

MFSR 多因子风格轮动策略 V 1.0

Multi-Factor Style Rotation Strategy V 1.0



2017/7

目录

1. 定义市场风格

2. 多因子风格轮动模型

3. 选股模型

4. 择时策略

5. 总结

6. 优化方向

1. 定义市场风格

根据有无分析师数据库和精简程度，我们可以将市场风格划分为以下 A-D 四类。极简风格下仅有 Risk-on (Size) 和 Risk-off (Valuation) 两类，甚至只做 Size 大小盘风格轮动。我们需要判断的是：通过模型不断择时切换，押注正确市场未来的风格。虽然不可能完全押对，但目的都是力求做了总比什么都不做的好，市场风格的转变比股价的转变缓慢的多，其延续性和可预测性也相对较好些。

A . 极简市场风格分类

编号	风格	因子	因子库编号	定义
1	Size	Total_market_cap	G1	当前总市值
2	Valuation	PB	V1	当前市净率
		Price_to_operating_EPS	V4	"总市值" /过去 4 季度 "营业利润" 总和

B . 精简市场风格分类

编号	风格	因子	因子库编号	定义
1	Size	Total_market_cap	G1	当前总市值
2	Momentum	12M_1M_total_return	M8	21 天前 "后复权收盘价" /252 天前 "后复权收盘价" -1
		Volatility_3M	M11	过去三个月日收益率的标准差
3	Valuation	PB	V1	当前市净率
		Price_to_operating_EPS	V4	"总市值" /过去 4 季度 "营业利润" 总和
4	Technical	VOL10	T133	10 日平均换手率 (Turnover Rate=当日成交金额/总市值)
		Realized_vol_1Y_daily	T2	过去 1 年日涨跌幅的标准差，以日为移动单位
5	Risk	DDNSR	M15	下跌波动 (Downside standard deviations ratio) ，过往 12 个月中，大盘指数日收益为负时，个股日收益标准差和大盘指数日收益标准差之比。
		Merton_distance_to_default	Q6	("资产总计" -("流动负债合计"+0.5*"长期负债合计"))/过去 1 年"总市值"的标准差

C. 分析师预测数据库下市场风格分类

编号	风格	因子	因子库编号	定义
1	Size	Total_market_cap	G1	当前总市值
2	Momentum	12M_1M_total_return	M8	21 天前 “后复权收盘价” /252 天前 “后复权收盘价” -1
3	Yield	FY1_dividend_yield_mean	V14	未来一年股息率预测中值
4	Valuation	PB	V1	当前市净率
		FY1_PE_mean	V16	未来一年 PE 预测中值
5	Revision	Revision index	S22	(过去 3 个月内乐观预计研究报告数-过去 3 个月内 v 悲观预计研究报告数) /过去 3 个月内研究报告总数
6	Growth	Sales growth (FY2)	S20	未来两年营业收入/未来一年营业收入
		EPS growth (FY2)	S21	未来两年 EPS/未来一年 EPS
7	Profitability	ROE FY1	S23	未来一年预测净利润/当前所有者权益
		Change in pretax profit margin	S24	未来两年预测毛利率 - 未来一年预测毛利率
8	Risk	Volatility_1Y	M10	过去一年日收益率的标准差
		Merton_distance_to_default	Q6	(“资产总计”- (“流动负债合计”+0.5*“长期负债合计”))/过去 1 年“总市值”的标准差
		Estimate dispersion	S25	未来一年 EPS 预测的方差/未来一年 EPS 预测的中位数

