

2016-10-23

金融工程 (专题报告)

选股模型

多因子模型系列报告之二 —— 利用纯因子组合检验因子有效性

分析师 秦瑶

☎ (8627)65799830

✉ qinyao@cjsc.com.cn

执业证书编号: S0490513080002

联系人 邓越

☎ 02765799830

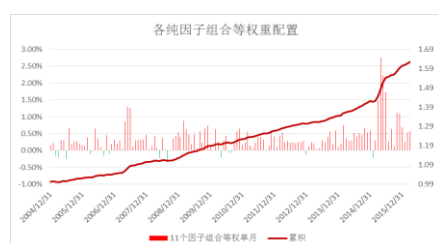
✉ dengyue1@cjsc.com.cn

联系人 杨靖凤

☎ (8621) 68751636

✉ yangjf@cjsc.com.cn

市场表现对比图



资料来源: 天软科技

相关研究

《基于盈利预测的选股与增强 (1)》2016-9-21

《筹码分布中的 Alpha: 交易行为的非对称性选股》2016-9-4

《择时视角: 多周期与多维度》2016-8-22

报告要点

- 纯因子组合对模型中某个因子具有单位暴露度, 对其他因子暴露度为 0

纯因子组合对模型中某个因子具有单位暴露度, 对其他因子暴露度为 0, 对因子显著性检验、投资收益事后分析和多因子组合构建均有重要作用。

- 纯因子组合避免了传统的排序多空组合可能无意识暴露到其它因子上的问题

检验因子效用的通常做法是将股票按照因子值分作多头组合和空头组合, 然后观察这个多/空组合的历史收益, 然而这样做的弊端是该多空组合有可能对其他因子有较多未意识到的暴露, 比如检验流动性因子或成长性因子时其实较多暴露在了规模因子上。纯因子组合是控制了对其他因子的暴露度后的多/空投资组合, 考察其历史收益会更有说服力。

- 从风险收益比来看, 流动性因子历史表现最佳, 反转因子和规模因子次之

流动性因子风险收益比为 2.24, 反转和规模因子风险收益比为 1.64 和 1.50。

目录

一、 构造纯因子组合	5
1.1 模型的形式	5
1.2 纯因子组合	5
二、 计算因子值的细节和模型具体形式	6
2.1 合成大类风格因子	6
2.2 模型的具体形式	8
三、 纯因子组合的历史收益	9
3.1 回测时间和样本空间	9
3.2 截距项因子收益	10
3.3 纯风格因子组合的收益	10
3.4 纯行业因子组合的收益	14
3.5 残差项的收益和模型的解释力	15
四、 纯因子组合风险特征	17
4.1 纯因子组合的历史波动率	17
4.2 纯因子组合的协方差	17
五、 检验一些常见因子	19
5.1 被排序法高估的因子	19
5.2 被排序法低估的因子	31
附录 I---- 纯行业因子组合的收益	38

图表目录

图 1: 样本空间股票数量	9
图 2: 截距项因子收益和中证全指对比	10
图 3: 纯流动性因子组合收益	11
图 4: 纯反转因子组合收益	11
图 5: 纯规模因子组合收益	12
图 6: 纯成长因子组合收益	12
图 7: 纯非线性规模因子组合收益	12
图 8: 纯 BP 因子组合收益	12
图 9: 纯 beta 因子组合收益	12
图 10: 纯动量因子组合收益	12
图 11: 纯质量因子组合收益	13
图 12: 纯 EP 因子组合收益	13
图 13: 纯残差波动率因子组合收益	13
图 14: 各纯因子组合收益汇总	14

图 15: 模型解释力 -- 调整 R 方	16
图 16: 2014 年 12 月, $R^2\text{-adj} = 0.69$	16
图 17: 2014 年 4 月, $R^2\text{-adj} = 0.065$	16
图 18: 纯因子组合历史波动率	17
图 19: 纯因子组合协方差矩阵热力图	18
图 20: 纯因子组合协方差矩阵热力图 3D	18
图 21: 个股乖离度与各风险因子相关系数	19
图 22: 个股乖离度因子收益率被高估	19
图 23: 换手率变化与各风险因子相关系数	20
图 24: 换手率变化因子收益率被高估	20
图 25: 10 日 RSI 与各风险因子相关系数	21
图 26: 10 日 RSI 因子收益率被高估	21
图 27: 固定资产周转率与各风险因子相关系数	22
图 28: 固定资产周转率因子收益率被高估	22
图 29: 现金比率与各风险因子相关系数	23
图 30: 现金比率因子收益率被高估	23
图 31: 每股自由现金流量比 eps 与各风险因子相关系数	24
图 32: 每股自由现金流量比 eps 因子收益率被高估	24
图 33: 个股股东数与各风险因子相关系数	25
图 34: 个股股东数因子收益率被高估	25
图 35: 流动比率与各风险因子相关系数	26
图 36: 流动比率因子收益率被高估	26
图 37: 总市值比 Ebitda 与各风险因子相关系数	27
图 38: 总市值比 Ebitda 因子收益率被高估	27
图 39: 股东权益比与各风险因子相关系数	28
图 40: 股东权益比因子收益率被高估	28
图 41: 应付账款占比与各风险因子相关系数	29
图 42: 应付账款占比因子收益率被高估	29
图 43: 净利润增速(季度同比)与各风险因子相关系数	31
图 44: 净利润增速(季度同比)因子收益率被低估	31
图 45: 三费占比与各风险因子相关系数	32
图 46: 三费占比因子收益率被低估	32
图 47: 送股比例与各风险因子相关系数	33
图 48: 送股比例因子收益率被低估	33
图 49: 销售毛利率与各风险因子相关系数	34
图 50: 销售毛利率因子收益率被低估	34
图 51: PE/G 与各风险因子相关系数	35
图 52: PE/G 因子收益率被低估	35
图 53: 存货周转率与各风险因子相关系数	36
图 54: 存货周转率因子收益率被低估	36
图 55: 每股自由现金流量比 EPS 与各风险因子相关系数	37
图 56: 每股自由现金流量比 EPS 因子收益率被低估	37
图 I-1: 纯行业因子收益----化工、有色、采掘、钢铁	38

图 I-2: 纯行业因子收益----传媒、计算机、通信	39
图 I-3: 纯行业因子收益----房地产、银行、非银	39
图 I-4: 纯行业因子收益----家电、建筑材料、建筑装饰	40
图 I-5: 纯行业因子收益----公用事业、农林牧渔、医药生物	40
图 I-6: 纯行业因子收益----交通运输、国防军工、汽车、电子	41
图 I-7: 纯行业因子收益----休闲服务、商业贸易、食品饮料	41
图 I-8: 纯行业因子收益----纺织服装、综合、轻工制造	42
图 I-9: 纯行业因子收益----机械设备、电气设备	42
表 1: 合成大类因子的子类因子	6
表 2: 11 种风格因子暴露度的相关系数 (2005.1 – 2016.4 月平均)	9
表 3: 各风格纯因子收益情况汇总	11
表 4: 各行业纯因子收益情况汇总	15
表 5: 各风格纯因子的残差相抵程度 (月度平均值)	16
表 6: 排序法高估的一些因子的 IR	30
表 7: 排序法低估的一些因子的 IR	38

一、构造纯因子组合

本系列的第一篇报告《多因子模型系列报告之一——模型理论随想和纯因子组合构建》已经详细说明了本篇报告所需要的诸如模型的解、因子暴露度、纯因子组合等概念，在此仅作一个简要回顾。之后的主要内容是展示如何使用纯因子组合检验因子有效性。检验因子效用的通常做法是将股票按照因子值分作多头组合和空头组合，然后观察这个多/空组合的历史收益，然而这样做的弊端是该多空组合有可能对其他因子有较多未意识到的暴露，比如检验流动性因子或成长性因子时会较多暴露于规模因子上。纯因子组合是控制了对其他因子的暴露度后的多/空投资组合，考察其历史收益会更有说服力。

1.1 模型的形式

我们采用的是线性多因子模型：

$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix} f_1 + \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix} f_2 + \cdots + \begin{bmatrix} x_{1m} \\ x_{2m} \\ \vdots \\ x_{nm} \end{bmatrix} f_m + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

其中， r_i 为某股票扣除无风险收益率的超额部分， x_{ij} 为股票 i 对因子 j 的**暴露度**， f_i 为**因子的收益**。

1.2 纯因子组合

一个投资组合的收益可以**分解**为 m 个因子收益的线性组合：

$$r_p = x_p^1 f_1 + x_p^2 f_2 + \cdots + x_p^m f_m + w_p^T u$$

其中 $x_p^m = w_p^T x_m = \sum w_i x_{im}$ 是组合 P 对该因子的因子暴露度。如果组合充分分散，应该有 $w_p^T u \rightarrow 0$ 。

如果一个投资组合，对模型中某因子暴露度为 1，对其他风险因子暴露度为 0，则称该组合为这个因子的**纯因子组合**。

假设因子 m 对应的纯因子组合的权重为 w_{P_m} 。由定义，该组合对各个因子的暴露为， $x_{P_m}^1 = 0, x_{P_m}^2 = 0, \dots, x_{P_m}^{m-1} = 0, x_{P_m}^m = 1$ ，因此因子 m 对应的纯因子组合的收益为：

$$r_{P_m} = w_{P_m}^T r = x_{P_m}^1 f_1 + x_{P_m}^2 f_2 + \cdots + x_{P_m}^m f_m + w_{P_m}^T u = 1 \cdot f_m + w_{P_m}^T u$$

所以，对于充分分散的组合 m 就是因子 m 对应的纯因子组合的投资收益。将 f_m 称作“因子收益”是自然而然的。

在《多因子系列报告之一》中，已经说明了如何通过广义最小二乘法（GLS）的中间步骤求出纯因子组合中个股的权重 w_{P_m} 。求解权重 w_{P_m} 并不需要股票收益率，即权重在建仓时即可确定，而无法预知的是这个投资组合对应的**收益**。

二、计算因子值的细节和模型具体形式

2.1 合成大类风格因子

决定是否将因子加入模型应当首要考虑因子间的共线性问题，向模型中加入一个共线性的因子对模型的伤害要远大于因子自身带来的信息的价值（详见系列报告一）。为了克服共线性，益将描述股票同一类性质的多种因子合成大类因子，此处分别为成长、EP、BP、流动性、规模、非线性规模、beta、残差波动率、反转、动量、质量，合成方式见表 1。

表 1：合成大类因子的子类因子

大类因子	子类因子	合成权重
成长	每股营收同比增速	1/3
	每股收益同比增速	1/3
	一致预期净利润增速	1/3
EP	1/PE-ttm	1/3
	1/PE-ftm（未来12月一致预期）	1/3
	1/PCF-ttm	1/3
BP	1/PB（最近报告期）	1
流动性	近21交易日平均换手率的自然对数	1/3
	近63交易日平均换手率的自然对数	1/3
	近252交易日平均换手率的自然对数	1/3
质量	ROE-ttm	1/2
	资产负债率ttm	1/2
规模	流通市值的自然对数	1
非线性规模	规模因子的三次方，并与规模因子正交化	1
Beta	CAMP模型的beta	1
残差波动率	CAMP模型残差的标准差	1
动量	过去两年并除去最近一月收益率	1
反转	最近一月收益率	1

资料来源：长江证券研究所

各子类因子详细描述如下。

成长因子

$1/3$ 每股营收同比增速 + $1/3$ 每股收益同比增速 + $1/3$ 一致预期净利润增速

每股营收同比增速：最近报告期比上一年度同类报告期营业收入

每股收益同比增速：最近报告期比上一年度同类报告期 eps

一致预期净利润增速：wind 一致预期

EP 因子

$$\frac{1}{3} \frac{1}{PE_{ttm}} + \frac{1}{3} \frac{1}{PE_{ftm}} + \frac{1}{3} \frac{1}{PCF_{ttm}}$$

PE-ftm: wind 一致预期

BP 因子

1/PB 最近报告期

流动性因子

1/3 近 21 交易日 + 1/3 近 63 交易日 + 1/3 近 252 交易日 平均换手率的自然对数

$$\frac{1}{3} \ln \left(\sum_{t=1}^{21} \frac{V_t}{S_t} \right) + \frac{1}{3} \ln \left(\frac{1}{3} \sum_{t=1}^{63} \frac{V_t}{S_t} \right) + \frac{1}{3} \ln \left(\frac{1}{12} \sum_{t=1}^{252} \frac{V_t}{S_t} \right)$$

规模因子

$$\ln(\text{流通市值})$$

非线性规模因子

非线性规模因子是规模因子的 3 次方, 然后经过施密特正交化处理, 以去掉共线性的部分, 非线性规模因子主要强调市值中等的股票。

非线性规模因子的计算公式为:

$$nls = cs - \frac{cs^T Ws}{s^T Ws} s$$

其中, s 是规模因子, nls 是非线性规模因子, cs 是规模因子的 3 次方。

Beta 因子

通过时间序列回归计算 beta:

$$r_t^i = \alpha + \beta_i r_t^m + \varepsilon_t$$

其中, 收益率均为扣除无风险收益率后的收益。通过广义最小二乘法 (GLS) 计算 beta 值, 数据取过去 252 个交易日的日收益率数据, 回归权重为指数衰减加权, 半衰期 63 个交易日。

残差波动率

计算 beta 因子时的残差序列的标准差 (ε_t)

动量

过去两年(扣除最近 1 个月)的对数收益率的加权平均:

$$\sum_{t=21}^{t=504+21} w_t [\ln(1 + r_t) - \ln(1 + r_f)]$$

权重为指数衰减加权, 半衰期 126 个交易日。

从子类因子合成为大类因子的方式是：先对各个子类因子值进行极值处理（±5% winsorization），再计算 z-score，然后将各子类因子的 z-score 按表中权重合成为大类因子值，并再一次对合成得到的大类因子值正态标准化，作为最终的因子暴露值。

大类因子值正态标准化的方式也有讲究：

$$x_i^S = \frac{x_{raw_i}^S - m}{\mu}$$

其中， m 是 $x_{raw_i}^S$ 按流通市值加权的平均值 $\sum w_{cap}^i x_{raw_i}^S$ ， μ 是样本标准差。

这种标准化方式可以使以流通市值为权重的投资组合是对所有风格因子 0 暴露的，因为基准组合对风格因子的暴露：

$$w_{cap}^T x^S = \sum w_{cap}^i x_i^S = \frac{\sum w_{cap}^i x_{raw_i}^S - m}{std(x_{raw})} = 0$$

简言之，基准组合对各个因子的因子暴露为 0，非零的因子暴露度就是投资组合偏离了基准组合多少倍标准差。

2.2 模型的具体形式

我们采用的模型将收益分解为截距项因子、行业因子（哑元）和风格因子：

$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} f_c + \left\{ \begin{bmatrix} x_1^{I1} \\ x_2^{I1} \\ \vdots \\ x_n^{I1} \end{bmatrix} f_{I_1} + \cdots + \begin{bmatrix} x_1^{I28} \\ x_2^{I28} \\ \vdots \\ x_n^{I28} \end{bmatrix} f_{I_{28}} \right\} + \left\{ \begin{bmatrix} x_1^{S1} \\ x_2^{S1} \\ \vdots \\ x_n^{S1} \end{bmatrix} f_{S_1} + \cdots + \begin{bmatrix} x_1^{S11} \\ x_2^{S11} \\ \vdots \\ x_n^{S11} \end{bmatrix} f_{S_{11}} \right\} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

