_{size}^T f_{size,i}^T + x_d^T f_{d,i}^T + u_i^T$$

r_i^{T+1} ：股票 i 在 T+1 期的收益率；

$f_{i,j}^T$ ：第 T 期，股票 i 的行业虚拟变量，如股票 i 在 j 行业，则取值为 1，如不在，则取值为 0。这里行业分类选用了申万一级行业分类；

$f_{size,i}^T$ ：第 T 期，股票 i 的流通市值；

$f_{d,i}^T$ ：第 T 期，股票 i 在 d 因子上的因子暴露；

X_j^T 、 x_{size}^T 、 x_d^T ：回归模型运算所得的因子收益率；

u_i^T ：回归模型运算所得股票 i 的残差；

在每一个横截面上使用上述模型进行加权最小二乘回归 (WLS)，权重采用流通市值的平方根，一定程度上消除了异方差性。经过回归模型，我们可以得到 t 检验的 t 值序列与因子收益序列 x_d^T 。

下一步，计算因子 IC 值。

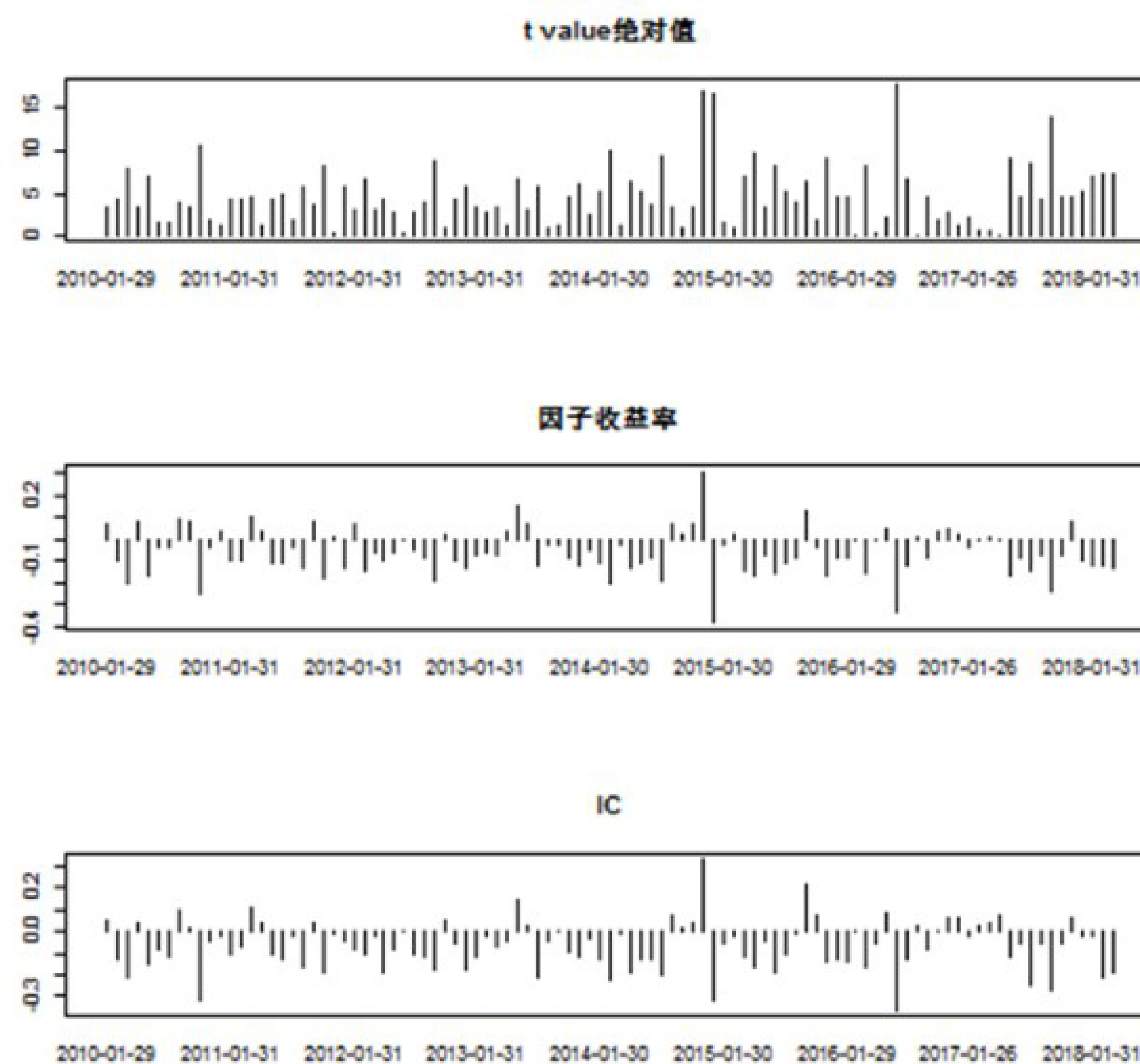
信息比率 IC (Information Coefficient) 是衡量因子收益预测能力的重要参数，它的计算方法是将每一期的因子值作为因变量，与行业哑变量和市值变量进行回归，取其残差，作为剔除行业与市值影响后的因子值。再计算新因子值与下一期股票收益序列间的 Spearman 相关系数。