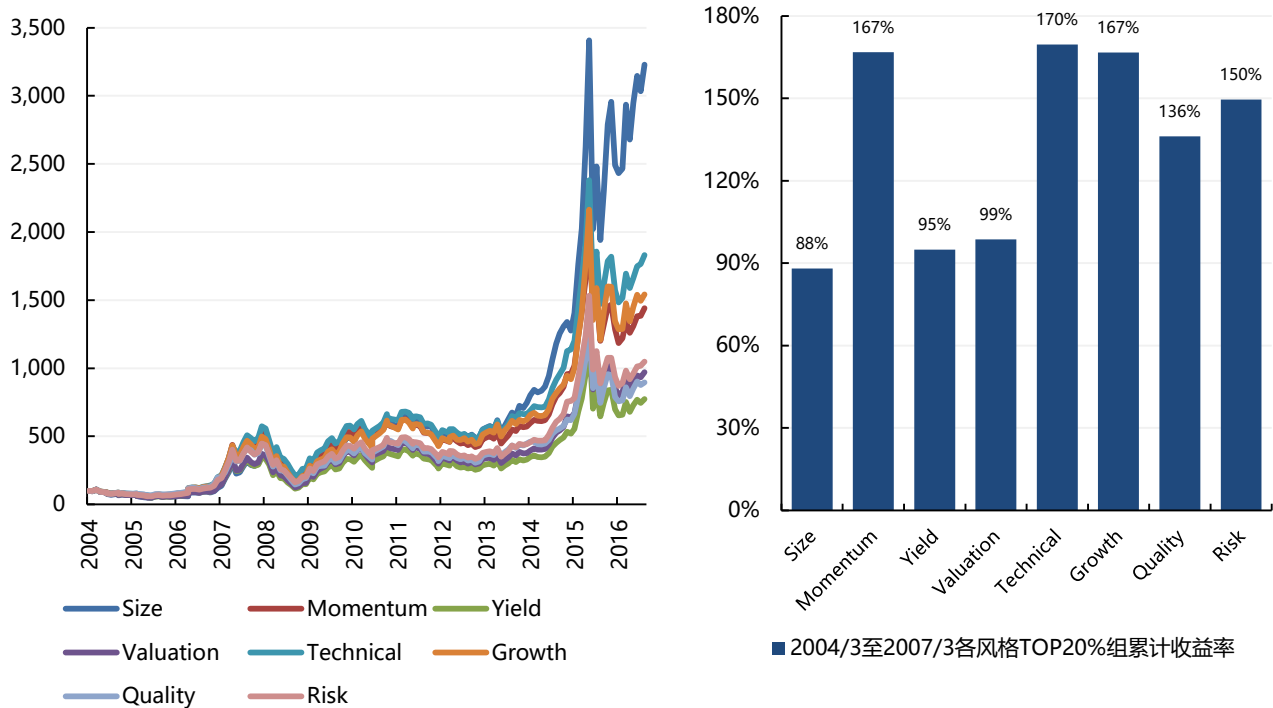
D. 无分析师预测数据库下市场风格分类

编号	风格	因子	因子库编号	定义
1	Size	Total_market_cap	G1	当前总市值
2	Momentum	12M_1M_total_return	M8	21 天前 “后复权收盘价” /252 天前 “后复权收盘价” -1
		Volatility_3M	M11	过去三个月日收益率的标准差
3	Yield	Dividend_yield	V9	过去 4 季度 “分配股利、利润或偿付利息支付的现金” /过去 4 季度 “净利润”
4	Valuation	PB	V1	当前市净率
		Price_to_operating_EPS	V4	“总市值” /过去 4 季度 “营业利润” 总和
5	Technical	VOL10	T133	10 日平均换手率 (Turnover Rate=当日成交金额/总市值)
		Realized_vol_1Y_daily	T2	过去 1 年日涨跌幅的标准差, 以日为移动单位
6	Growth	Sales_SQ_YoY	G8	最新季度 “营业总收入” /过去第 5 个季度 “营业总收入” -1
		Net_profit_SQ_YoY	G7	最新季度 “净利润” /过去第 5 个季度 “净利润” -1
7	Quality	Return_on_assets	Q8	过去 4 季度 “净利润” /最近季度 “资产总计”
		Berry_ratio	Q11	(过去 4 季度 “营业总收入” -过去 4 季度 “营业总成本”) / 过去 4 季度 “营业总成本”
8	Risk	DDNSR	M15	下跌波动 (Downside standard deviations ratio), 过往 12 个月中, 大盘指数日收益为负时, 个股日收益标准差和大盘指数日收益标准差之比。
		Merton_distance_to_default	Q6	(“资产总计”- (“流动负债合计”+0.5*“长期负债合计”))/过去 1 年“总市值”的标准差

2. 多因子风格轮动模型

对于只有一个因子的市场风格，我们生成该风格指数的方法为：在每个行业¹内部根据该因子值²进行排序，取行业分组内排名前 20%分位的股票等权组合计算收益率（或做多排名前 20%分位的股票等权组合，同时做空排名最后 20%分位的股票未来一个月的等权组合的收益率。虽然 A 股不能做空，但此处仅为风格择时计算各市场风格的收益率）。该收益率即可理解为该风格单边做多³（或双边多空）的收益率。

对于有多个因子的市场风格，我们在每个行业内部对每个股票的因子值做 Zscore，然后等权加总股票在行业内部各因子的分值，取行业分组内排名前 20%分位的股票等权组合计算收益率（或做多 Top 20%，做空 Bottom 20%，同上）。在行内分组内进行因子值排序的做法实际上是行业中性化处理的最好方案，能使得最终的组合保持较高的行业平衡度。计算各风格历史收益率曲线后发现市值因子的收益遥遥领先。但值得注意的是市值因子仅在 08 年后开始发力，此前曾长期落后于其他风格。



¹ 证监会 2012 版二级行业分类

² 不做中性化处理，后续会在行业内部做中性化

³ 由于不能做空，此处我们最终采用单边做多作为风格的收益率

我们对各风格 Top 组的收益率总结如下表，计算超额收益率相关系数可以发现：某些风格间相关性较弱，甚至有不少风格是负相关的（互补）。尝试风格轮动的现实依据源于此。倘若各风格之间相关系数较高，互为替代关系，就失去了风格轮动的意义。

收益率统计								
	Size	Momentum	Yield	Valuation	Technical	Growth	Quality	Risk
年化收益率	31.5%	23.2%	17.4%	19.5%	25.5%	23.9%	18.7%	20.2%
标准差	18.6%	16.9%	15.9%	16.7%	16.5%	16.1%	14.0%	15.7%
IR	1.69	1.38	1.10	1.17	1.54	1.48	1.34	1.29

月度超额收益率*相关系数矩阵								
	Size	Momentum	Yield	Valuation	Technical	Growth	Quality	Risk
Size	1.00	-0.55	0.53	0.48	-0.74	-0.11	-0.58	-0.79
Momentum		1.00	-0.59	-0.34	0.65	-0.37	0.04	0.68
Yield			1.00	0.56	-0.75	0.05	-0.58	-0.50
Valuation				1.00	-0.52	-0.39	-0.77	-0.32
Technical					1.00	-0.22	0.43	0.69
Growth						1.00	0.46	-0.31
Quality							1.00	0.23
Risk								1.00

注：* 此处的超额收益率是各风格和 8 个收益率的平均值之差（也可以做成相对全市场平均收益率的超额收益率）。由于风格收益率的计算采用 Top 组的收益率，并非采用 Top-Bottom 组的收益率之差作为风格收益率，因此风格相对指数的超额收益率相关性较大，不适合用于此处。

和其他市场现象一样，导致市场风格转变的是：反转和趋势。首先，中长期来看，价值均值回归的反转是不变的主题，例如市场价格连续长期偏离其历史平均水平（暂理解为价值），价值因子的区分度较高，则较为可能发生价值均值回归的反转现象；其次，短期来看，风格一般具有较为明显的趋势性，近期表现强劲的风格大概率在未来短期表现依然良好。除了反转和趋势，学术研究表明，过往收益率可能影响市场的风险偏好情绪。若最近指数的累计收益率大幅为正，投资者普遍盈利，更有可能愿意冒险承担更多风险；反之则市场情绪较为谨慎，熊市中更愿意防御为主。由此，我们就可以通过反转、趋势和风险偏好这三个层面构建风格择时模型。

2.1 价值反转

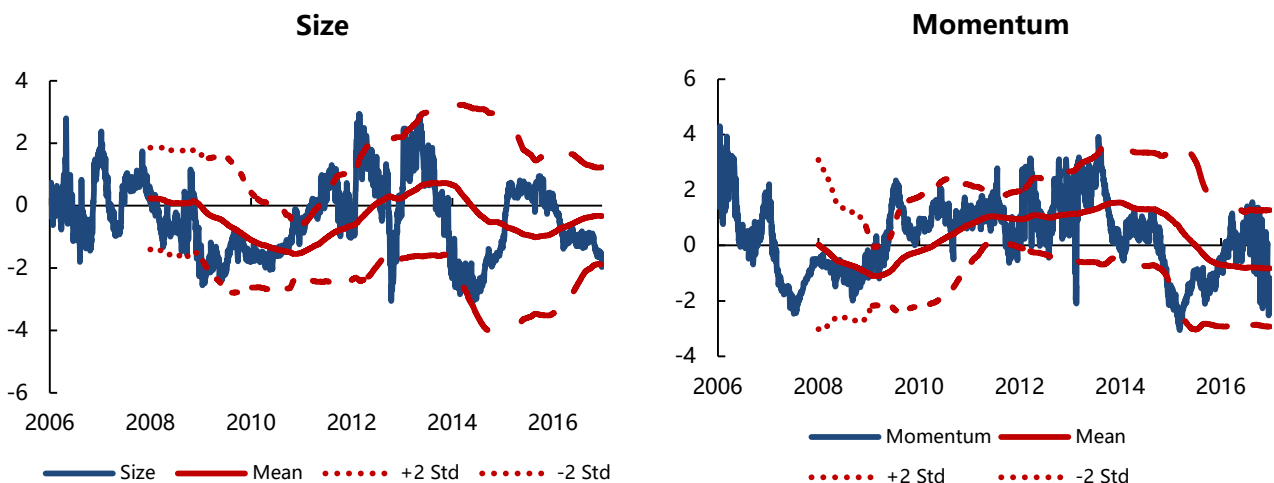
我们首先计算 1-8 个市场风格 Top 20%和 Bottom 20%组合的 BP 中位数，然后将这两个 Top 20%和 Bottom 20%组合的 BP 中位数作差，再减去历史均值（以月度为单位，过去所有时间样本的均值，逐步扩大的时间窗口⁴，但至少要有 24 个月数据），最后除以历史标准差。最后再对 BP Spread 做图，以其 2 个标准差为上下阈值。当 BP Spread 超过其 2 个标准差⁵下界（上界）时往往是价值因子极度高估、无区分度（低估、明显区分度）之时。