并有限制条件：

$$w_{I_1} f_{I_1} + w_{I_2} f_{I_2} + \cdots + w_{I_{28}} f_{I_{28}} = 0$$

其中 r 是扣除无风险收益率后的当期股票收益率， f_c 是截距项因子， I_1, \dots, I_{28} 是 28 个申万一级行业因子的因子收益， S_1, \dots, S_{11} 是表 1 中的 11 个大类风格因子，共计 40 个因子。行业因子暴露度非 0 即 1，风格因子暴露度为上述因子值。限制条件中 w_{I_n} 表示属于该行业的所有股票的流通市值之和。

在此种限制条件下，流通市值加权的组合（基准组合）的收益就是截距项 f_c ，更详细的描述见系列报告一。

加入截距项的好处是让求解出的纯因子组合的权重之和一定为 0（现金中性），因为任一因子的纯因子组合对截距项因子的暴露为： $w_{p_m}^T x_1 = w_{p_m}^T \vec{1} = \text{sum}(w_{p_m}) = 0$ 。加入 beta 因子的好处是可以构建 beta 中性的纯因子组合。

三、纯因子组合的历史收益

在历史样本中回测观察各个因子对应的纯因子组合收益，解决了通常的排序多/空组合同时对其它因子暴露的弊端，可以更具说服力的检验一个因子是否有效。下面分别考察截距项因子、纯风格因子、纯行业因子和残差项的历史收益情况。

3.1 回测时间和样本空间

回测时间为 2005 年 1 月至 2016 年 4 月共计 136 个月份，每月计算各因子对应的纯因子组合，并计算它们的收益。由于只是检测因子的有效性，而非真实可投资的组合（多空组合无法真实投资），故未考虑任何成本。样本空间为回测时间点非 ST、非停牌的全部 A 股。

图 1：样本空间股票数量



资料来源：天软科技，长江证券研究所

表 2：11 种风格因子暴露度的相关系数（2005.1 – 2016.4 月平均）

	BP	EP	Beta	成长	流动性	动量	非线性规模	质量	非线性规模	反转	规模
BP	1.00	0.31	0.05	-0.08	-0.23	-0.40	0.08	-0.08	-0.47	-0.16	0.07
EP	0.31	1.00	-0.13	0.25	-0.31	-0.03	0.08	0.50	-0.41	-0.05	0.38
Beta	0.05	-0.13	1.00	-0.05	0.37	-0.09	0.05	-0.10	0.22	-0.08	-0.18
成长	-0.08	0.25	-0.05	1.00	-0.08	0.17	0.08	0.31	-0.06	0.05	0.12
流动性	-0.23	-0.31	0.37	-0.08	1.00	0.11	-0.21	-0.27	0.57	0.08	-0.54
动量	-0.40	-0.03	-0.09	0.17	0.11	1.00	0.07	0.12	0.35	0.16	0.16
非线性规模	0.08	0.08	0.05	0.08	-0.21	0.07	1.00	0.10	-0.10	0.03	0.34
质量	-0.08	0.50	-0.10	0.31	-0.27	0.12	0.10	1.00	-0.18	0.02	0.41
残差波动率	-0.47	-0.41	0.22	-0.06	0.57	0.35	-0.10	-0.18	1.00	0.11	-0.31
反转	-0.16	-0.05	-0.08	0.05	0.08	0.16	0.03	0.02	0.11	1.00	0.06
规模	0.07	0.38	-0.18	0.12	-0.54	0.16	0.34	0.41	-0.31	0.06	1.00

注：红、绿色字体为每列最大值和最小值

资料来源：天软科技，wind，长江证券研究所

检验共线性的方法除了观察相关系数外，还可以计算 VIF 值和方差分解比例，方法

见系列报告一。

3.2 截距项因子收益

截距项因子收益 f_c 对应的组合是基准组合，基准组合对各风格因子的暴露均为 0，对自身暴露为 1（全额投资）。如果风险因子解释力充分，因子收益 f_c 的累计收益应该和市值加权指数的收益十分接近，与中证全指对比如图 2。

图 2：截距项因子收益和中证全指对比



资料来源：天软科技，长江证券研究所

3.3 纯风格因子组合的收益

风格因子对应的纯因子组合，对截距项暴露为 0（零额投资或叫现金中性），即组合为多空组合，且多头/空头为 100/100，对各行业因子暴露为 0（在各行业中也是零额投资），对其他风格因子暴露为 0，对自身风格因子偏离基准组合 1 个标准差。此处展示成长、EP、BP、流动性、规模、非线性规模、beta、残差波动率、反转、动量、质量等 11 个风格因子的纯因子收益。

通过 t 值来检验观察到的各纯因子收益率是否显著不为零，并计算历史上发生 $|t| > 2$ 的次数所占比例来考察因子对股价产生影响的稳定性，这种影响若是完全随机的，那么占比应该小于 5%。通过 $t > 2$ 和 $t < -2$ 分别占比，来体现因子对股价在同一方向上的影响的持续性。

选取的各风格纯因子组合均能超越基准组合，流动性、规模纯因子、反转因子组合超额收益明显高于其他因子。波动率方面，规模因子最大，成长因子最小。综合起来，流动性因子的信息比率是最高的。在因子同方向持续性方面，流动性、规模、反转出现 $t < -2$ 的概率和 BP、成长出现 $t > 2$ 的概率都较高，持续性也十分优秀。具体数据见表 3 和图 3 到图 14。

表 3: 各风格纯因子收益情况汇总

	年化 收益	年化 波动率	收益/风险	mean(t)	t >2 占比	(t>2占比/ t<-2占比) ^d	因 子 方 向
流动性	9.25%	4.12%	2.24	2.82	53.7%	9.43	-1
反转	6.46%	3.93%	1.64	3.54	58.8%	3.21	-1
规模	12.16%	8.13%	1.50	4.97	77.2%	4.00	-1
成长	1.59%	1.51%	1.06	1.30	24.3%	5.60	+1
非线性规模	3.41%	4.13%	0.82	2.63	52.9%	2.00	-1
BP	3.12%	3.99%	0.78	2.53	53.7%	1.70	+1
Beta	3.13%	4.15%	0.75	3.22	53.7%	1.43	+1
动量	2.19%	3.46%	0.63	2.57	57.4%	1.60	+1
质量	1.20%	2.16%	0.55	1.53	25.0%	1.43	+1
EP	1.24%	2.56%	0.48	1.74	36.8%	1.78	+1
残差波动率	0.99%	3.09%	0.32	1.99	44.9%	1.26	-1

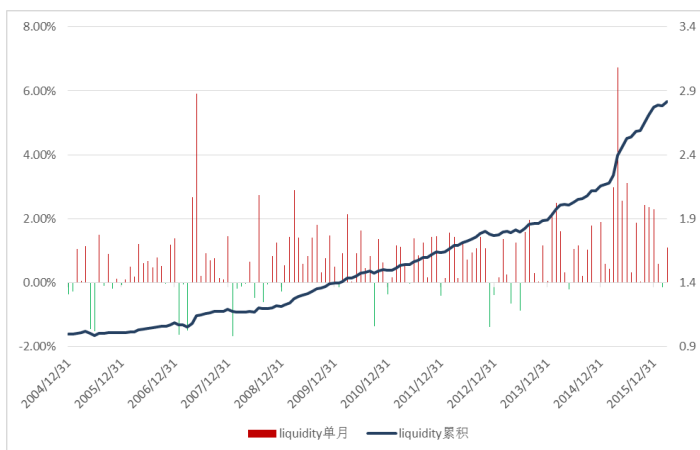
 注 1: (t>2 占比/ t<-2 占比)^d 中 d 为因子方向, 取 1 或 -1,

注 2: 年化收益均已剔除设定为 3% 的无风险收益率

资料来源: 天软科技, wind, 长江证券研究所

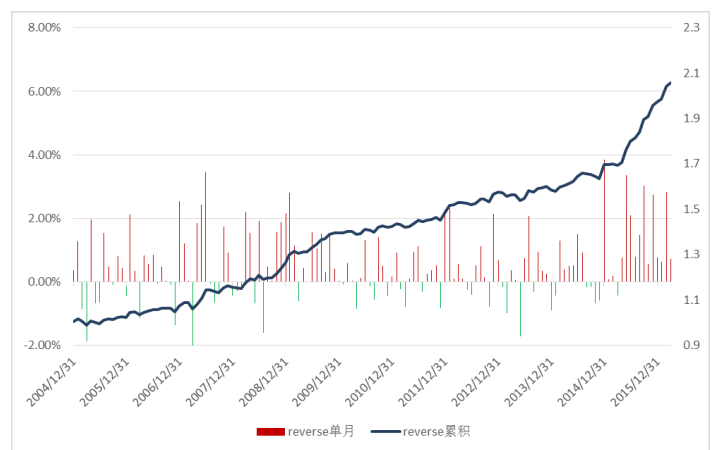
各风格因子收益曲线(图 3 -- 14):