最终，回归模型输出以下几个指标：

- a) t 值绝对值平均值：衡量因子整体显著性的指标；
- b) t 值绝对值 > 2 概率：衡量因子显著性是否稳定；
- c) 因子收益平均值：衡量因子收益能力大小的指标；
- d) 因子收益标准差：衡量因子收益能力波动率的指标；
- e) 因子收益 t 值：衡量因子收益率统计上是否显著不为 0 的指标；
- f) 因子收益 > 0 概率：衡量因子收益率方向性是否一致的指标；
- g) IC 平均值：衡量模型预测能力的指标；
- h) IC 标准差：衡量模型预测能力是否稳定的指标；
- i) IRIC: IC 平均值/IC 标准差；
- j) IC > 0 概率：衡量模型预测收益方向性是否一致的指标。

在选取因子时，我们希望首先筛选出 t 值绝对值平均值大于或接近 2，因子收益与 IC 值的平均值较大，标准差较小，因子收益与 IC 大于 0 的概率接近 0 或 1（而不是在 0.5 附近）的因子。当然，选择因子的标准也不是绝对的，例如 Barra 手册中收录的 beta 因子与市值因子，IC 大于 0 的概率都接近 50%，但却具有较高的 t 值，这意味着因子对选股收益的影响显著，但是方向却并不稳定，在以风险控制为主要目标的模型中依然是需要考虑的一类因子。