$$\text{BP Spread} = [(\text{Bottom 20\%组合的 BP 中位数} - \text{Top 20\%组合的 BP 中位数}) - \text{历史 BP Spread 均值}] / \text{历史 BP Spread 标准差}$$

$$\text{EP Spread} = [(\text{Bottom 20\%组合的 EP 中位数} - \text{Top 20\%组合的 EP 中位数}) - \text{历史 EP Spread 均值}] / \text{历史 EP Spread 标准差}$$

对每个市场风格我们都进行 BP\EP Spread 的计算，并计算 Zscore 值，根据 Zscore 值对 1-8 个市场风格分别予以 1-8 的排名分值（排名越高，分值越大，越好）。以上方法的潜在原理是：BP Spread Zscore 越大的市场风格，其组内上下 20%分位组的价值差距越大，价值因子越容易对组内股票进行区分进而产生超额收益，正所谓组内价值层面的维度相对清晰。

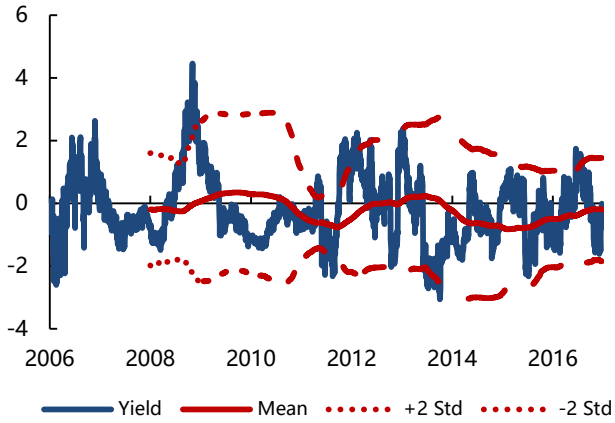
以 **D. 无分析师预测数据库** 下市场风格分类为例，各风格的历史 BP Spread 如下图：



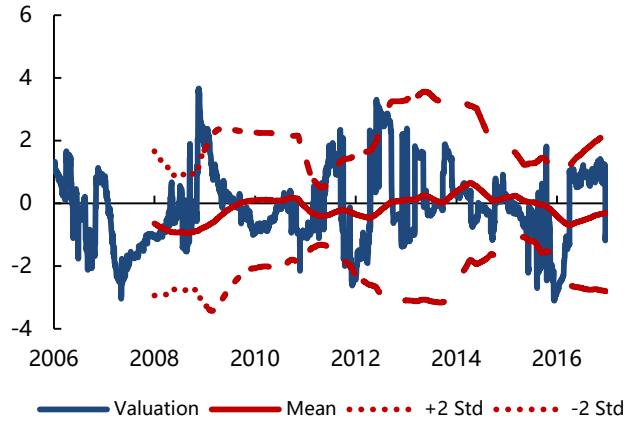
⁴ 历史均值的时间窗口尝试过改为过去 24 个月移动时间窗口，但效果不如扩展的全历史样本时间窗口