图 3: 纯流动性因子组合收益



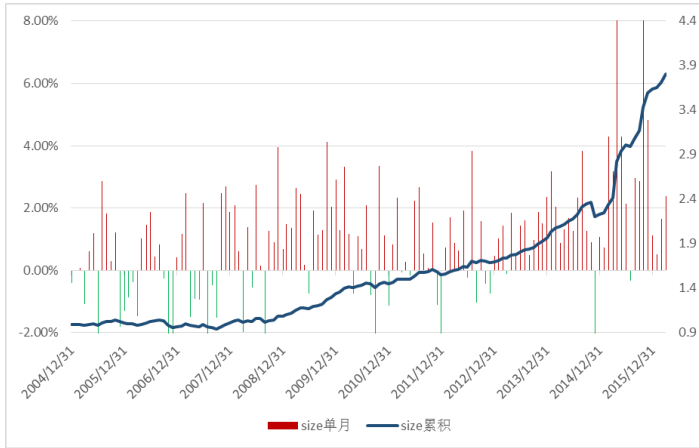
资料来源: 天软科技, 长江证券研究所

图 4: 纯反转因子组合收益



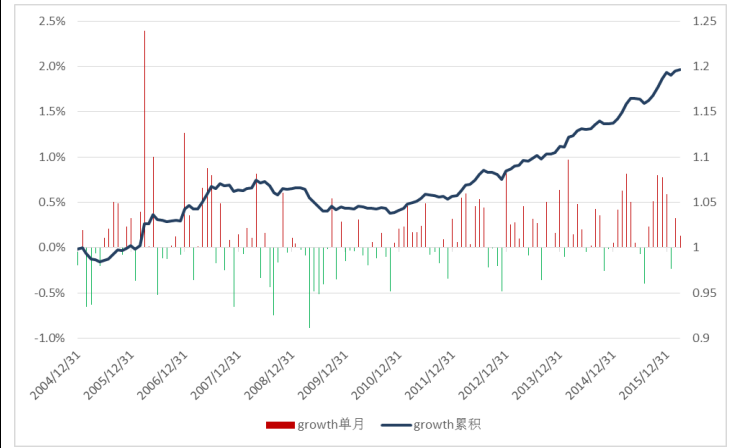
资料来源: 天软科技, 长江证券研究所

图 5: 纯规模因子组合收益



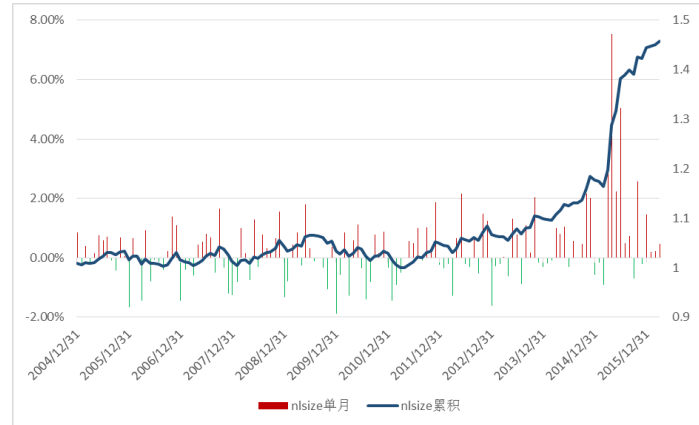
资料来源：天软科技，长江证券研究所

图 6: 纯成长因子组合收益



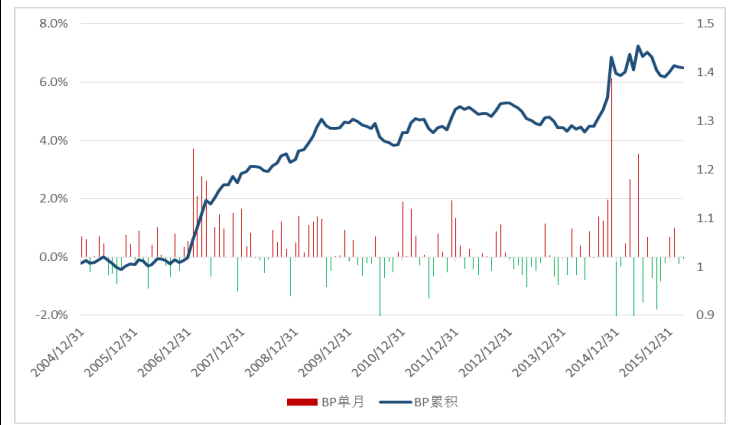
资料来源：天软科技，wind，长江证券研究所

图 7: 纯非线性规模因子组合收益



资料来源：天软科技，长江证券研究所

图 8: 纯 BP 因子组合收益



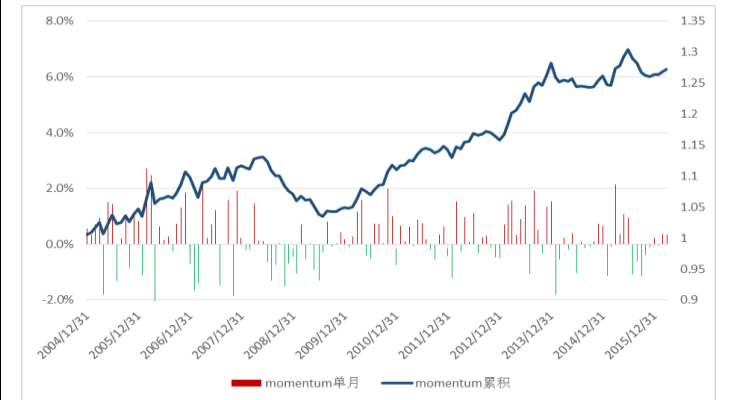
资料来源：天软科技，长江证券研究所

图 9: 纯 beta 因子组合收益



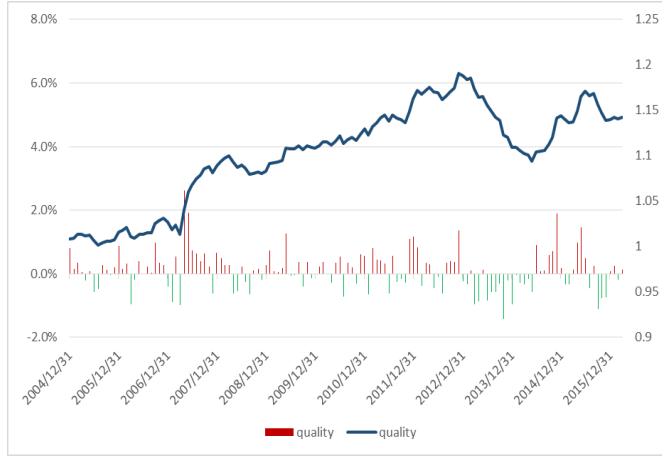
资料来源：天软科技，长江证券研究所

图 10: 纯动量因子组合收益



资料来源：天软科技，长江证券研究所

图 11: 纯质量因子组合收益



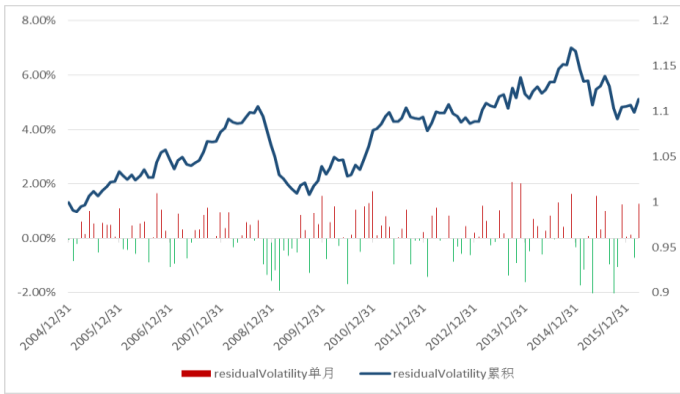
资料来源：天软科技，长江证券研究所

图 12: 纯 EP 因子组合收益



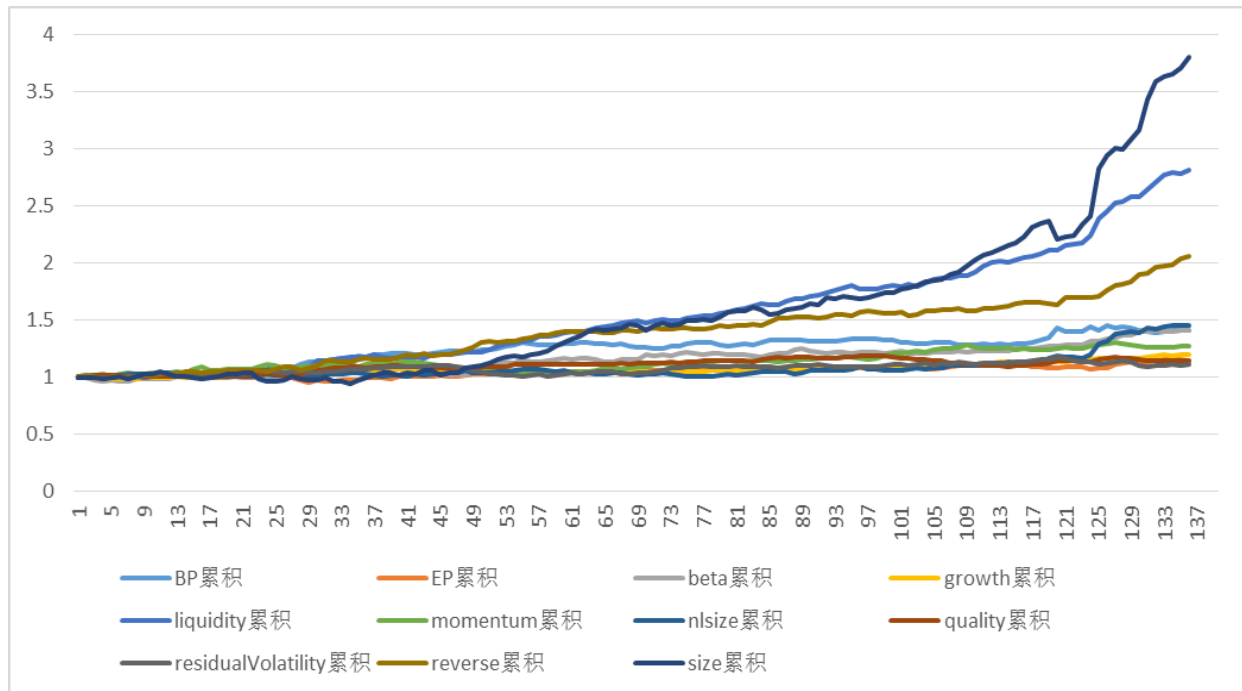
资料来源：天软科技，wind，长江证券研究所

图 13: 纯残差波动率因子组合收益



资料来源：天软科技，长江证券研究所

图 14：各纯因子组合收益汇总



资料来源：天软科技, wind, 长江证券研究所

3.4 纯行业因子组合的收益

行业因子收益 f_{I_n} 对应的纯因子组合对本行业暴露为 1（即权重 100%），对截距项暴露为 0（零额投资, 100%自身行业多头，同时 100%基准组合空头，即行业因子收益本身就是超额收益），对各个风格因子暴露为 0。考察 28 个申万一级行业的因子收益，可以了解在控制风格暴露后，各行业的收益情况。

表 4：各行业纯因子收益情况汇总

	年化超额	年化波动率	信息比率	mean(t)	t >2占比	t>2占比	t<-2占比
申万医药生物	9.54%	12.52%	0.76	2.99	53.68%	32.4%	21.3%
申万有色金属	11.98%	16.61%	0.72	2.83	51.47%	30.1%	21.3%
申万食品饮料	7.77%	12.27%	0.63	2.11	39.71%	22.8%	16.9%
申万计算机	9.20%	14.93%	0.62	2.56	47.79%	28.7%	19.1%
申万国防军工	8.44%	15.61%	0.54	2.29	45.59%	27.9%	17.6%
申万传媒	7.31%	13.86%	0.53	2.31	45.59%	27.2%	18.4%
申万通信	6.95%	13.55%	0.51	2.19	44.85%	28.7%	16.2%
申万非银金融	9.07%	19.45%	0.47	3.52	58.82%	32.4%	26.5%
申万电气设备	4.25%	9.29%	0.46	2.16	45.59%	24.3%	21.3%
申万农林牧渔	6.06%	13.81%	0.44	1.96	38.24%	19.9%	18.4%
申万机械设备	3.92%	9.10%	0.43	2.20	48.53%	25.0%	23.5%
申万汽车	3.56%	10.59%	0.34	2.14	41.91%	25.7%	16.2%
申万银行	4.46%	15.83%	0.28	3.20	58.82%	31.6%	27.2%
申万综合	1.83%	10.58%	0.17	1.15	18.38%	11.8%	6.6%
申万建筑装饰	2.04%	12.04%	0.17	1.95	36.03%	16.2%	19.9%
申万房地产	2.72%	16.39%	0.17	3.37	62.50%	33.8%	28.7%
申万家用电器	0.47%	11.80%	0.04	1.50	29.41%	16.9%	12.5%
申万商业贸易	0.36%	10.28%	0.03	1.85	34.56%	14.7%	19.9%
基准组合	0.00%	33.60%	0.00	--	--	--	--
申万电子	-0.45%	12.13%	-0.04	2.15	44.12%	22.8%	21.3%
申万休闲服务	-1.36%	13.21%	-0.10	1.24	22.06%	9.6%	12.5%
申万采掘	-2.04%	19.23%	-0.11	2.96	57.35%	22.1%	35.3%
申万建筑材料	-1.75%	10.73%	-0.16	1.85	35.29%	14.7%	20.6%
申万纺织服装	-2.22%	10.01%	-0.22	1.36	27.21%	10.3%	16.9%
申万化工	-2.58%	7.86%	-0.33	2.07	44.12%	16.9%	27.2%
申万轻工制造	-4.89%	9.30%	-0.53	1.23	23.53%	8.8%	14.7%
申万公用事业	-7.36%	10.70%	-0.69	2.15	44.85%	14.7%	30.1%
申万交通运输	-8.29%	10.22%	-0.81	1.99	35.29%	8.8%	26.5%
申万钢铁	-12.57%	12.41%	-1.01	1.95	41.91%	10.3%	31.6%

资料来源：天软科技,长江证券研究所

总体来说，在控制了风格因子之后，医药、有色、食品饮料和计算机贡献较高超额收益，而钢铁、交运、公用事业和轻工没有超额收益。从波动率可以看出，行业因子收益的轮动更加频繁和难以把握（可能是因为行业属性比风格因子属性更显而易见，price-in 越快）。各纯行业因子的收益曲线见附录 I。

3.5 残差项的收益和模型的解释力

a. 检验各纯因子组合中不能被模型解释的收益是否相互抵消。对各纯因子组合，计算 $w_f^T u$ ，各月平均值见表 5。

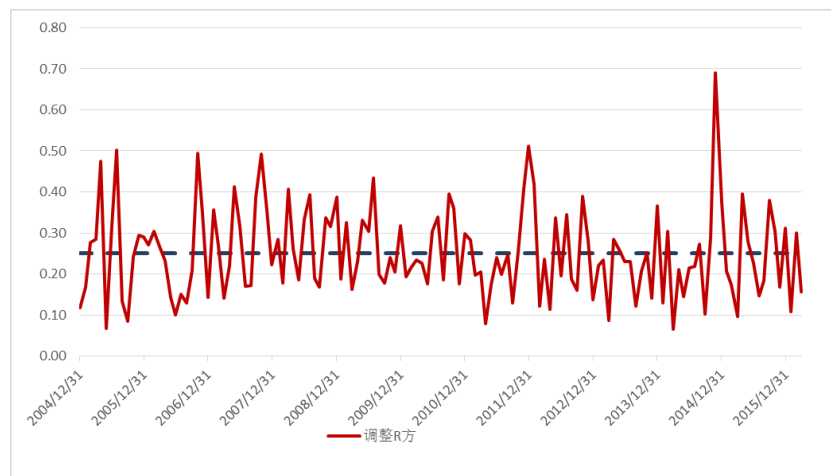
表 5: 各风格纯因子的残差相抵程度 (月度平均值)

BP	EP	beta	成长	流动性	动量
-2.12E-18	5.01E-18	2.67E-17	1.17E-17	1.03E-17	-3.00E-18
非线性规模	质量	非线性规模	反转	规模	
-2.50E-17	3.67E-18	-1.80E-17	-2.62E-18	1.18E-17	

资料来源: 天软科技, 长江证券研究所

b. 模型解释力

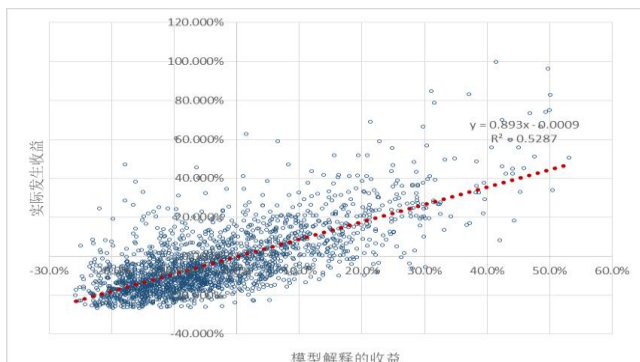
图 15: 模型解释力 -- 调整 R 方



资料来源: 天软科技, 长江证券研究所

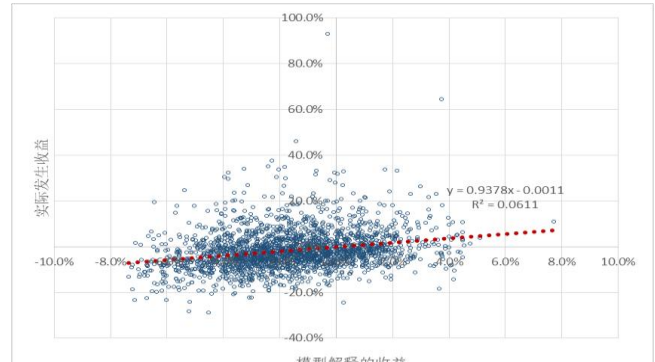
模型的调整 R 方的均值为 0.25 ± 0.11 ，也就是说大概股票收益率 r 和模型对收益的解释结果 $x\beta$ 之间的相关系数大约为 0.5，其余部分均为模型无法解释的部分。以调整 R 方最大的 2014 年 12 月 ($R^2\text{-adj} = 0.69$) 和最小的 2014 年 4 月 ($R^2\text{-adj} = 0.065$) 来具体展示。

图 16: 2014 年 12 月, $R^2\text{-adj} = 0.69$



资料来源: 天软科技, 长江证券研究所

图 17: 2014 年 4 月, $R^2\text{-adj} = 0.065$



资料来源: 天软科技, 长江证券研究所

表 5 和图 16、图 17 说明，虽然模型的解释力有限（被解释的收益和实际收益在相

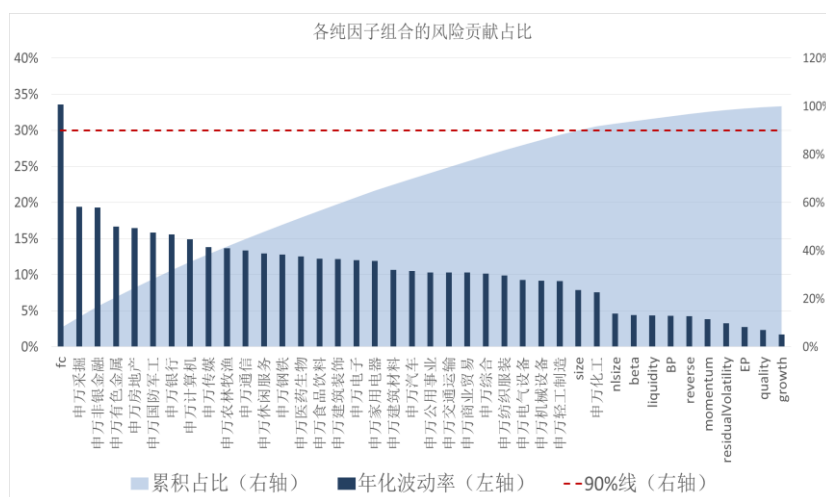
关系数在 0.25 到 0.73 之间), 但只要组合充分分散, 不能被解释的部分能够相互抵消, 不用做特别处理。但若是集中持股, 多因子模型的特异性收益率带来的风险将上升。

四、纯因子组合风险特征

4.1 纯因子组合的历史波动率

计算样本区间内各因子的年化波动率（图 18），可见，截距项因子的波动率（即市场波动率）最大，成长因子的波动率最小，除规模因子外，纯风格因子组合的波动率普遍小于纯行业因子组合的波动率。市场因子、纯行业因子和规模因子贡献了大约 90% 的波动率，一个市场中性、行业中性的组合可以规避掉大部分风险。

图 18: 纯因子组合历史波动率



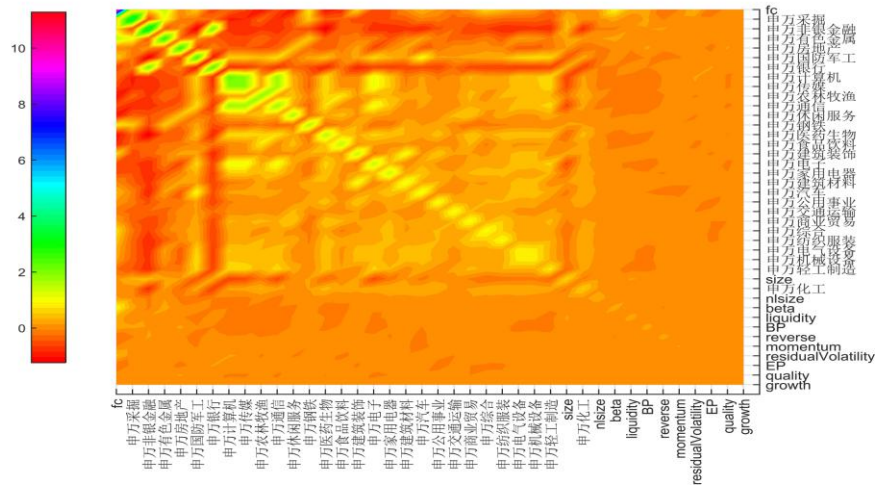
资料来源：天软科技，长江证券研究所

4.2 纯因子组合的协方差

当考虑运用均值-方差优化，将多种纯因子组合合成可投资的组合时，不仅要估计各组合各自的方差（4.1 节），还需要考察各纯因子组合间的协方差。

由于我们展示的这个模型包含有 40 个因子，即含有 $40 \times 40 = 1600$ 个协方差，数量较多，为了展示的便利，我们采用热力图和 3D 热力图来观察协方差矩阵的特征，如图 19 和图 20。