图 3：单因子统计检测结果展示（以成交量 1 月波动率为例）



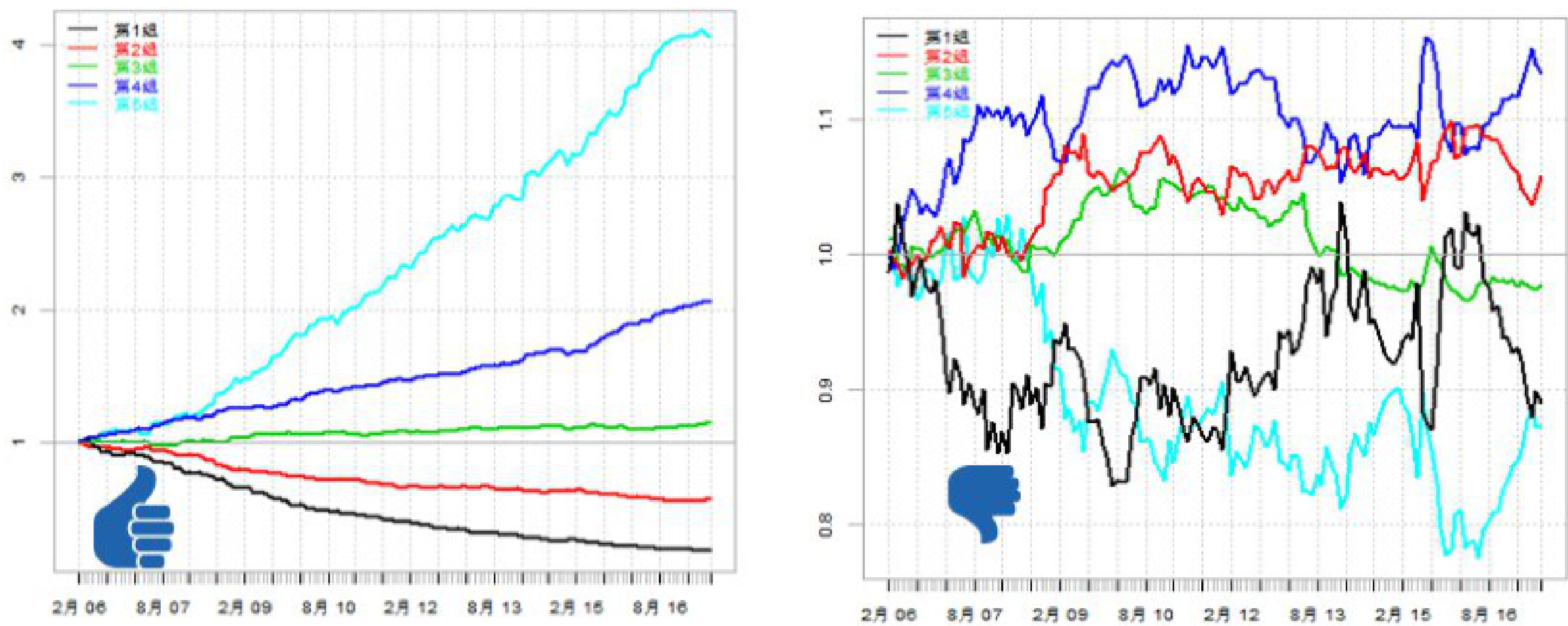
资料来源：Wind，渤海证券研究所

除了统计检验之外，一个更为直观观察因子选股能力的方式就是观察因子值高的股票与因子值低的股票在走势上的不同之处，即分层回测模型。模型构建方法如下：

在每个截面期的最后一个交易日，提取样本内股票因子值，并剔除因子值缺失的股票。按照因子将样本内股票排序，并按照序号从大到小平均分为 N 组（本报告中 $N=5$ ）。在下一个截面期的首个交易日，以当天的收盘价换仓并剔除当天因停牌、涨停等因素不能交易的股票。股改股票由于在复牌当日不受涨跌停板限制，可能出现极端涨幅，影响回测结果，故在复牌当月同样剔除股票池。对 N 组股票的历史收益率进行回测，并计算其年化收益率、波动率、夏普比率等值。

如果情况理想， N 组股票的收益率会呈现较好的单调性，如 12345 或 54321，且每一组股票间的收益差距较大。这样的因子在选股上体现为较好的区分度。

图 4：单因子分层回测结果对比



资料来源：Wind，渤海证券研究所

2.3 收益预测模型

经过单因子检测后，我们将以下因子列入因子库：

表 1：因子定义

大类因子	小类因子	因子解释
反转	RSTR_barra	Barra 因子； $\sum_{t=L}^{T+L} w_t [\ln (1+r_t)]$ ，L=21，T=500，半衰期 120 日
	RSTR_m24	$\sum_{t=L}^{T+L} w_t [\ln (1+r_t)]$ ，L=1，T=240，半衰期 120 日
	RSTR_m12	$\sum_{t=L}^{T+L} w_t [\ln (1+r_t)]$ ，L=1，T=240，半衰期 60 日
	RSTR_m6	$\sum_{t=L}^{T+L} w_t [\ln (1+r_t)]$ ，L=1，T=120，半衰期 30 日
	RSTR_m3	$\sum_{t=L}^{T+L} w_t [\ln (1+r_t)]$ ，L=1，T=60，半衰期 15 日
	RSTR_m1	$\sum_{t=L}^{T+L} w_t [\ln (1+r_t)]$ ，L=1，T=20，半衰期 5 日
动量	RS_1	最新收盘价/21 个交易日前收盘价
	RS_3	最新收盘价/63 个交易日前收盘价
	RS_6	最新收盘价/126 个交易日前收盘价
	RS_12	最新收盘价/252 个交易日前收盘价
	Alpha	alpha 系数；个股收益率序列与沪深 300 指数收益率序列以半衰期指数加权，得到 alpha 系数，半衰期为 60 日
市值	Size	总市值
	non-linear-size	Barra 因子；中等市值；将总市值的对数与总市值立方的对数回归得到残差，再对残差做标准化处理
	marketcap	流通市值
流动性	STOM	月度平均换手率；最近一个月的交易量/流通股数
	STOQ	季度平均换手率；最近一季度的交易量/流通股数
	STOS	半年平均换手率；最近半年的交易量/流通股数
	STOA	年度平均换手率；最近一年的交易量/流通股数