⁵ 近似 99%置信度下 1.96 的 t 检验阈值

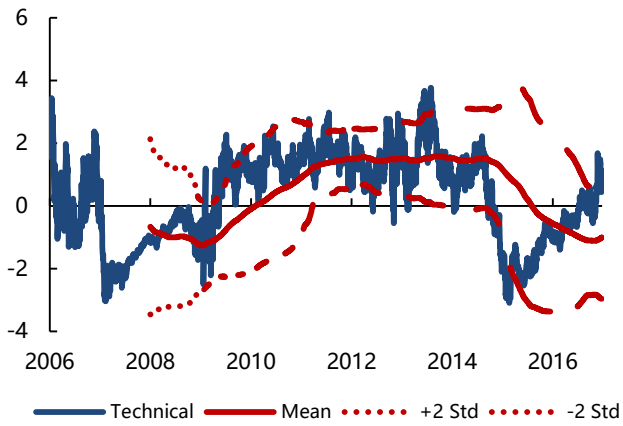
Yield



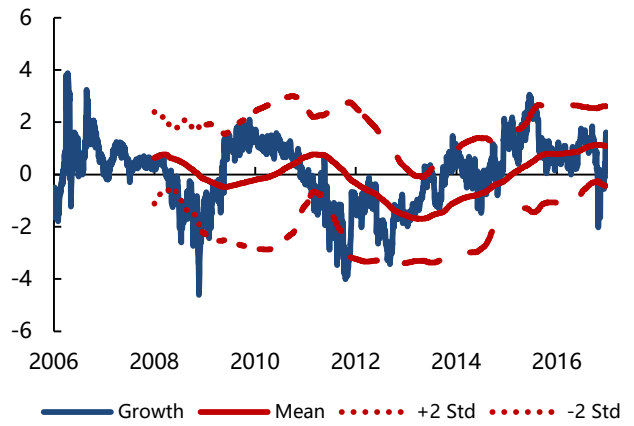
Valuation



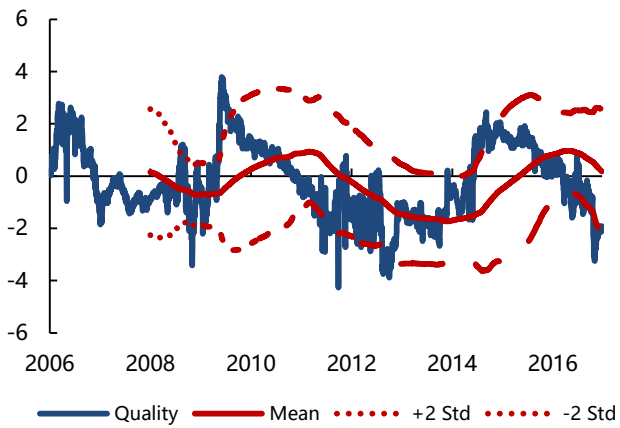
Technical



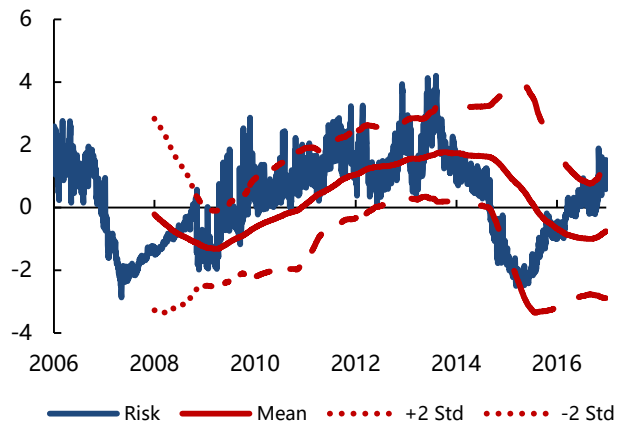
Growth



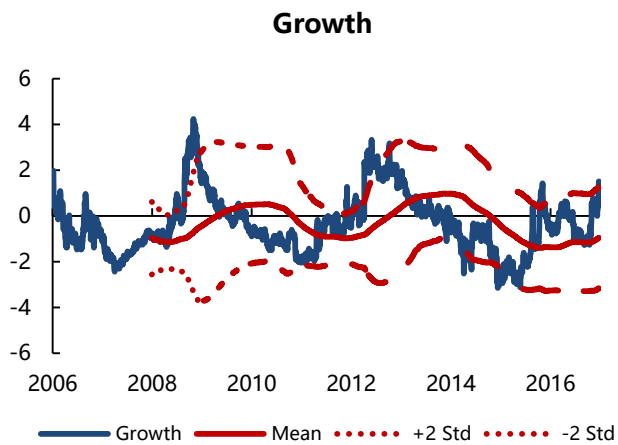
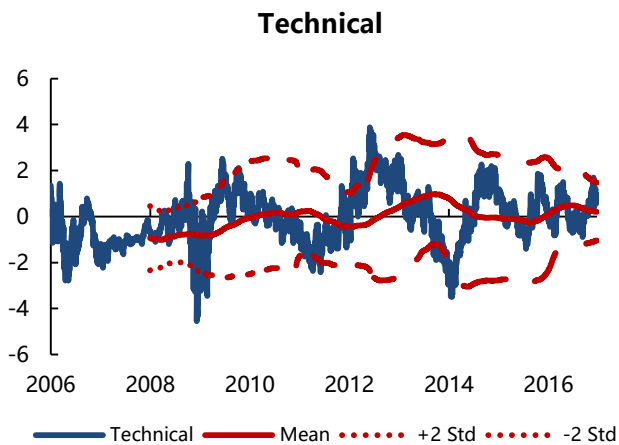
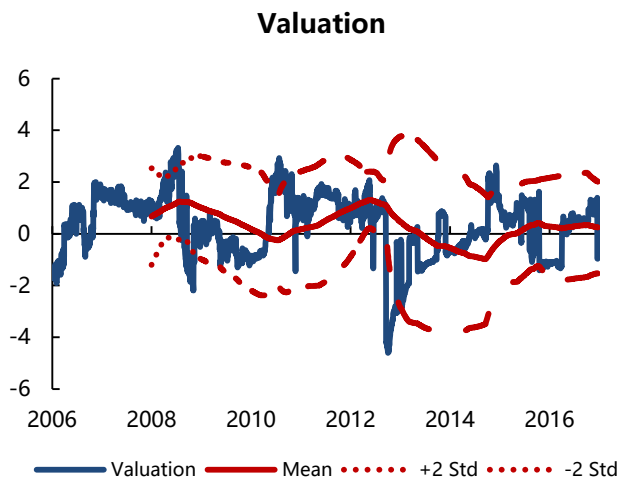
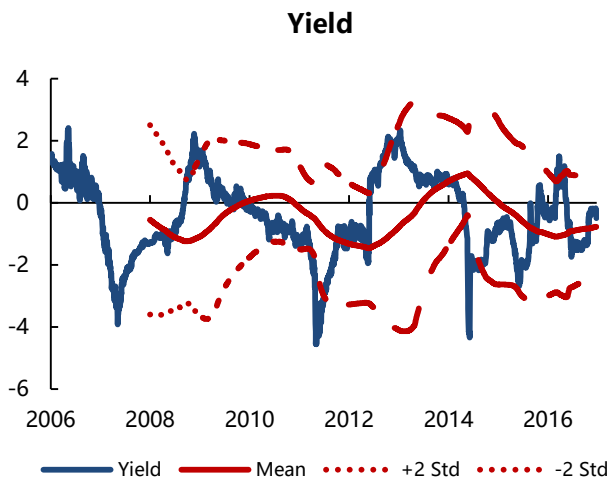
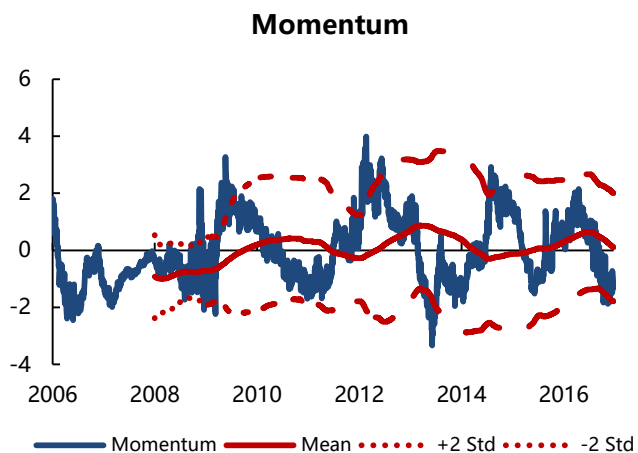
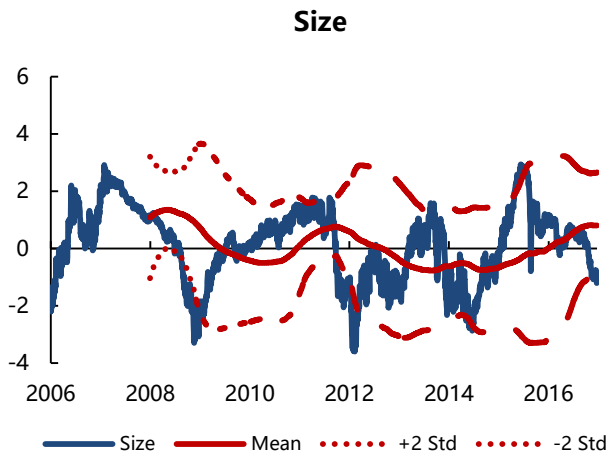
Quality