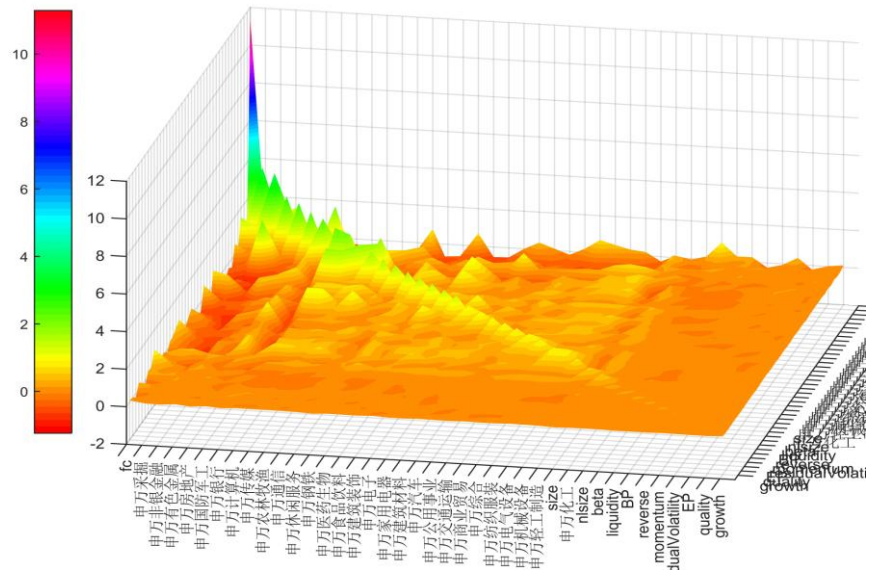
图 19: 纯因子组合协方差矩阵热力图



注：颜色栏标注数值为月协方差*12*100

资料来源：天软科技，长江证券研究所

图 20: 纯因子组合协方差矩阵热力图 3D



资料来源：天软科技，长江证券研究所

从以上两图可以看出，方差和协方差的来源也主要来自市场和行业因子，风格因子收益之间的协方差相对平坦。另外，有趣的是，纯银行因子组合和纯非银因子组合的收益，与大部分其他纯因子组合的收益的协方差都小于零。

五、检验一些常见因子

以上罗列的风格因子至少大约可以解释掉 50% (模型 $R^2 \approx 0.3$) 的收益来源, 如果另外的某个因子和这些因子相关度较高, 那么简单的排序法会高估因子本身带来的收益; 另一方面, 如果因子的方向和这些因子相反(正序或逆序), 或者这个因子和这些因子之间有复杂的相关性而变为噪音, 那有可能该因子的收益会被原来的因子掩盖掉, 排序法会低估因子带来的收益。

下面检验 18 种常见的因子, 并和排序分组法的前后 10% 组的多空组合做对比。

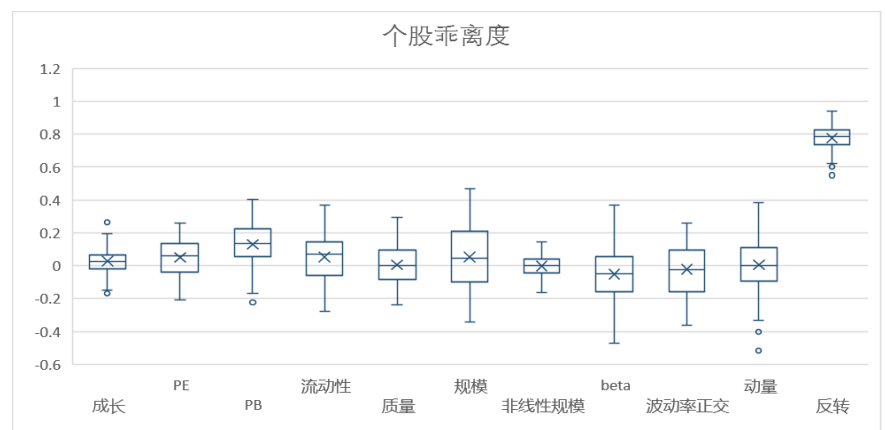
5.1 被排序法高估的因子

个股乖离度

个股乖离度与反转因子相关度高, 在纯因子组合中剔除了反转因子的效用后, 因子收益率有明显降低。

a. 与各风险因子相关系数

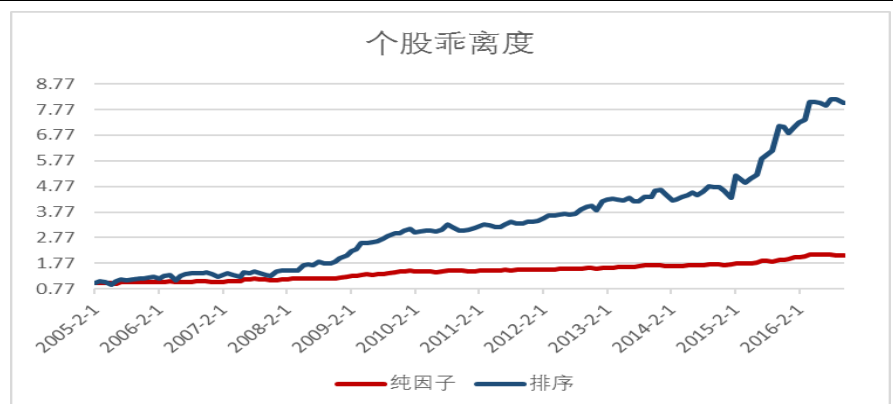
图 21: 个股乖离度与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 22: 个股乖离度因子收益率被高估



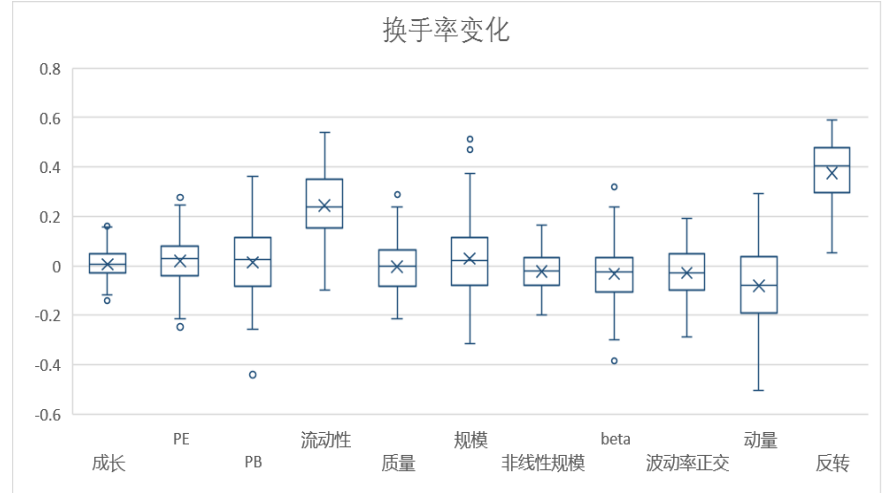
资料来源：天软科技，长江证券研究所

换手率变化

换手率变化是近 20 日换手率与近 60 日换手率之比值。该因子相关度与反转因子相关性较高，并和流动性因子有些许相关，在纯因子组合中，因子收益率有明显降低。

a. 与各风险因子相关系数

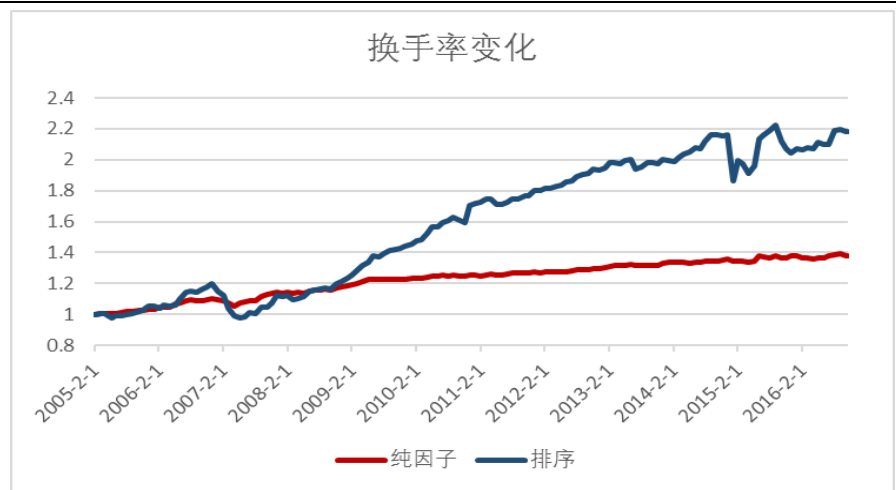
图 23: 换手率变化与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 24: 换手率变化因子收益率被高估



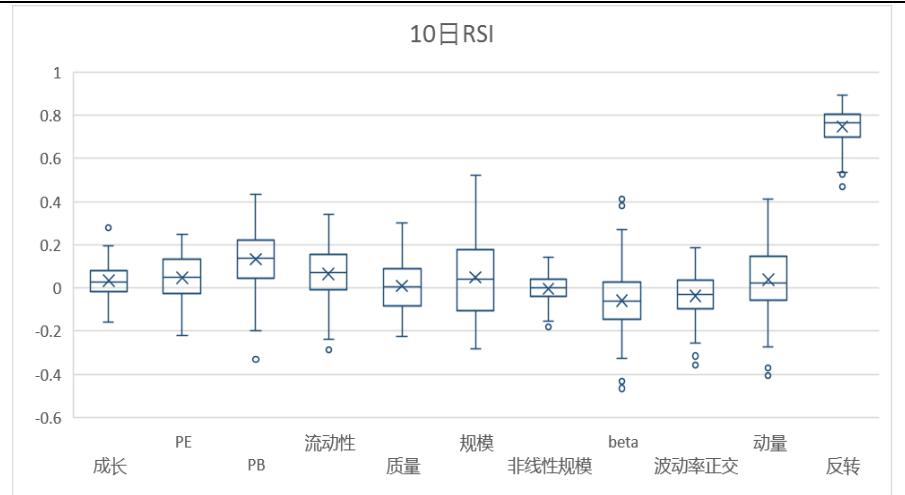
资料来源：天软科技，长江证券研究所

10 日 RSI

该因子相关度与反转因子相关性较高，在纯因子组合中，因子收益率有明显降低。

a. 与各风险因子相关系数

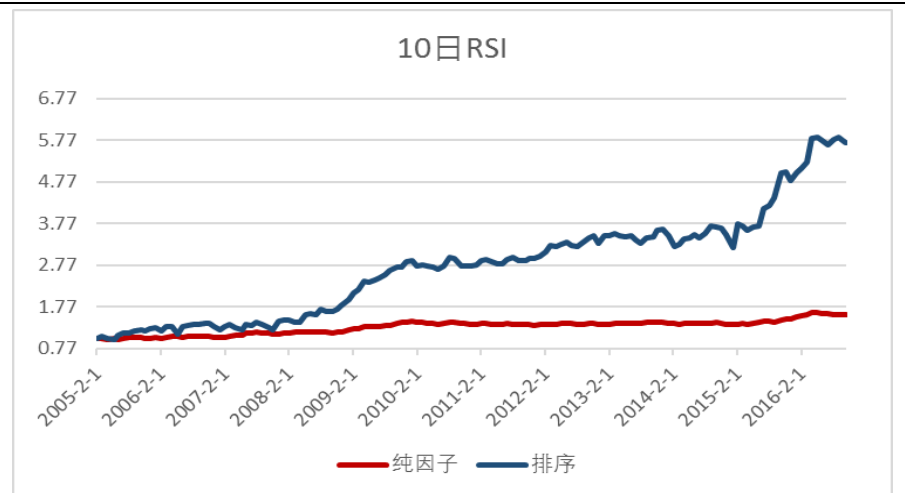
图 25: 10 日 RSI 与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 26: 10 日 RSI 因子收益率被高估



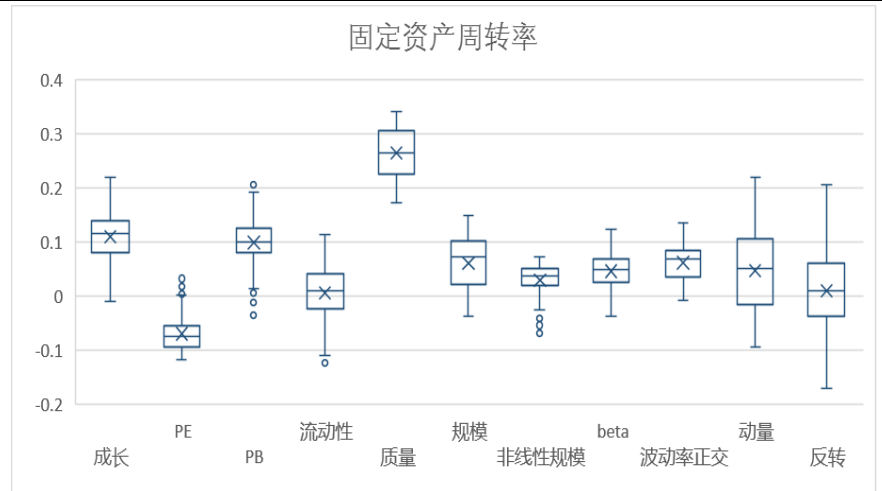
资料来源：天软科技，长江证券研究所

固定资产周转率

该因子与质量因子相关性较高，在纯因子组合中，因子收益率有明显降低。

a. 与各风险因子相关系数

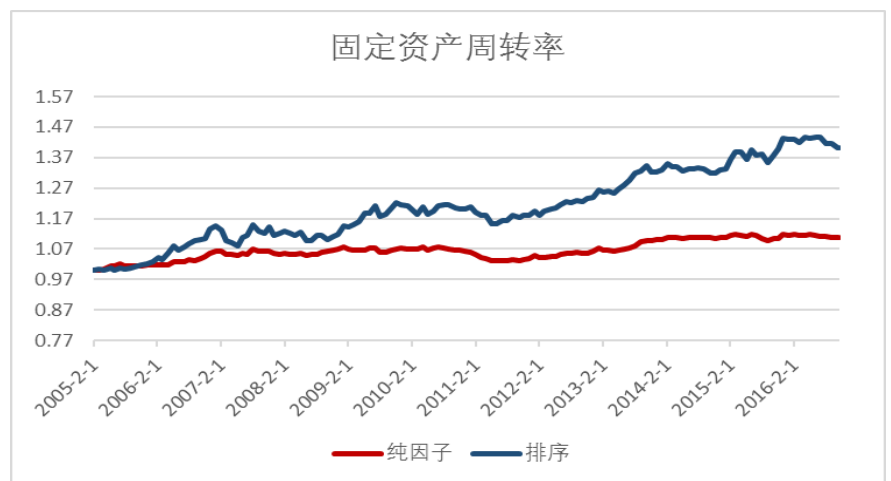
图 27: 固定资产周转率与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 28: 固定资产周转率因子收益率被高估