波动率	STOM_barra	Barra 因子; 公式: $\ln \left(\sum_{t=1}^{21} \frac{V_t}{S_t} \right)$, V_t 为 t 日成交金额, S_t 为 t 日流动市值
	STOQ_barra	Barra 因子; 公式: $\ln \left[\frac{1}{T} \sum_{t=1}^T \exp (STOM_t) \right]$, T=63 个交易日
	STOA_barra	Barra 因子; 公式: $\ln \left[\frac{1}{T} \sum_{t=1}^T \exp (STOM_t) \right]$, T=244 个个交易日
	Ins	机构持股比例; 机构持股变动/总股本
	ins_c	机构持股比例变动
	MSM	一个月换手率变动; 最近 1 个月换手率/最近 1 年换手率
	MSQ	季度换手率变动; 最近 3 个月换收益率/最近 1 年换手率
	MSS	半年换手率变动; 最近 6 个月换收益率/最近 1 年换手率
	DASTD	Barra 因子; 年度平均波动率; 累计日波动率以半衰期指数加权, 半衰期为 40 日
	CMRA	Barra 因子; 年度收益率波动
	HSIGMA	Barra 因子; sigma: 个股收益率序列与沪深 300 指数收益率序列以半衰期指数加权, 得到残差, 对残差求标准差得到 sigma, 半衰期为 60 日
	Beta	Barra 因子; 贝塔系数; 个股收益率序列与沪深 300 指数收益率序列以半衰期指数加权, 得到 beta 系数, 半衰期为 60 日
	yieldvol_1	月度日收益率波动率; 一个月日收益率标准差
	yieldvol_3	季度日收益率波动率; 三个月日收益率标准差
	yieldvol_6	半年日收益率波动率; 半年日收益率标准差
	high_low_1	月度股价波动; 最高价 /最低价 (最近 一个月 内价格)
	high_low_3	季度股价波动; 最高价 /最低价 (近三个月 内价格)
	high_low_6	半年股价波动; 最高价 /最低价 (近六个月 内价格)
	high_low_12	全年股价波动; 最高价 /最低价 (近十二个月 内价格)
成长	VOL_1	成交量月度波动率; 1 月波动率标准差
	VOL_3	成交量季度波动率; 3 月波动率标准差
	VOL_6	成交量半年波动率; 6 月波动率标准差
	VOL_12	成交量年度波动率; 12 月波动率标准差
	growth_ttm_or	营业收入 ttm 一年增长率
	growth_ttm_profit	净利润 ttm 一年增长率
	qfa_yoy sales_qq	单季度营业收入一年增长率
	qfa_yoy netprofit_qq	单季度归母净利润一年增长率
	qfa_yoy ocf_qq	单季度经营性现金流一年增长率
	qfa_yoy profit_qq	单季度净利润一年增长率
	qfa_roe_qq	单季度 roe 一年增长率

盈利	yoy_profit_qq	净利润一年增长率
	yoy_growth_netprofit_qq	归母净利润一年增长率
	yoy_or_qq	营业收入一年增长率
	yoyroe_qq	roe 一年增长率
	yoyocf_qq	经营性现金流一年增长率
	growth_roe_qq_3	roe 三年增长率
	growth_netprofit_qq_3	归母净利润三年增长率
	growth_or_qq_3	营业收入三年增长率
	growth_profit_qq_3	净利润三年增长率
	growth_ocf_qq_3	经营性现金流三年增长率
	growth_profit_qq_5	净利润五年增长率
	growth_ocf_qq_5	经营性现金流五年增长率
	growth_roe_qq_5	roe 五年增长率
	SGRO	Barra 因子; 过去 5 年企业营业收入复合增长率
	EGRO_5	Barra 因子; 过去 5 年企业归属母公司净利润复合增长率
	EGIB	Barra 因子; 未来 3 年企业一致预期净利润增长率
	EGIB_S	Barra 因子; 未来 1 年企业一致预期净利润增长率
	CETOP	Barra 因子: 个股现金收益比股票价格
	roe_q	当季净资产收益率
	roe_ttm	滚动 ROE
	roa_q	当季资产收益率、资产回报率
	roa_ttm	滚动 ROA
	qfa_grossprofitmargin	当季毛利率
	grossprofitmargin_ttm	滚动毛利率
	profitmargin_q	当季扣非后利润率
	profitmargin_ttm	滚动扣非后利润率
	assetsturn_q	当季资产周转率
	assetturn_ttm	滚动资产周转率
	operationcashflowratio_q	当季经营活动现金净流比率
	operationcashflowratio_ttm	滚动经营活动现金净流比率
	ROIC	投入资本回报率
	EBIT2EV	投资收益率; 息税前利润/投入资本
	CASHROIC	现金 ROIC