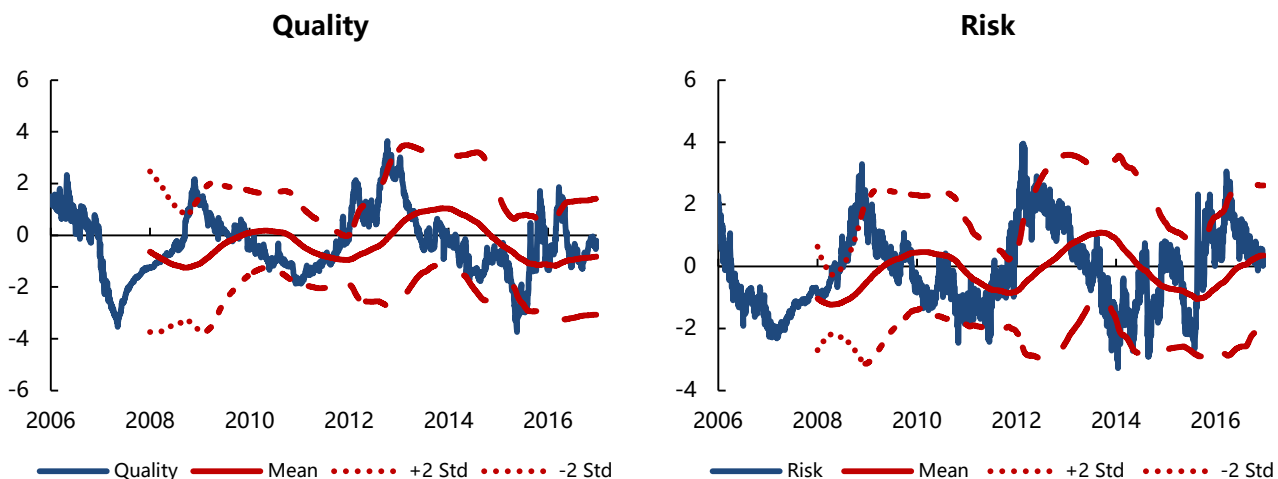


Risk



各风格的历史 EP Spread 如下图：





2.2 趋势延续

趋势的计算较为简单：计算各市场风格过去 N 个月的累计收益率为指标进行排名，根据累计收益率大小对 1-8 个市场风格分别予以 1-8 的排名分值（排名越高，分值越大，越好）。

2.3 有效性检验

事实上我们将 BP Spread、EP Spread 和趋势这三个因子作为挑选风格的分类器，和用因子结合挑选股票的多因子模型是类似的，分类器所运用的层面从股票转移到了因子或风格。因此，我们也可以对这三个因子计算风格层面的 IC 时间序列图，以初步检测其有效性。首先是周频率，BP Spread、EP Spread 的 IC 总体上均为负，效果欠佳。

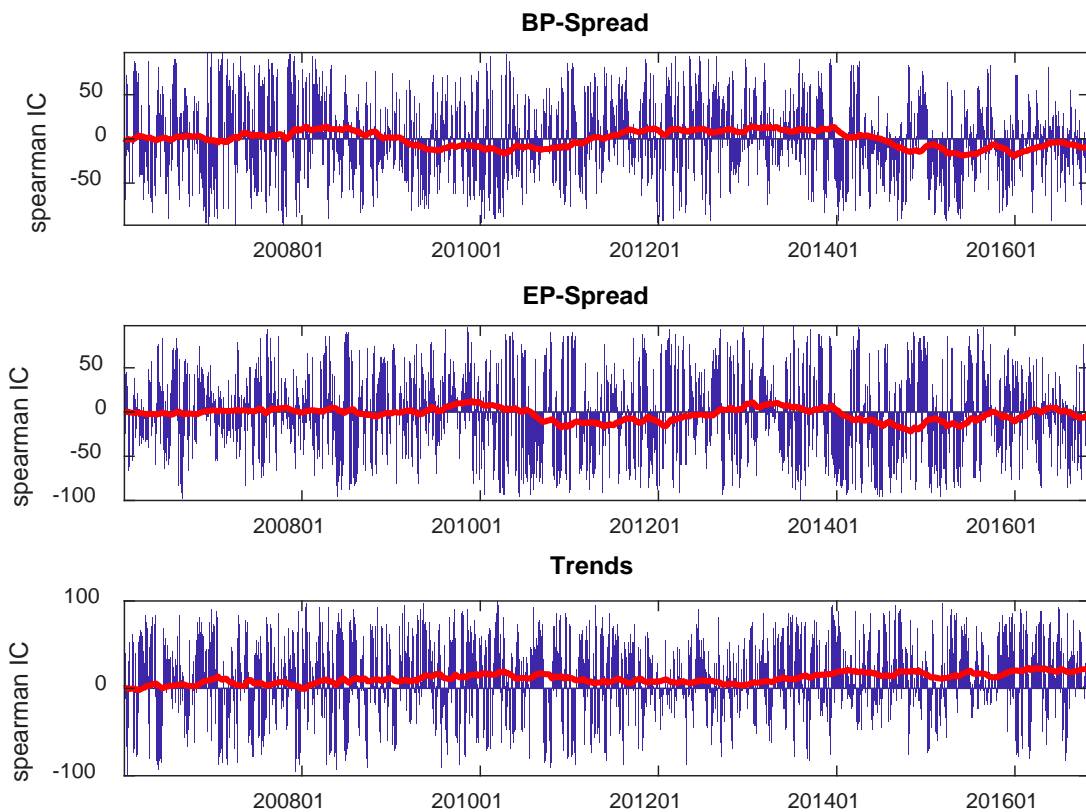
周度 IC

	原始值		12M 均线	
	Mean	Std	Mean	Std
BP Spread	-0.81	45.07	-0.45	9.14
EP Spread	-2.96	48.55	-2.49	7.24
Trends	12.19	43.94	11.24	5.70