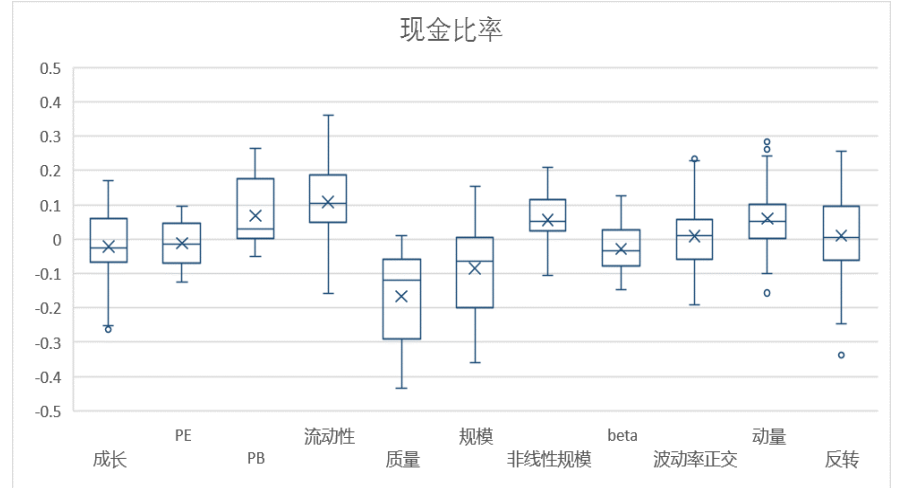
资料来源：天软科技，长江证券研究所

现金比率

该因子相关度与质量、规模因子相关性较高，在纯因子组合中，因子收益率有明显降低。

a. 与各风险因子相关系数

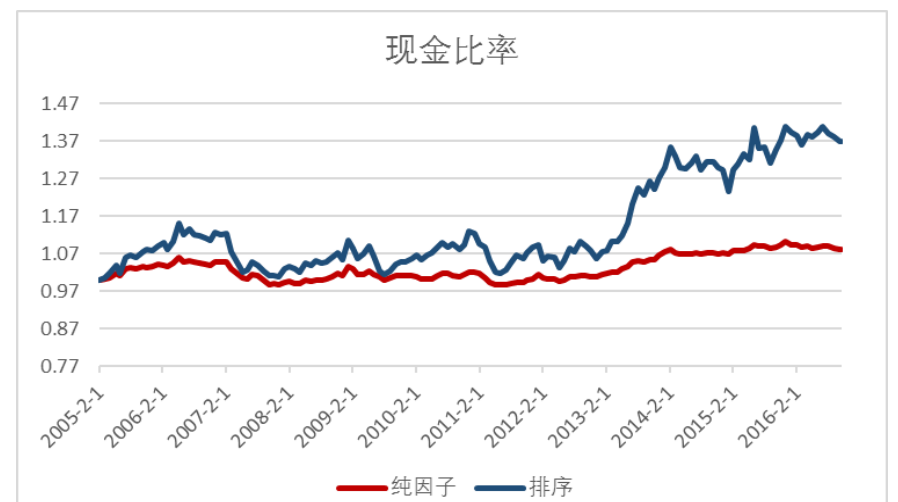
图 29: 现金比率与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 30: 现金比率因子收益率被高估



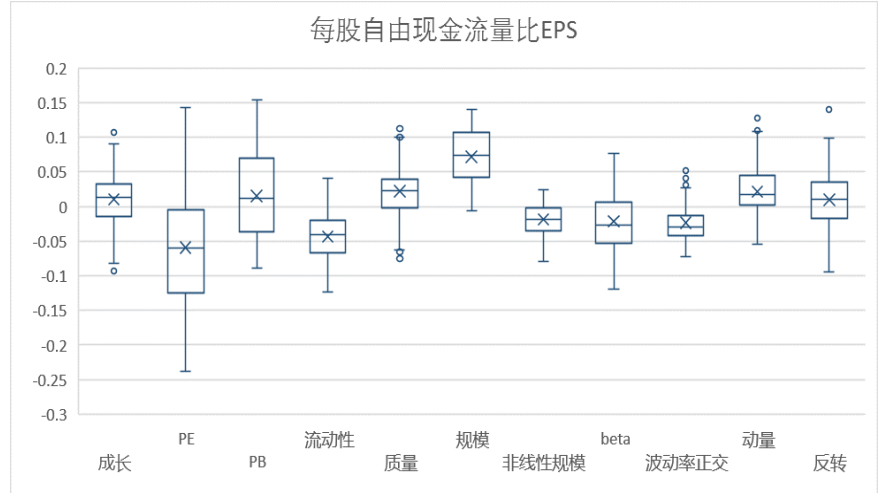
资料来源：天软科技，长江证券研究所

每股自由现金流量比 eps

该因子与各常见风格因子均不高，在控制了风格因子后，因子收益减少，但因子收益与原排序组合的相关性较高，可以体现出这部分是独立的因子收益。

a. 与各风险因子相关系数

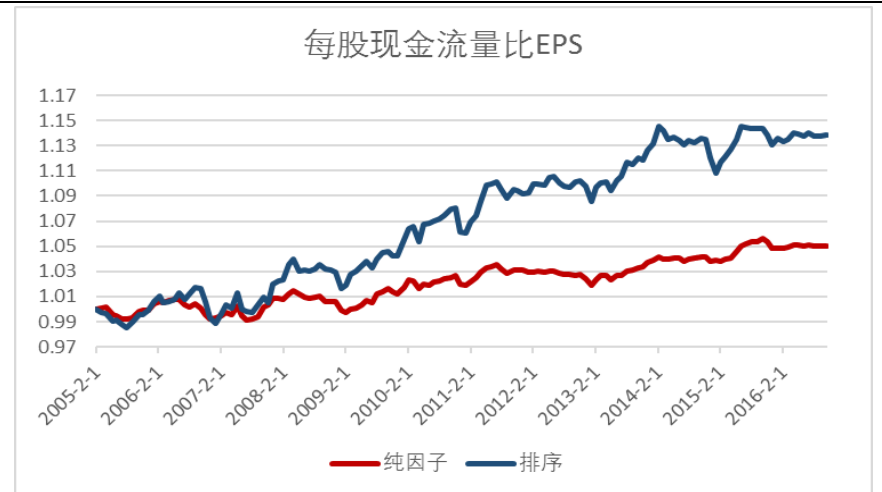
图 31：每股自由现金流量比 eps 与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 32：每股自由现金流量比 eps 因子收益率被高估



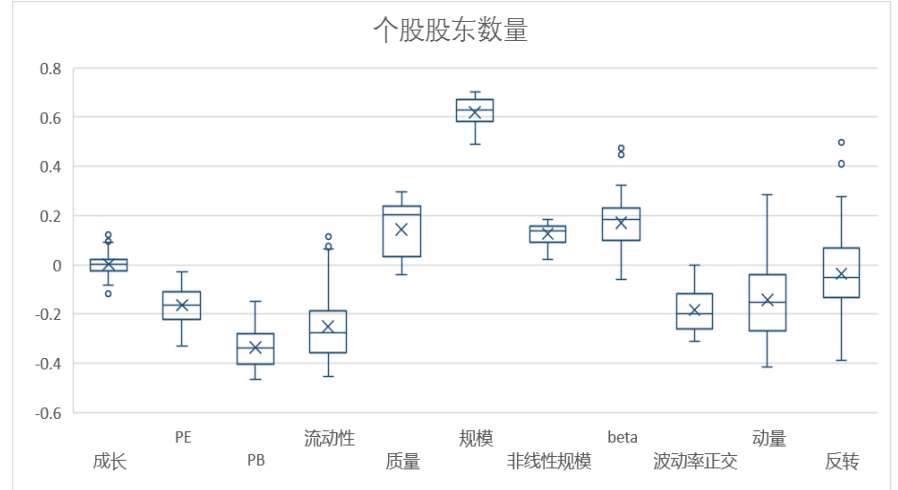
资料来源：天软科技，长江证券研究所

个股股东数

该因子有严重规模因子暴露，并有 PB、流动性等复杂关系的干扰，剔除风格因子贡献后的纯因子组合中，超额收益大幅减少。

a. 与各风险因子相关系数

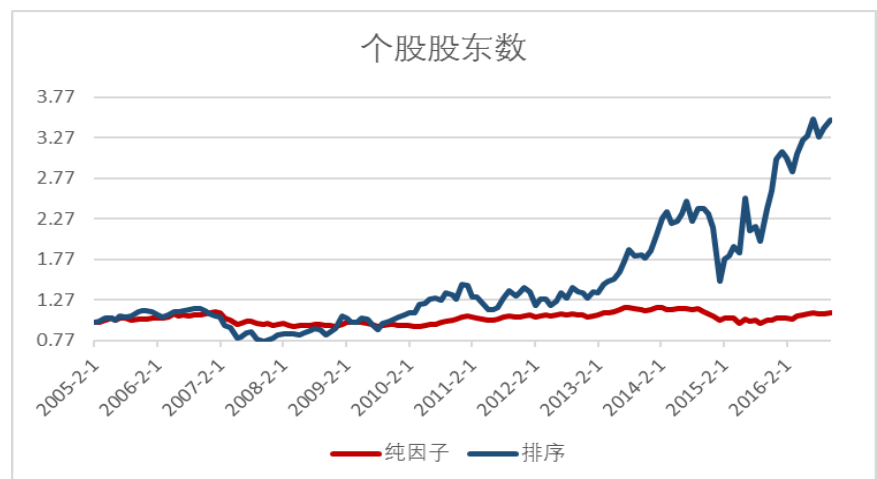
图 33: 个股股东数与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 34: 个股股东数因子收益率被高估



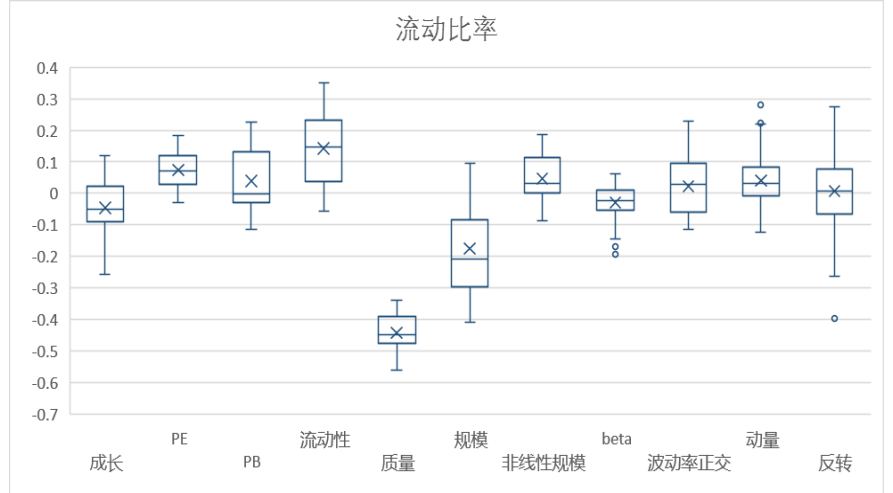
资料来源：天软科技，长江证券研究所

流动比率

该因子严重暴露到质量因子上，并有一定规模因子暴露，在剔除风格因子贡献后的纯因子组合中，几乎没有超额收益。

a. 与各风险因子相关系数

图 35: 流动比率与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 36: 流动比率因子收益率被高估



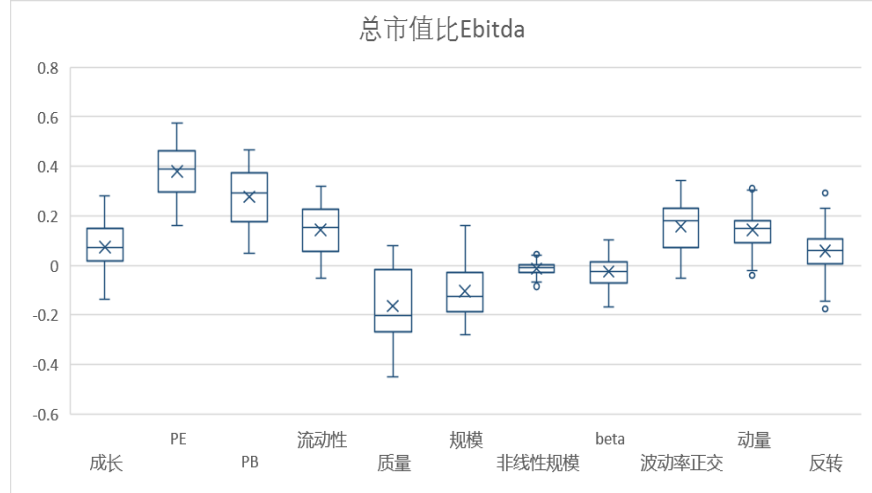
资料来源：天软科技，长江证券研究所

总市值比 Ebitda

该因子暴露到 PB、PE 因子上，在剔除风格因子贡献后的纯因子组合中，同样几乎没有超额收益。

a. 与各风险因子相关系数

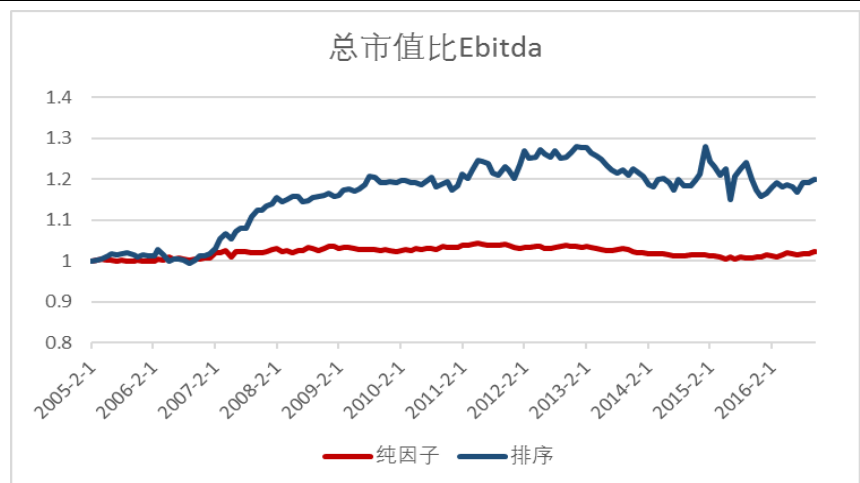
图 37: 总市值比 Ebitda 与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 38: 总市值比 Ebitda 因子收益率被高估