价值	FreeCashflowYield	自由现金流比率；经营性活动产生的净现金流-构建其他长期资产支付的现金/总市值
	sales2EV	营业收益率；营业收入_TTM/总市值+非流动负债
	cashflow1	经营活动净现金流/总市值
	cashflow2	经营活动净现金流/营业收入
	cashflow3	经营活动净现金流/营业收入净收益
	invturn_qq	存货周转率；存货成本/平均存货余额
	arturn_qq	应收账款周转率；当期销售净收入/平均应收账款
	fatur_n_qq	固定资产周转率；销售收入/平均固定资产
	assetturnover_ttm	滚动总资产周转率；营业收入_ttm / [(期初资产总额+期末资产总额) / 2]
	assetsturn_qq	总资产周转率；营业总收入 / [(期初资产总额+期末资产总额) / 2]
	longdebtworkingcapit al_qq	长期债务与营运资金比率；长期债务/营运资本
	finaexpensetogr_qq	财务费用比率；财务费用/主营业务收入
	gctogr_qq	营业费用比率；营业费用/主营业务收入
	ETOP	Barra 因子；历史 EP 值；利用过去 12 个月个股净利润除以当前市值。
	BTOP	Barra 因子；历史 BP 值；普通股权益价值/市值
	epcut	市盈率（扣除非经营性损益部分，即公司经营性盈利与市值之比值）
	bp_lf	最近公告日 BP
	ncfp	净现金市值比；净现金流 / 总市值
	ocfp	营业现金流比率；经营性现金流 / 总市值
	dividendyield	股息率；过去一年分红/总市值
	stop1	营收市值比，市销率 PS（TTM）倒数
	stop2	营收市值比，市销率 PS（LYR）倒数
	ep_rel	相对 PE；PE/行业 PE
	bp_rel	相对 PB；PB/行业 PB
	PEG	市盈率相对盈利增长比率
	EPIBS	Barra 因子；预期 EP 值

资料来源：渤海证券研究所、The Barra China Equity Model (CNE5)

在线性回归时，过多因子会导致过度拟合的问题，从而降低模型预测的准确率。因此在构建收益预测模型时，需要对因子进行二次筛选。这里我们采用逐步回归法，逐个引入新因子，构建回归模型，计算拟合优度（如 AIC、BIC、交叉验证等）。根据拟合优度的统计量决定新引入的因子对整个模型的贡献程度，如果贡献程度大，则将该因子加入模型，如果不够大，则跳过该因子，测试下一因子。

筛选因子与计算因子收益不同，因子收益每月更新，单考虑到模型的稳定性，因子变动不宜过于频繁，我们采用每年更新一次的频率。

选定模型所用因子后，需要判断因子间是否存在多重共线性。判断多重共线性可以使用相关性矩阵，也可以使用 VIF 检验。对于相关性过高的因子，有如下三种处理方案：

请务必阅读正文之后的免责条款部分

- 1) 对过于相似的因子（如 PE 和扣非 PE），可选取单因子检验效果最好的因子进入模型，剔除其他因子；
- 2) 对内在相关因子（如月度、季度、年度换手率），可选用等权、历史收益率加权、ICIR 加权等方法，对因子进行合并，生成新的因子；
- 3) 对经济学不相关因子（如估值与市值），可选用因子正交方法，剔除其相互影响。

接下来，根据上文所提到的收益预测模型公式，回归得到因子的历史收益。并采用移动平均，或指数加权移动平均的方式，预测下一期因子收益，从而预测下一期股票的收益率。

2.4 风险预测模型

风险模型的一种估计方法，是直接选取组合中股票的历史收益率协方差矩阵作为 Σ 。但由于最终优化问题中要求协方差矩阵可逆，历史收益率需要考虑的时间跨度个数 N 必须大于组合中的股票数量 M 。对于全市场几千只股票来说，如果使用月度收益率来计算协方差矩阵，则所需的回溯期 N 要超过几千个月，这对于历史相对短暂、市场风格变化迅速的 A 股来说比较困难。即使改为使用日度因子，需要估计的参数数量依然在百万级别，误差过大，回测结果并不理想。