但观察 IC 时间序列图⁶时发现：趋势 IC 较弱时 BP/EP Spread 的 IC 较强，相关系数矩阵的结果也印证了两者之间的互补性。

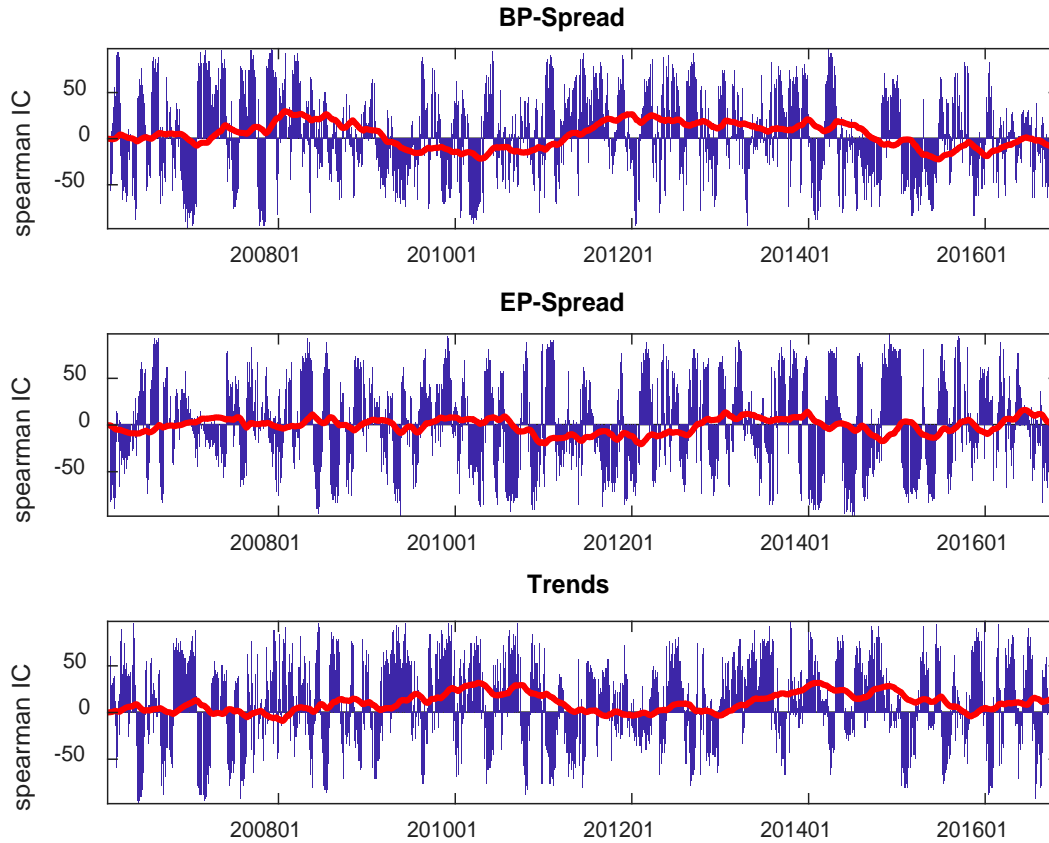
周度相关系数

	原始值			年度均线		
	BP Spread	EP Spread	Trends	BP Spread	EP Spread	Trends
BP Spread	1.00			1.00		
EP Spread	0.37	1.00		0.28	1.00	
Trends	-0.35	-0.16	1.00	-0.59	-0.18	1.00



月度频率下的统计值也证明趋势是一个强因子，BP 和 EP Spread 因子与趋势因子互补。BP 因子本身在月度统计下尚微弱有效（月度结果较周度结果有所改善，可能是价值类因子发散速度较慢），EP 因子基本无效，但存在的重要意义是和趋势互补。

⁶ 图中每个点都是用当日的月度收益和一个月前（20日）的 BP Spread、EP Spread、Trends 等的 spearman 相关系数。统计时每日结果都进行了计算，画在了一张图上。每日结果都进行了统计，图中红线是根据每日的 IC 计算的平均值，和做因子分析的时候略有不同，因子分析的时，统计时间间隔是每个月，不像现在是每天都进行统计，取样频率不同。



月度 IC

	原始值		年度均线	
	Mean	Std	Mean	Std
BP Spread	3.43	45.82	3.76	13.29
EP Spread	-1.31	47.32	-0.97	7.91
Trends	11.02	44.75	10.27	9.71

月度相关系数

	原始值			年度均线		
	BP Spread	EP Spread	Trends	BP Spread	EP Spread	Trends
BP Spread	1.00			1.00		
EP Spread	0.37	1.00		0.05	1.00	
Trends	-0.29	-0.18	1.00	-0.43	0.12	1.00

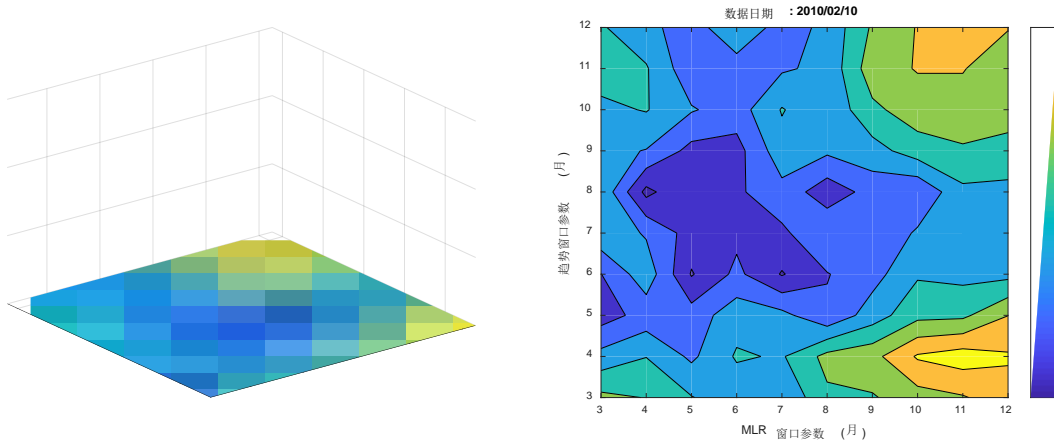
2.4 混合面板回归模型

在得到 EP Spread、BP Spread 和趋势后，我们可以对风格的未来收益排名作风格和时间序列两个维度的面板回归。由于风格本身有异质性，且随时间轮动，因此很可能同时存在固定效应和随机效应，混合面板回归是较为合适的方法⁷。 $Rank_{return,i,t+1}$ 为因子 i 在 t+1 时间的收益排名， $Rank_{EP,i,t}$ 为 EP Spread 排名。在求得 EP Spread、BP Spread 和趋势项的回归系数⁸后，用以下一期的排名预测。

$$Rank_{return,i,t+1} = \alpha + \beta_1 Rank_{EP,i,t} + \beta_2 Rank_{BP,i,t} + \beta_3 Rank_{Trend,i,t} + e$$

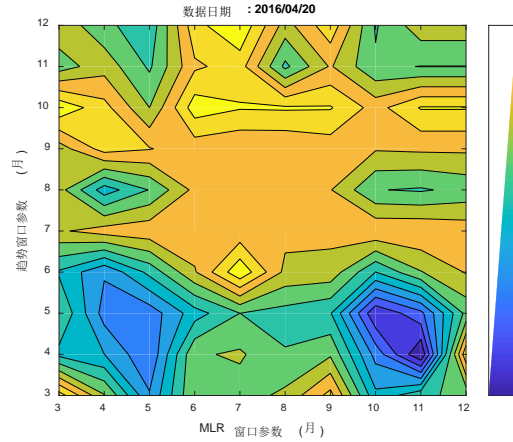
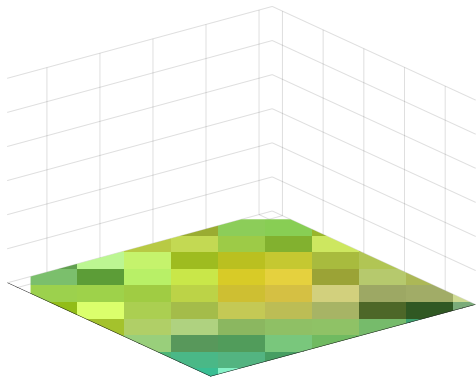
2.5 参数选择