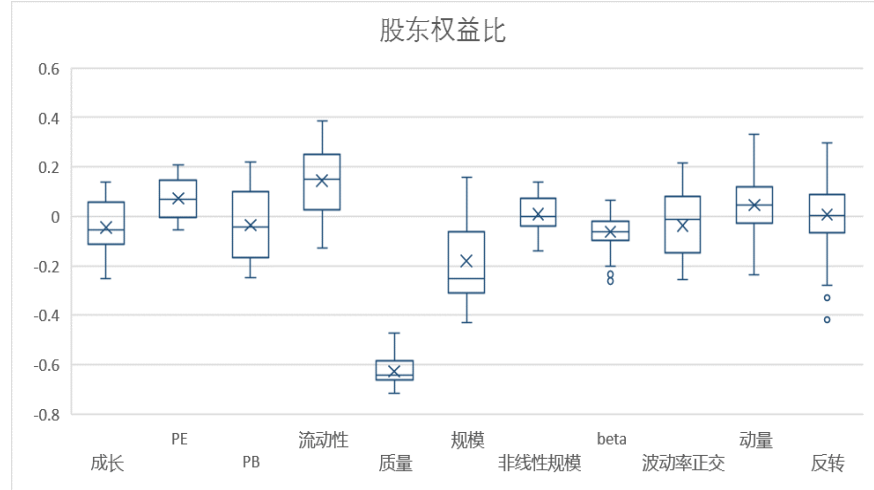
资料来源：天软科技，长江证券研究所

股东权益比

该因子很大程度暴露到质量上,并有一定规模因子暴露,在剔除风格因子贡献后的纯因子组合中,甚至体现出负的超额收益。

a. 与各风险因子相关系数

图 39: 股东权益比与各风险因子相关系数



资料来源: 天软科技, 长江证券研究所

b. 因子收益

图 40: 股东权益比因子收益率被高估



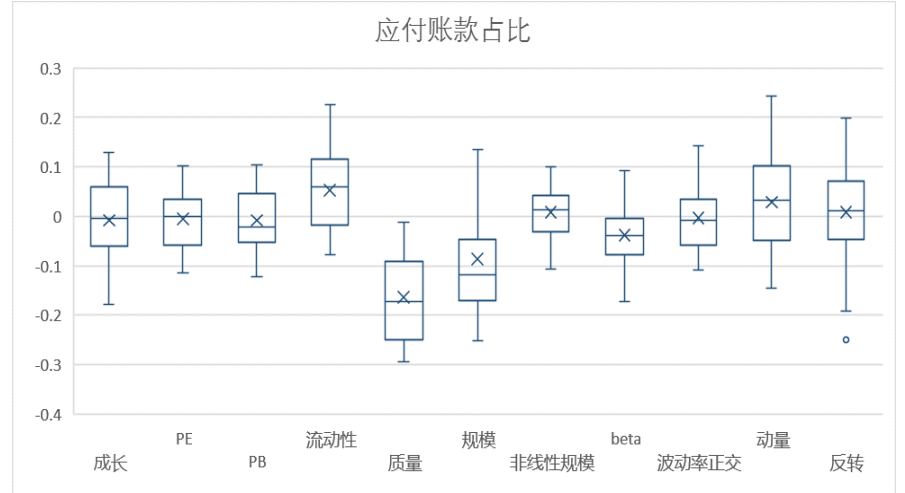
资料来源: 天软科技, 长江证券研究所

应付账款占比

该因子有些许暴露到质量和规模因子上，在剔除风格因子贡献后的纯因子组合中，同样体现出负的超额收益。

a. 与各风险因子相关系数

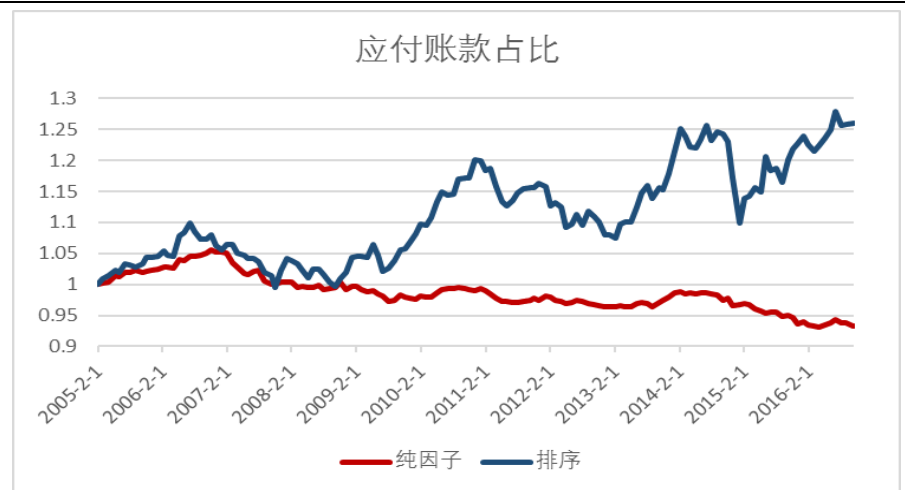
图 41: 应付账款占比与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 42: 应付账款占比因子收益率被高估



资料来源：天软科技，长江证券研究所

表 6 为以上各因子的纯因子组合和排序多空组合的 IR 对比。虽然排序法高估了个股乖离度因子和换手率变化因子的收益，但通过构建纯因子组合，它们的 IR 有所提升，仍不失为是优秀的因子。10 日 RSI 和固定资产周转率的 IR 虽然降低，但仍然可以接受。而应收账款占比和股东权益比等因子在剔除了常见风格因子的贡献后已经完全失效。

表 6：排序法高估的一些因子的 IR

	纯因子法IR	排序法IR	二者百分差
个股乖离度	6.08	5.09	19%
换手率变化	5.61	3.92	43%
10日RSI	3.81	4.45	-14%
固定资产周转率	2.06	3.13	-34%
现金比率	1.36	1.77	-23%
每股现金流量比EPS	1.30	2.22	-41%
个股股东数	0.93	2.75	-66%
流动比率	0.54	1.99	-73%
总市值比Ebitda	0.18	1.44	-88%
股东权益比	-1.35	1.30	-203%
应付账款占比	-2.60	1.30	-300%

资料来源：天软科技,长江证券研究所

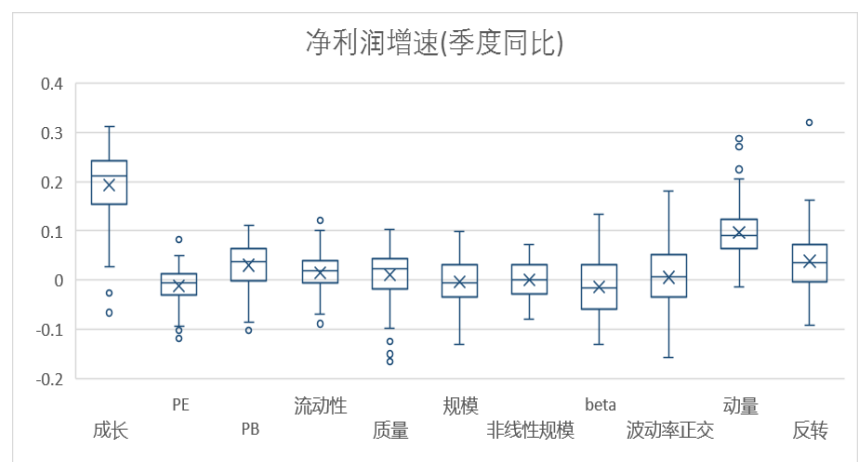
5.2 被排序法低估的因子

净利润增速（季度同比）

该因子与成长因子有些许相关性，与动量因子有细微的相关性，在控制了风格因子后，因子收益有少许增加，并且更加平稳，因子收益与原排序组合的相关性较高，可以体现出这部分是独立的因子收益。

a. 与各风险因子相关系数

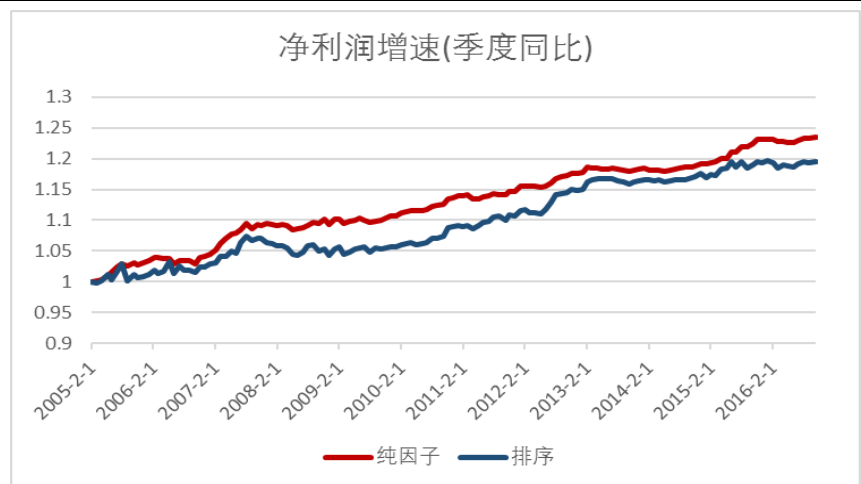
图 43：净利润增速（季度同比）与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 44：净利润增速（季度同比）因子收益率被低估



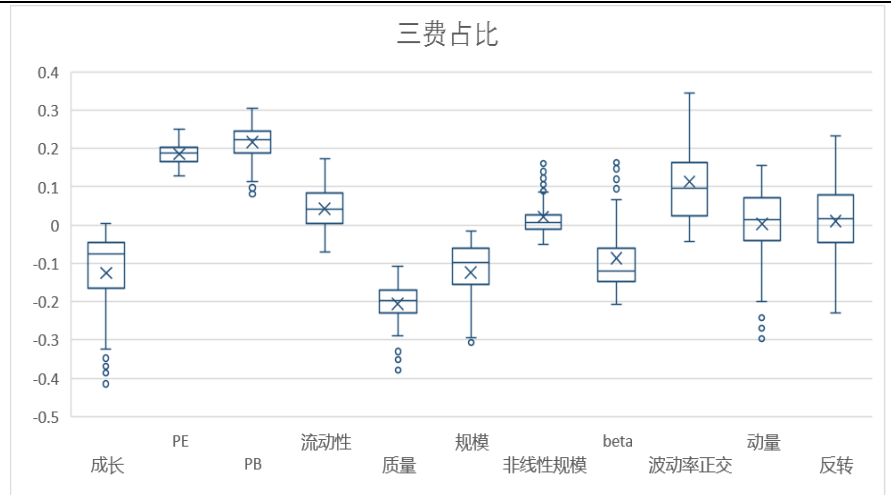
资料来源：天软科技，长江证券研究所

三费占比

该因子与 PB、PE 因子正相关，又与质量因子负相关；与规模因子负相关，又与成长因子负相关，多种因子对其产生方向不一的复杂影响。在剔除风格因子贡献后的纯因子组合中，体现出更加稳定的的因子收益。

a. 与各风险因子相关系数

图 45：三费占比与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 46：三费占比因子收益率被低估



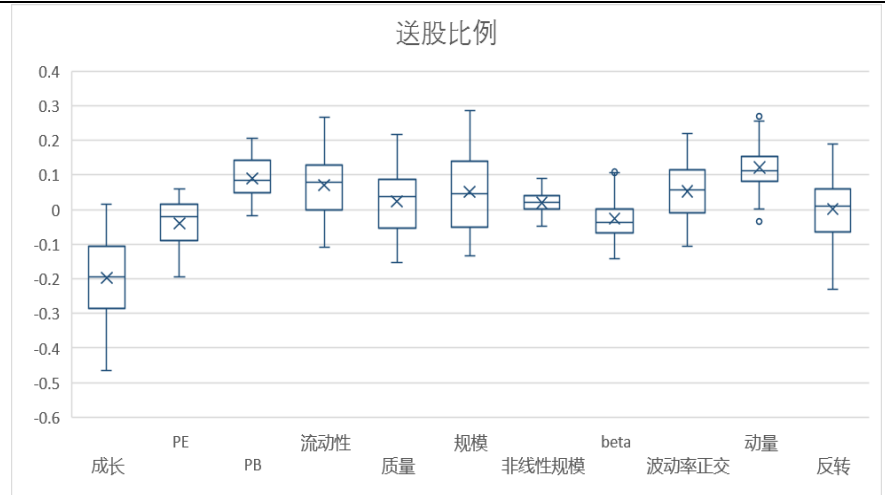
资料来源：天软科技，长江证券研究所

送股比例

送股比例暴度到了低成长的股票上面，着有可能其排序多空组合的收益被低成长的股票(通常表现不好)掩盖。在剔除风格因子贡献后的纯因子组合中，因子收益遂体现出来。

a. 与各风险因子相关系数

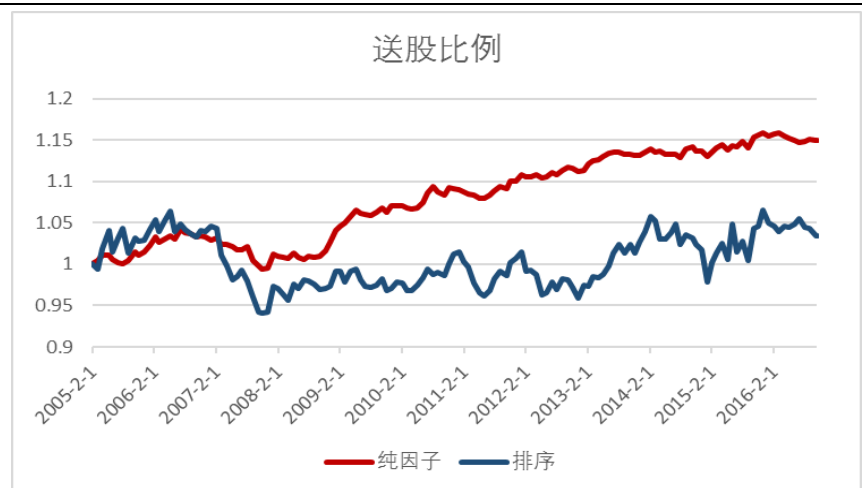
图 47：送股比例与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 48：送股比例因子收益率被低估



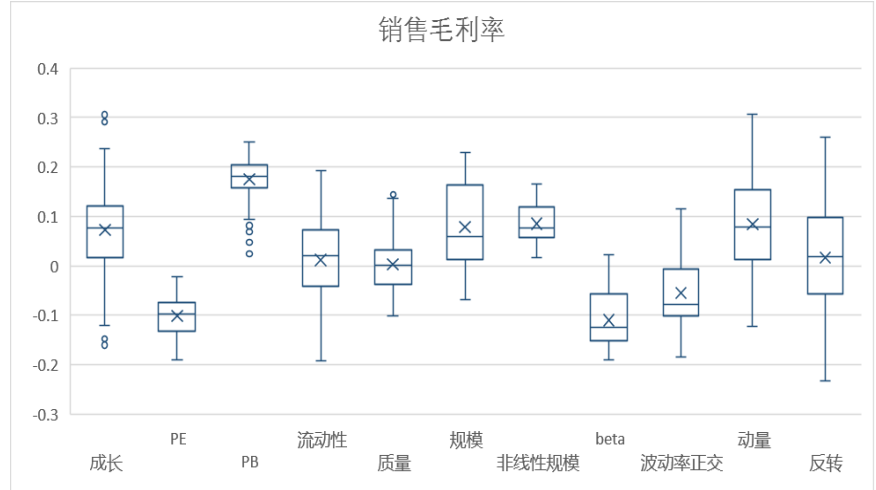
资料来源：天软科技，长江证券研究所

销售毛利率

该因子也有复杂的因子暴露：低 PE、高 PB、低 beta，在剔除了这些风格因子暴露后，因子收益明显稳定许多。

a. 与各风险因子相关系数

图 49：销售毛利率与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 50：销售毛利率因子收益率被低估