另一种风险模型的估计方法来源于 Barra 模型，它将组合协方差矩阵 Σ 分为两个部分：

$$\Sigma = X_f F X_f' + \Delta$$

其中 X_f 为组合中股票的因子暴露矩阵， F 为因子收益率之间的协方差矩阵， Δ 为个股残差波动率组成的对角矩阵。

Barra 风险模型将个股的风险转化为可以用多因子模型的系统性风险以及股票本身的残差风险之和。这样做可以大量减少需估计参数的数量，从估计几千支股票的收益协方差矩阵（数量级为 N^2 ），变为估计几十个因子收益协方差矩阵以及个股的残差（数量级为 N ）。

经过上文提到的因子筛选过程，我们最终确定了模型所选用的收益因子和风险因子：

表 2：最终入选因子

	因子类型	最终入选因子
收益因子	估值因子	BP
	盈利因子	单季度 ROE
	成长因子	单季度营业收入增速、单季度归母净利润增速、单季度 ROE 增速
	动量因子	20 日特色反转因子、60 日特色反转因子、180 日特色反转因子
	流动性因子	Barra 月度换手率、Barra 季度换手率、Barra 年度换手率
风险因子	波动率因子	1、3、6、12 月成交量波动率
	市值因子	流通市值对数
	中性市值因子	非线性市值对数
	Beta 因子	相对业绩基准 Beta 因子
	行业因子	中信行业因子