此处趋势项和混合面板回归所用的移动样本窗口长度是两个关键参数。参数的选择虽然多少代有些主观性，但应当最大可能地避免使用未来信息。所以，我们采用移动时间窗口下遍历双参数选取最优的方法。具体来说，每周均利用最近 6 个月的历史样本信息，罗列所有双参数的排列组合，回测并选取最近 6 个月押注排名前两位风格所得累计收益率（评价函数）最高的组合作为下一期的参数。两个参数的范围均以 1 个月为间隔选取 3 至 12 个月中的 10 个参数。12 个月以上的时间窗口会让模型反应迟缓，3 个月以下的间隔单位会使计算量成倍增长且信号可能过于闪烁。下左图为某一天所有 100 (10*10) 个参数组合下的历史收益净值，图中的每一个点（每次历史净值收益）均为一次单独回测，倘若有 2000 个交易日，那么需要进行 20 万次回测。以 2010/2/10 和 2016/4/20 为例，最优参数会随着时间而剧烈变化，结果最优的黄色区域发生了明显的变动。

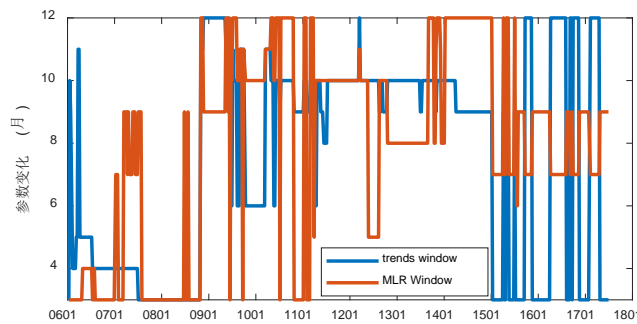


⁷ 后续需要做 Hausman 检验，以证实上述说法

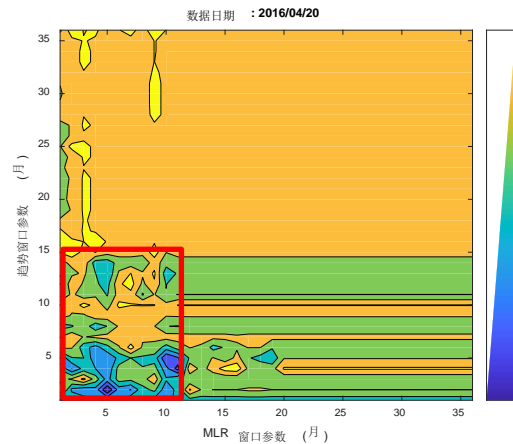
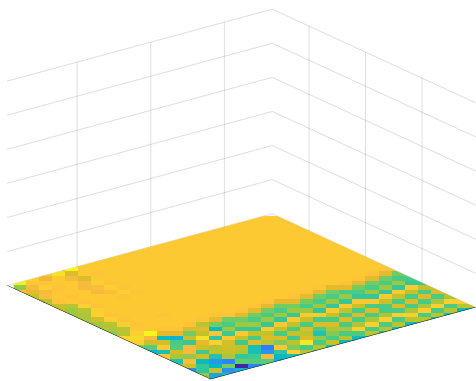
⁸ 负值系数归 0 处理，意义是不做空该因子

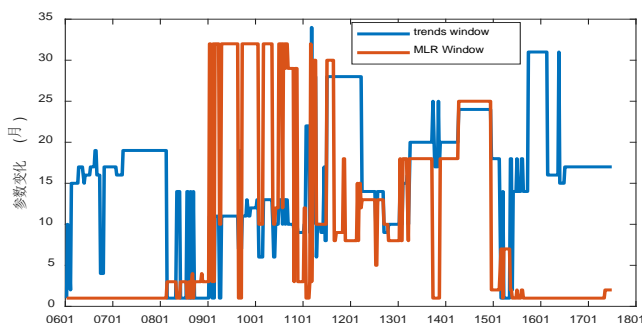


下图为所选择的双参数的时间序列图，最优参数的波动性较强。



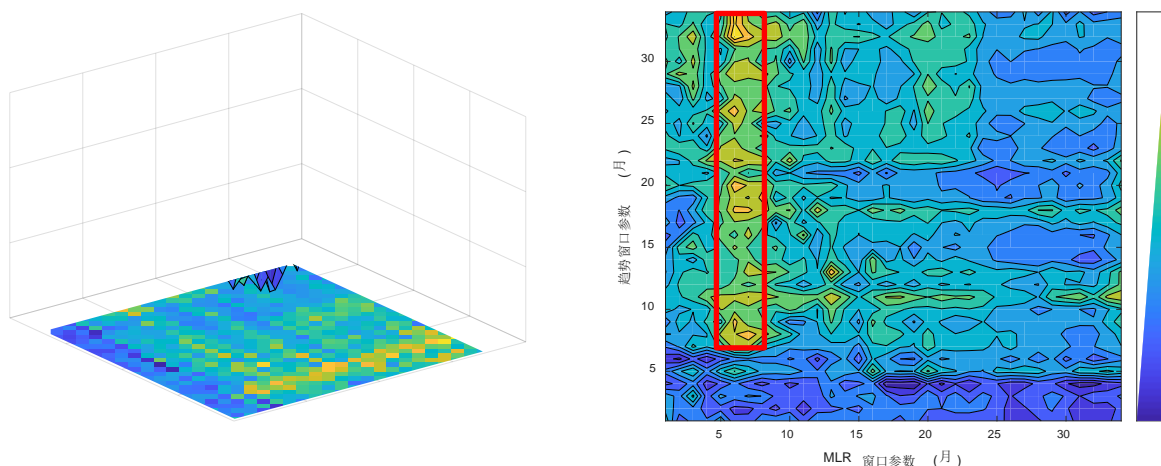
虽然以上方法动态地选取了参数组合，但依然人为主观设置了双参数 3 个月和 12 个月的上下限。为考察合理性，我们可以尝试保持其他条件不变，仅将 12 个月的双参数上限提高至 36 个月，下限降低至 1 个月。随机抽取 2016/4/20 的结果显示：右下图 15*12 红色区域以外的参数区域对结果的影响非常小。拉长参数范围并不能提高模型历史回测效果，而且面对未来的中短期变化可能非常迟缓，有弊而无利。因此双参数的上下限设置有一定的合理性。



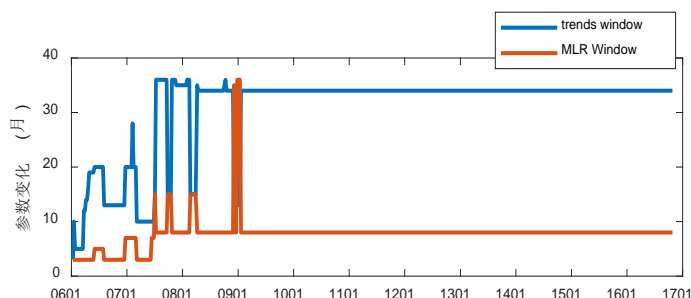


值得注意的是：上面的做法并非历史回测结果最好的，我们力求从逻辑出发，权衡短期闪烁和中长期趋势，合理设定参数范围，而非完全根据回测选取能让历史结果最好的设定。完全避免任何主观设定的难度较大，例如上述评价函数（最近 6 个月押注排名前两位风格所得累计收益率）的设定就带有一定的主观性，如此设定并非为了追求最好的历史回测效果，而是意在进一步提高模型的反应灵敏度。