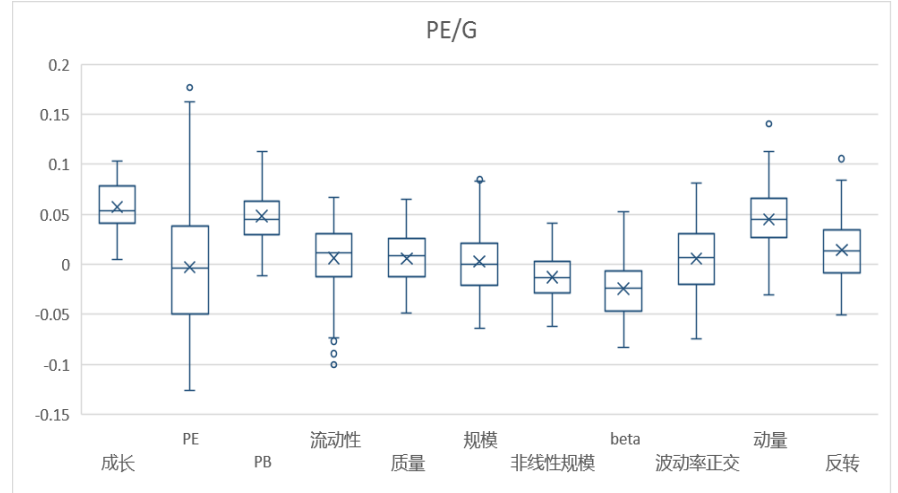
资料来源：天软科技，长江证券研究所

PE/G

该因子与风格因子相关性都较小，在剔除了风格因子暴露后，因子收益并没有太多改变，仅仅是稳定性增加，是不错的因子。

a. 与各风险因子相关系数

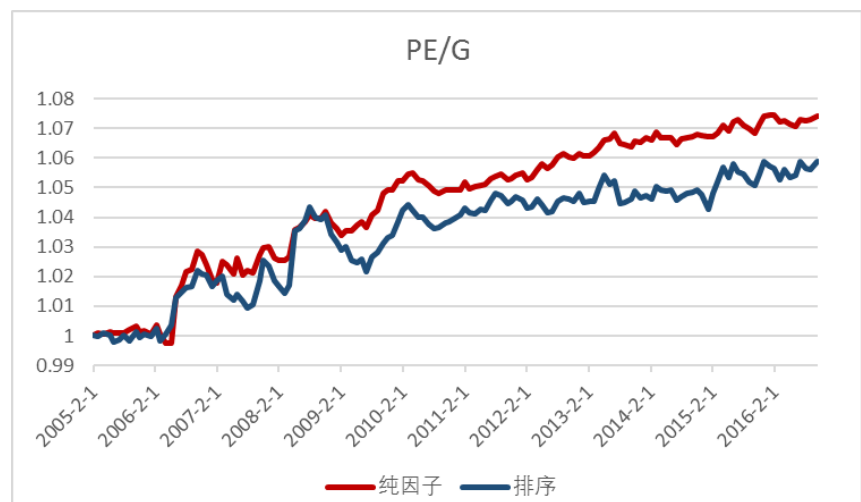
图 51: PE/G 与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 52: PE/G 因子收益率被低估



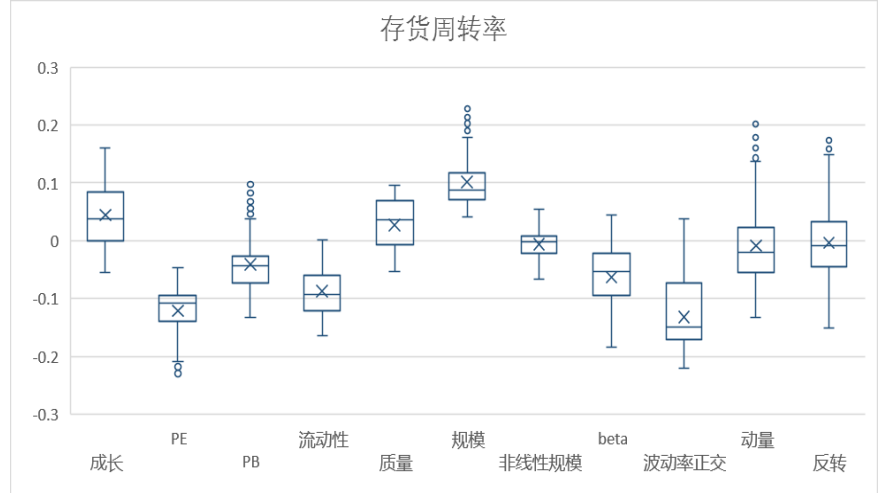
资料来源：天软科技，长江证券研究所

存货周转率

该因子主要偏向于小 PE、规模大的股票，在剔除了风格因子暴露后，因子收益并没有太多改变，仅仅是稳定性增加，是不错的因子。

a. 与各风险因子相关系数

图 53: 存货周转率与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 54: 存货周转率因子收益率被低估



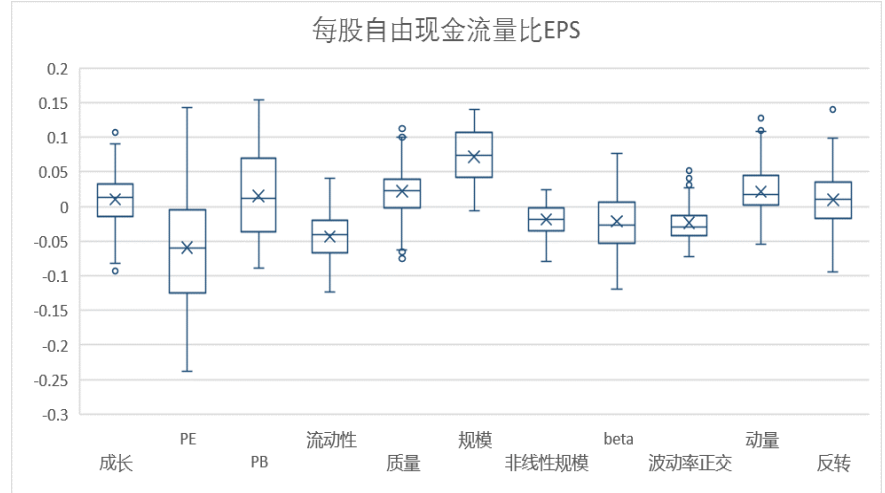
资料来源：天软科技，长江证券研究所

每股自由现金流量比 EPS

该因子的排序多空组合和纯因子组合反向，说明之前的一些因子暴露完全覆盖了因子本身的收益，在剔除了风格因子暴露后，该因子还稍微剩下一点超额收益。

a. 与各风险因子相关系数

图 55: 每股自由现金流量比 EPS 与各风险因子相关系数



资料来源：天软科技，长江证券研究所

b. 因子收益

图 56: 每股自由现金流量比 EPS 因子收益率被低估



资料来源：天软科技，长江证券研究所

表 7 为以上各因子的纯因子组合和排序多空组合的 IR 对比。净利润增速季度同比、三费占比和 PE/G 这样的因子本身已经很优秀的因子，做了因子纯化之后 IR 提升显著，更能体现因子本身带来的超额收益；送股比例、销售毛利率、存货周转率和每股自由现

金流量比 EPS 由于本身有一定风格偏差, 导致本身超额收益被风格掩盖, 因子纯化之后 IR 也变得显著。

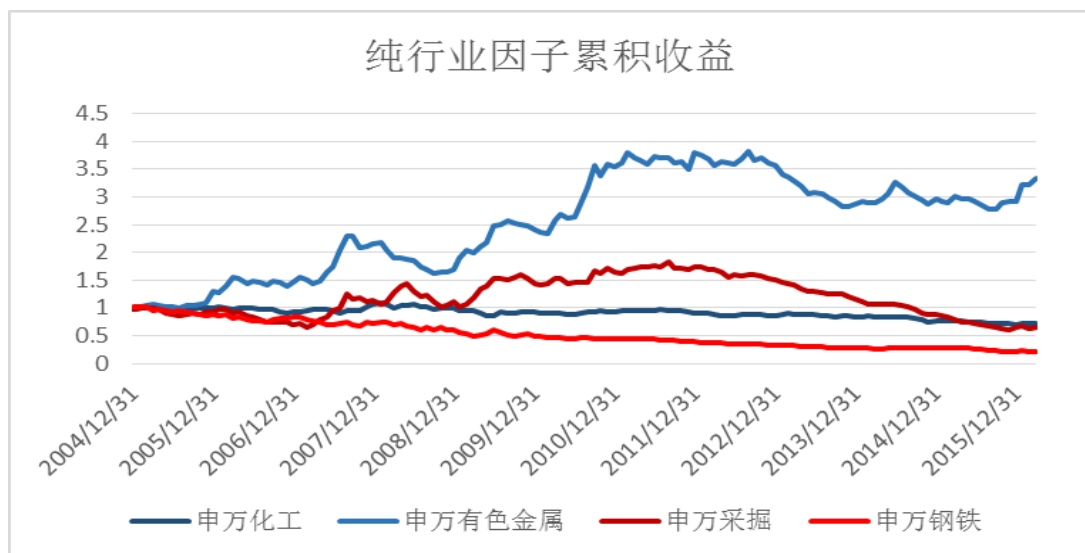
表 7: 排序法低估的一些因子的 IR

	纯因子法IR	排序法IR	二者百分差
净利润增速季度同比	6.42	3.04	111%
三费占比	5.75	2.06	179%
送股比例	2.93	0.25	1095%
销售毛利率	2.49	0.73	240%
PE/G	2.48	1.46	69%
存货周转率	1.35	0.90	51%
每股自由现金流量比 EPS	1.31	-1.26	204%