数据来源：渤海证券研究所、Wind

3. 沪深 300 指数增强模型建立

根据上述理论，我们针对沪深 300 成分股，以月度调仓的频率构建投资组合。回测数据为 2011 年 2 月-2019 年 1 月。

我们选取表 2 因子构建收益模型与风险模型，其中收益模型采用 12 个月移动平均的方式预测因子收益，风险模型中，在二次规划的条件部分对风险因子的暴露设定为 0。

根据指数增强模型的特点，在取得超额收益的同时，尽量控制跟踪误差：

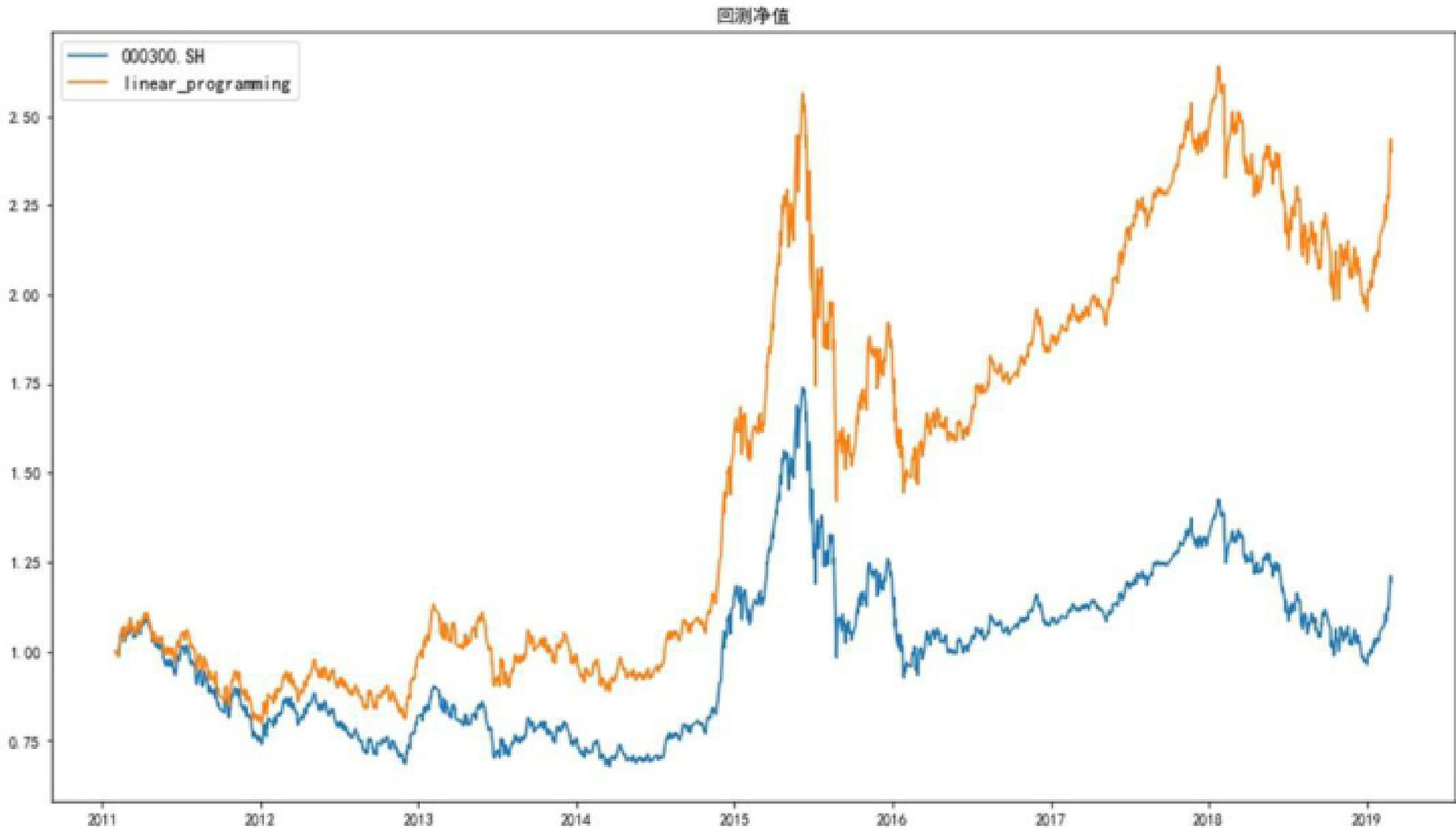
$$TD_{ti} = R_{ti} - R_{tb}$$

$$TE = \sqrt{\frac{1}{n-1} \sum_{t=1}^n [TD_{ti} - \overline{TD_{ti}}]^2}$$

化跟踪误差要求在 7.75%以内。

经过回测，模型年化超额收益 10.15%，并在 2011-2019 期间每一年都跑赢指数，平均跟踪误差 3.55%，平均换手率 4.09 倍。模型回测结果如下：

图 5：选股模型相对沪深 300 净值曲线



资料来源：Wind，渤海证券研究所

表 3：模型分年度收益情况

	年度超额收 益	年度收益	年度波动	夏普比率	信息比率	最大回撤	日胜率	跟踪误差	换手率
2011	10.03%	-16.20%	19.73%	-1.02	2.39	26.62%	61.61%	2.53%	3.94
2012	12.46%	23.38%	20.44%	0.95	3.19	17.08%	59.67%	2.65%	3.87
2013	8.15%	-0.33%	22.66%	-0.19	1.03	20.49%	53.36%	4.04%	4.45
2014	6.49%	62.13%	19.07%	3.05	0.71	8.63%	56.33%	3.49%	3.76
2015	10.14%	11.91%	41.28%	0.19	1.14	43.64%	54.51%	5.43%	4.14
2016	14.87%	9.82%	21.13%	0.28	3.55	18.01%	61.07%	3.07%	3.40
2017	8.29%	30.59%	10.43%	2.55	1.48	5.78%	59.43%	2.89%	3.67
2018	9.27%	-19.48%	21.05%	-1.12	1.94	25.44%	61.73%	2.72%	3.54
总计	10.15%	10.44%	23.93%	0.27	1.73	43.64%	58.41%	3.55%	4.09

数据来源：渤海证券研究所、Wind

对回测结果进行业绩归因，可以看出模型在各因子上表现较为平均，在大部分年份会比较偏重于 ROE 和成长因子。

表 4：模型分年度业绩归因

	BP	ROE_q	growth	liq_barra	mom
2011/12/31	-0.034	0.481	0.287	0.021	-0.241
2012/12/31	0.014	0.418	0.245	0.067	-0.221
2013/12/31	-0.058	0.394	0.203	0.064	-0.317
2014/12/31	0.088	0.359	0.203	0.138	-0.180
2015/12/31	0.605	0.150	0.152	0.028	-0.009
2016/12/31	0.549	0.207	0.051	0.035	-0.092
2017/12/31	0.187	0.240	0.116	-0.132	-0.193
2018/12/31	0.051	0.269	0.075	-0.108	-0.150

数据来源：渤海证券研究所、Wind

4. 总结与未来研究方向展望

本篇报告中，我们使用多因子框架构建了沪深 300 指数增强模型。未来，我们将持续关注该模型的运行情况，定期公布净值。除了沪深 300 指数增强模型，我们还将构建中证 500 指数增强模型，以及其他指数增强模型。我们还将构建使用机器学习方法预测股票收益的指数增强模型。

未来，我们会继续研究更多的量化模型，探索其适用范围、运行机制、收益来源以及成果的延续能力，并将其应用扩展到更多领域，如行业轮动、资产配置等。

风险提示：随着市场环境变化，模型存在失效风险。

投资评级说明

项目名称	投资评级	评级说明
公司评级标准	买入	未来 6 个月内相对沪深 300 指数涨幅超过 20%
	增持	未来 6 个月内相对沪深 300 指数涨幅介于 10%~20%之间
	中性	未来 6 个月内相对沪深 300 指数涨幅介于-10%~10%之间
	减持	未来 6 个月内相对沪深 300 指数跌幅超过 10%
行业评级标准	看好	未来 12 个月内相对于沪深 300 指数涨幅超过 10%
	中性	未来 12 个月内相对于沪深 300 指数涨幅介于-10%-10%之间
	看淡	未来 12 个月内相对于沪深 300 指数跌幅超过 10%

免责声明：本报告中的信息均来源于已公开的资料，我公司对这些信息的准确性和完整性不作任何保证，不保证该信息未经任何更新，也不保证本公司做出的任何建议不会发生任何变更。在任何情况下，报告中的信息或所表达的意见并不构成所述证券买卖的出价或询价。在任何情况下，我公司不就本报告中的任何内容对任何投资做出任何形式的担保，投资者自主作出投资决策并自行承担投资风险，任何形式的分享证券投资收益或者分担证券投资损失书面或口头承诺均为无效。我公司及其关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供或争取提供投资银行或财务顾问服务。我公司的关联机构或个人可能在本报告公开发表之前已经使用或了解其中的信息。本报告的版权归渤海证券股份有限公司所有，未获得渤海证券股份有限公司事先书面授权，任何人不得对本报告进行任何形式的发布、复制。如引用、刊发，需注明出处为“渤海证券股份有限公司”，也不得对本报告进行有悖原意的删节和修改。

渤海证券股份有限公司研究所

所长&金融行业研究

张继袖

+86 22 2845 1845

副所长&产品研发部经理

崔健

+86 22 2845 1618

计算机行业研究小组

王洪磊（部门经理）

+86 22 2845 1975

张源

+86 22 2383 9067

汽车行业研究小组

郑连声

+86 22 2845 1904

陈兰芳

+86 22 2383 9069

食品饮料行业研究

刘瑀

+86 22 2386 1670

电力设备与新能源行业研究

张冬明

+86 22 2845 1857

刘秀峰

+86 10 6810 4658

滕飞

+86 10 6810 4686

医药行业研究小组

赵波

+86 22 2845 1632

甘英健

+86 22 2383 9063

陈晨

+86 22 2383 9062

通信行业研究小组

徐勇

+86 10 6810 4602

公用事业行业研究

刘蕾

+86 10 6810 4662

餐饮旅游行业研究

刘瑀

+86 22 2386 1670

杨旭

+86 22 2845 1879

非银金融行业研究

洪程程

+86 10 6810 4609

中小盘行业研究

徐中华

+86 10 6810 4898

机械行业研究

张冬明

+86 22 2845 1857

传媒行业研究

姚磊

+86 22 2383 9065

电子行业研究

王磊

+86 22 2845 1802

固定收益研究

冯振

+86 22 2845 1605

夏捷

+86 22 2386 1355

朱林宁

+86 22 2387 3123

李元玮

+86 22 2387 3121

金融工程研究

宋旻

+86 22 2845 1131

李莘泰

+86 22 2387 3122

张世良

+86 22 2383 9061

金融工程研究

祝涛

+86 22 2845 1653

郝惊

+86 22 2386 1600

流动性、战略研究&部门经理

周喜

+86 22 2845 1972

策略研究

宋亦威

+86 22 2386 1608

严佩佩

+86 22 2383 9070

机构销售•投资顾问

朱艳君

+86 22 2845 1995

刘璐

宏观研究

宋亦威

+86 22 2386 1608

孟凡迪

+86 22 2383 9071

博士后工作站

张佳佳 资产配置

+86 22 2383 9072

张一帆 公用事业、信用评级

+86 22 2383 9073

综合管理&部门经理

齐艳莉

+86 22 2845 1625

合规管理&部门经理

任宪功

+86 10 6810 4615

风控专员

白骐玮

+86 22 2845 1659

渤海证券研究所

天津

天津市南开区水上公园东路宁汇大厦 A 座写字楼

邮政编码: 300381

电话: (022) 28451888

传真: (022) 28451615

北京

北京市西城区西直门外大街甲 143 号 凯旋大厦 A 座 2 层

邮政编码: 100086

电话: (010) 68104192

传真: (010) 68104192

渤海证券研究所网址: www.ewww.com.cn