资料来源: 天软科技, 长江证券研究所

附录 I---- 纯行业因子组合的收益

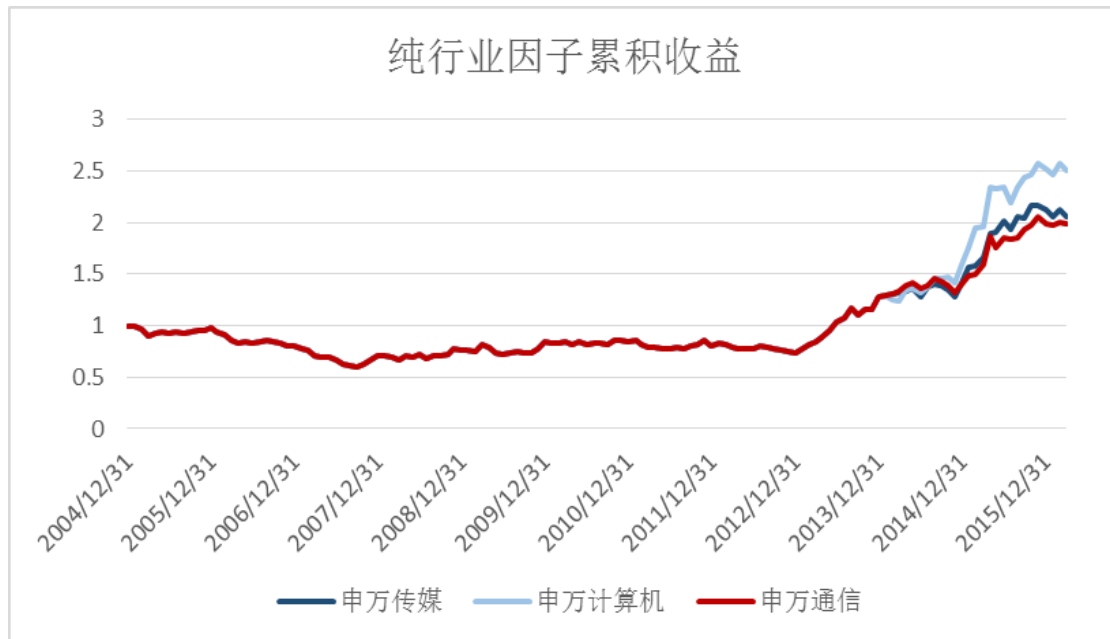
图 I-1: 纯行业因子收益----化工、有色、采掘、钢铁



资料来源: 天软科技, 长江证券研究所

注: 2014 年以前, 申万钢铁取原申万黑色金属行业

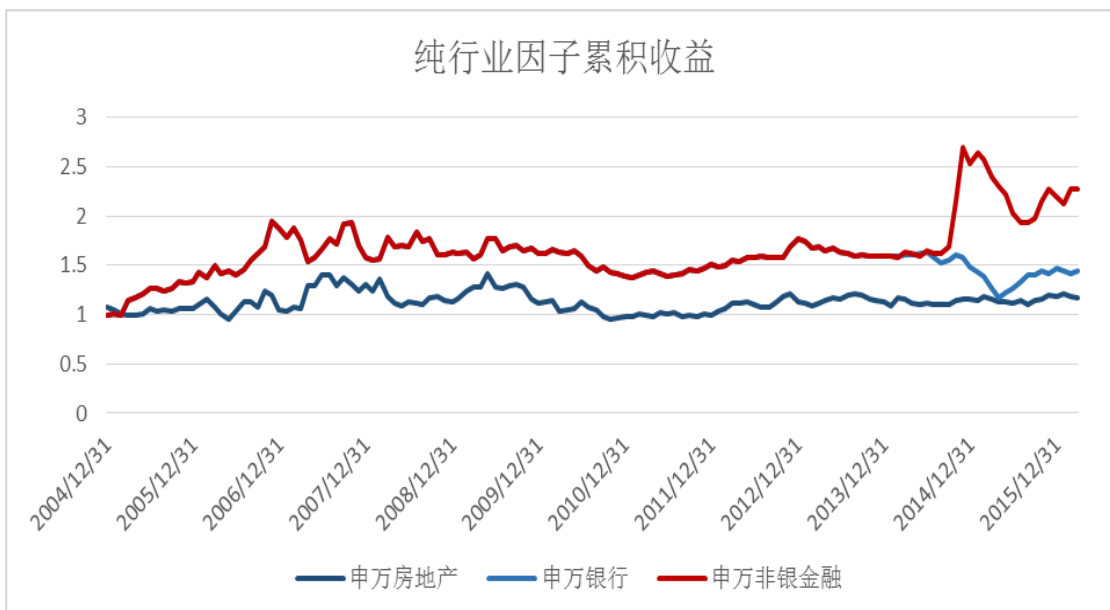
图 I-2: 纯行业因子收益----传媒、计算机、通信



资料来源：天软科技，长江证券研究所

注：2014 年以前，申万传媒、计算机、通信取原申万信息设备、申万信息服务的平均值。

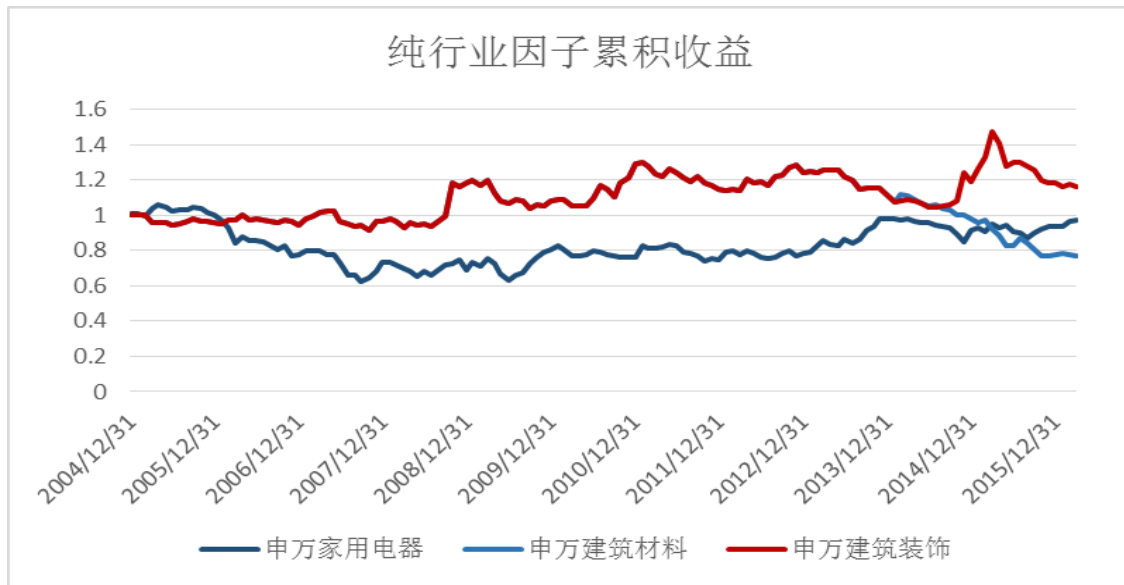
图 I-3: 纯行业因子收益----房地产、银行、非银



资料来源：天软科技，长江证券研究所

注：2014 年以前，申万银行和非银金融同属原申万金融服务行业

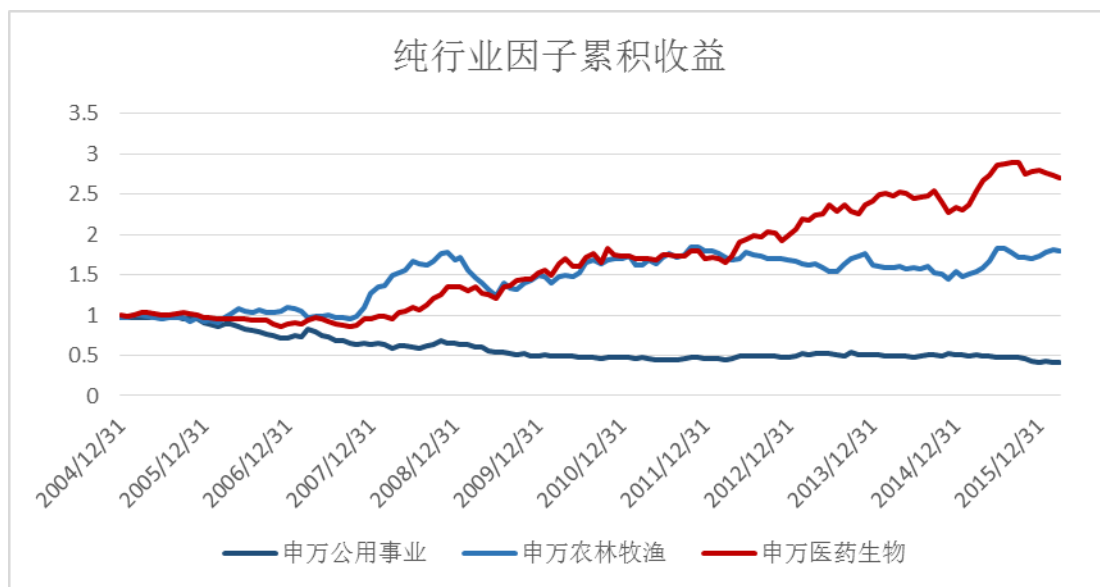
图 I-4: 纯行业因子收益----家电、建筑材料、建筑装饰



资料来源：天软科技，长江证券研究所

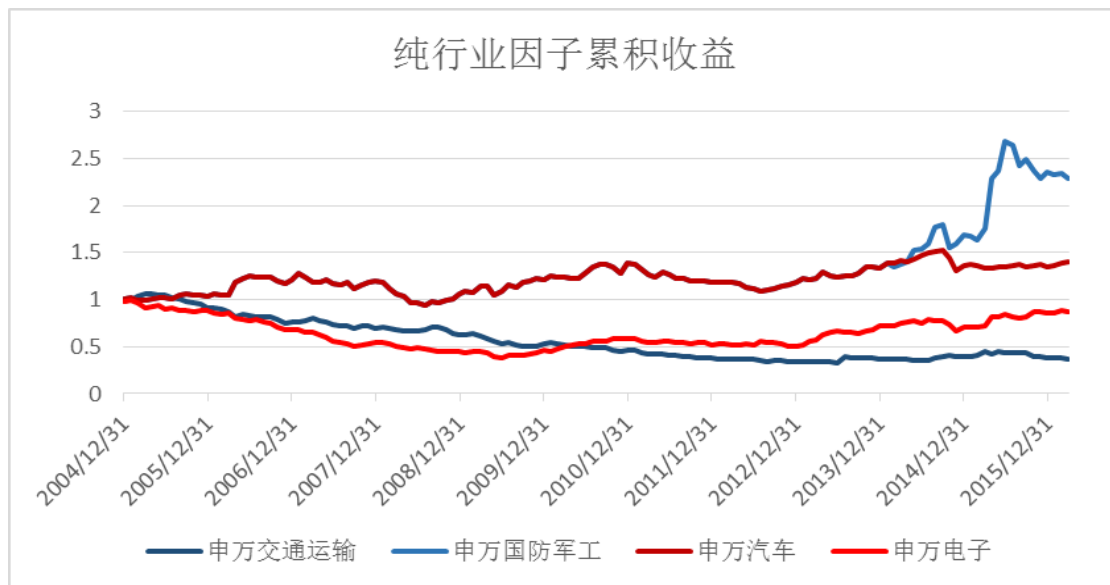
注：2014 年以前，申万建筑材料和建筑装饰同属原申万建筑建材行业

图 I-5: 纯行业因子收益----公用事业、农林牧渔、医药生物



资料来源：天软科技，长江证券研究所

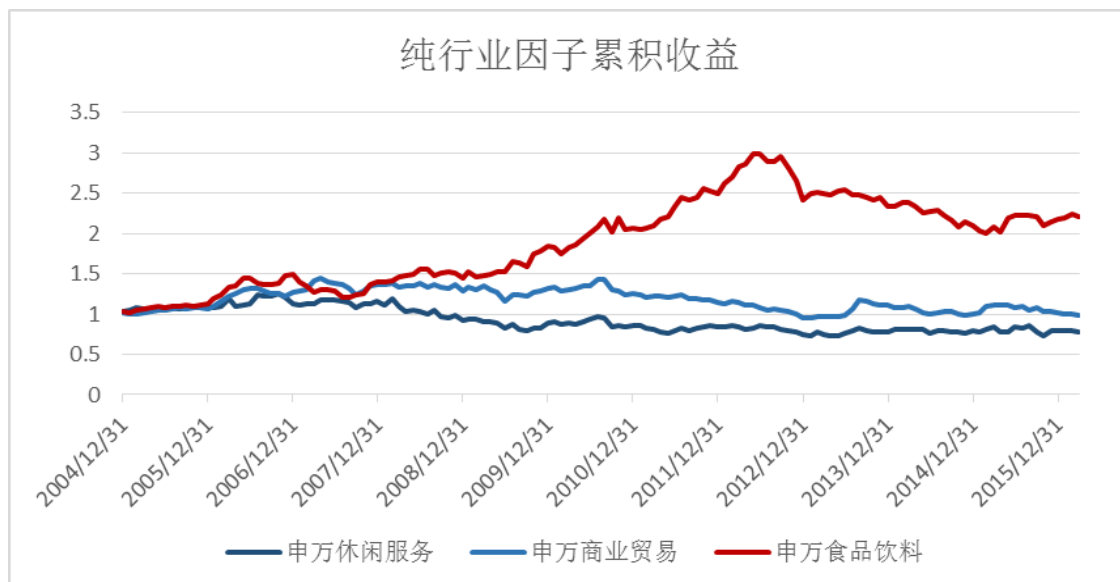
图 I-6: 纯行业因子收益----交通运输、国防军工、汽车、电子



资料来源：天软科技，长江证券研究所

注：2014 年以前，申万国防军工和汽车同属原申万交运设备行业

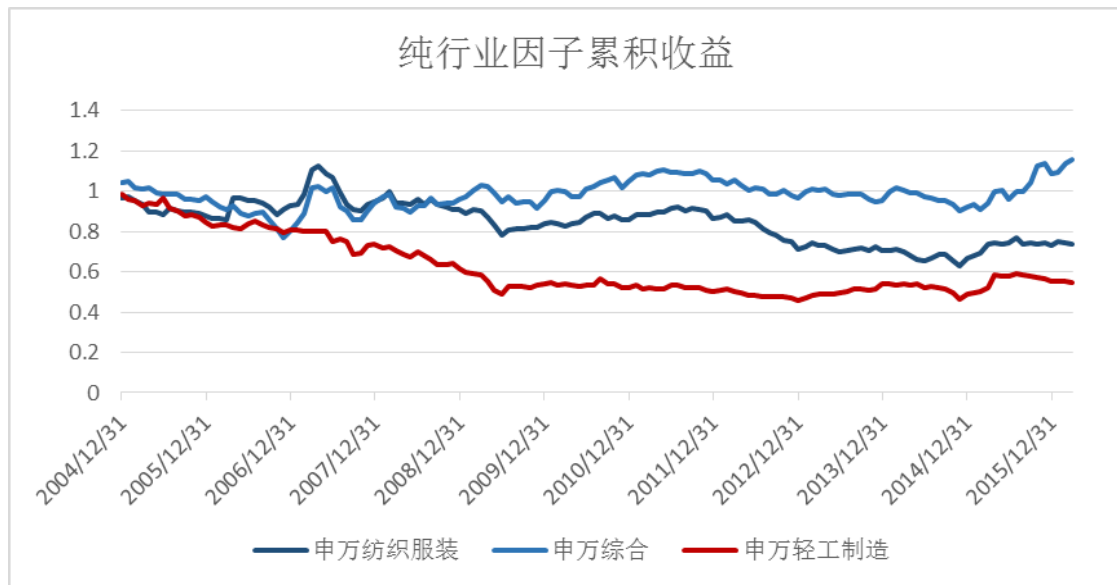
图 I-7: 纯行业因子收益----休闲服务、商业贸易、食品饮料



资料来源：天软科技，长江证券研究所

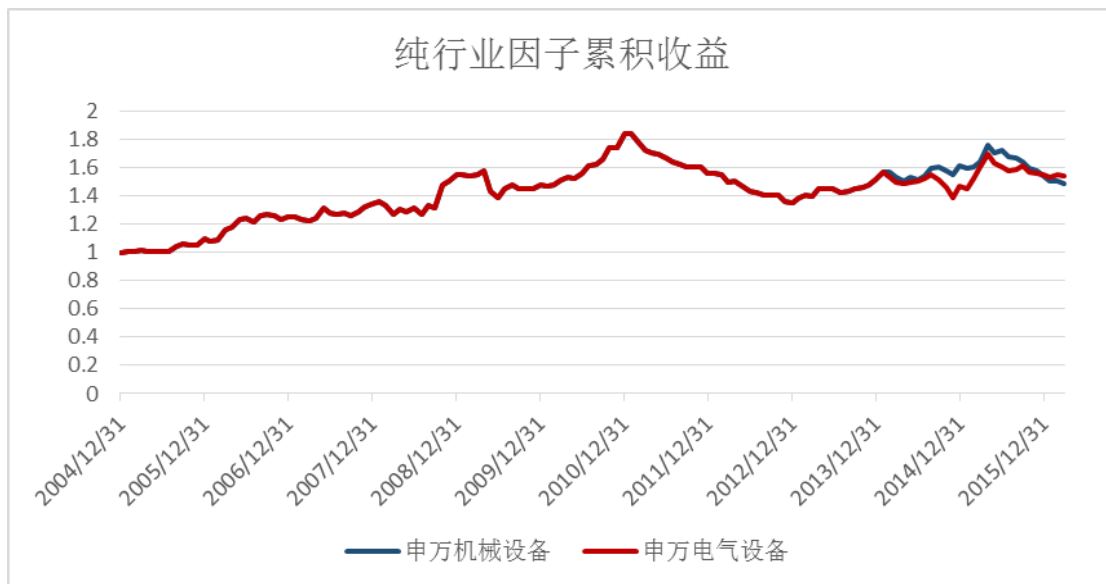
注：2014 年以前，申万休闲服务为原申万餐饮旅游

图 I-8: 纯行业因子收益----纺织服装、综合、轻工制造



资料来源：天软科技，长江证券研究所

图 I-9: 纯行业因子收益----机械设备、电气设备



资料来源：天软科技，长江证券研究所

注：2014 年以前，申万机械设备和电气设备同属原申万机械设备行业

风险提示：数理模型存在风险；历史收益不保证未来回报；回测模拟不代表真实情况

投资评级说明

行业评级	报告发布日后的 12 个月内行业股票指数的涨跌幅度相对同期沪深 300 指数的涨跌幅为基准，投资建议的评级标准为：		
	看 好：	相对表现优于市场	
	中 性：	相对表现与市场持平	
	看 淡：	相对表现弱于市场	
公司评级	报告发布日后的 12 个月内公司的涨跌幅度相对同期沪深 300 指数的涨跌幅为基准，投资建议的评级标准为：		
	买 入：	相对大盘涨幅大于 10%	
	增 持：	相对大盘涨幅在 5%~10%之间	
	中 性：	相对大盘涨幅在-5%~5%之间	
	减 持：	相对大盘涨幅小于-5%	
	无投资评级： 由于我们无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使我们无法给出明确的投资评级。		

联系我们

上海

浦东新区世纪大道 1589 号长泰国际金融大厦 21 楼（200122）
电话：021-68751100 传真：021-68751151

武汉

武汉市新华路特 8 号长江证券大厦 11 楼（430015）
传真：027-65799501

北京

西城区金融大街 17 号中国人寿中心 606 室（100032）
传真：021-68751791

深圳

深圳市福田区福华一路 6 号免税商务大厦 18 楼（518000）
传真：0755-82750808, 0755-82724740

重要声明

长江证券股份有限公司具有证券投资咨询业务资格，经营证券业务许可证编号：10060000。

本报告的作者是基于独立、客观、公正和审慎的原则制作本研究报告。本报告的信息均来源于公开资料，本公司对这些信息的准确性和完整性不作任何保证，也不保证所包含信息和建议不发生任何变更。本公司已力求报告内容的客观、公正，但文中的观点、结论和建议仅供参考，不包含作者对证券价格涨跌或市场走势的确定性判断。报告中的信息或意见并不构成所述证券的买卖出价或征价，投资者据此做出的任何投资决策与本公司和作者无关。

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、价值及投资收入可升可跌，过往表现不应作为日后的表现依据；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；本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司及作者在自身所知知情范围内，与本报告中所评价或推荐的证券不存在法律法规要求披露或采取限制、静默措施的利益冲突。

本报告版权仅仅为本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。如引用须注明出处为长江证券研究所，且不得对本报告进行有悖原意的引用、删节和修改。刊载或者转发本证券研究报告或者摘要的，应当注明本报告的发布人和发布日期，提示使用证券研究报告的风险。未经授权刊载或者转发本报告的，本公司将保留向其追究法律责任的权利。