为进一步证明上述说法，可以变更评价函数为：采用扩大时间窗口下押注排名第一位风格所得累计收益率历遍双参数选取最优的方法。即每周均利用截至当前时刻的所有历史样本信息，罗列所有双参数的排列组合，回测并选取仅押注一种风格下历史累计收益率最高的组合作为下一期的参数。两个参数的范围均以 1 个月为间隔选取 1 至 36 个月中的 36 个参数。



由于历史上 Size 风格趋势性极强，间接导致趋势项的参数窗口越长越好，一段时间后趋势项最优参数一直维持在最大的 36 个月，回归样本长度维持在 8 个月（右上图红框区域出现了参数平原）。这样的结果虽然能提高整体风格预测正确率和策略的收益率，但实不可取。其对突如其来的市值反转需要非常长的时间做出反应，且有过拟合的嫌疑。



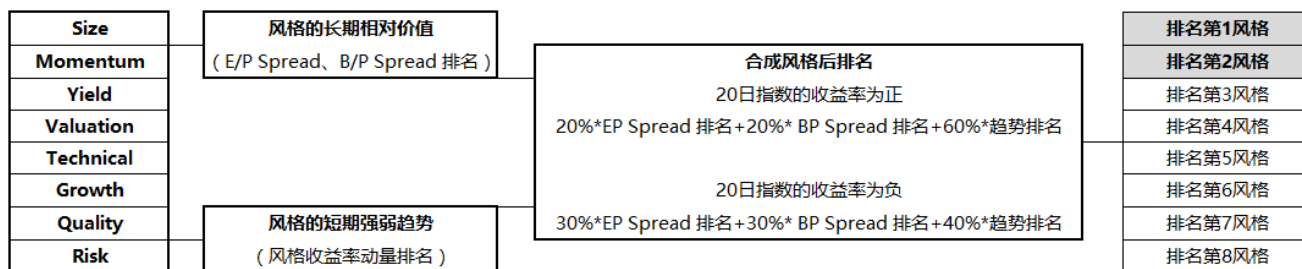
2.6 初始静态权重

在初始的样本期内，由于无法做面板回归，因此需要设定一个同意的初始静态权重。在未知的情况下等权价值⁹和趋势是较为自然的选择。与此同时，我们结合之前提到的市场风险偏好：若过去 20 个交易日指数的收益率为正，投资者普遍盈利，更有可能愿意冒险承担更多风险，因此我们予以趋势延续更多的比重；反之，若过去 20 个交易日指数的收益率为负，则予以价值反转更多的比重。由此，初始风格择时模型可以表述为：

当过去 20 个交易日指数的收益率为正，各市场风格的综合得分 = 20% * EP Spread 排名 + 20% * BP Spread 排名 + 60% * 趋势排名；

当过去 20 个交易日指数的收益率为负，各市场风格的综合得分 = 30% * EP Spread 排名 + 30% * BP Spread 排名 + 40% * 趋势排名；

最后根据当月各市场风格的综合得分，选定得分第 1 名（或前 2 名）¹⁰ 的市场风格作为下月市场风格，并等权单边做多该两组市场风格的组合。

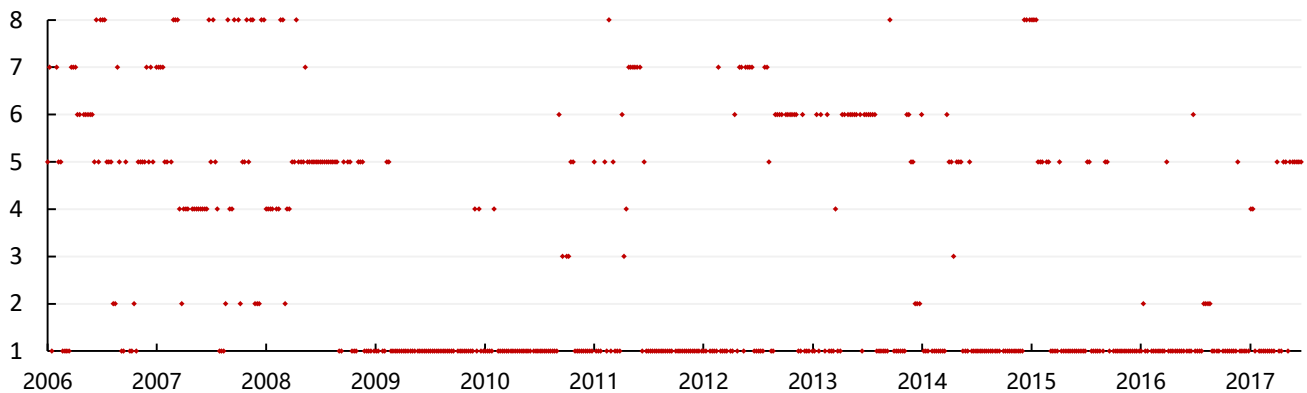
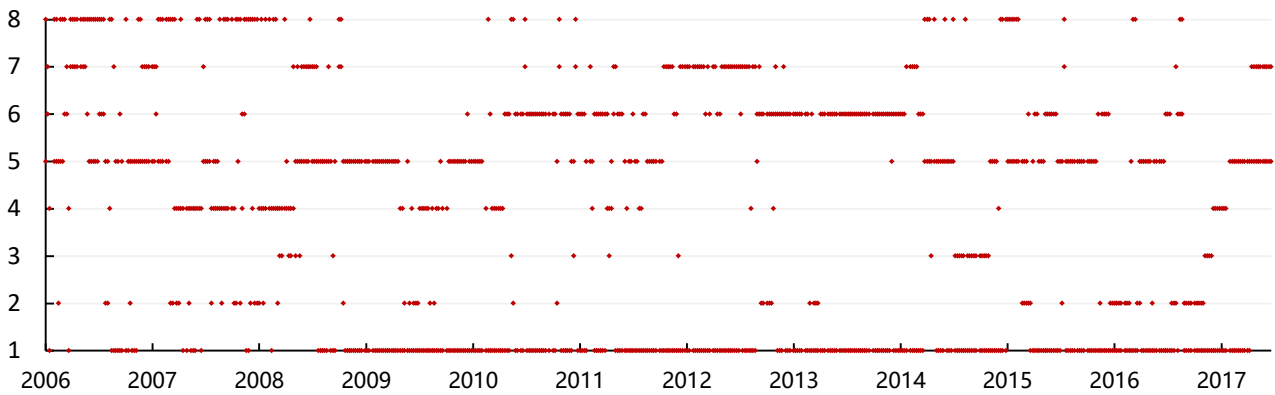


⁹ 价值反转的变量 BP、EP 单独的效果不如 BP 和 EP 结合的方式

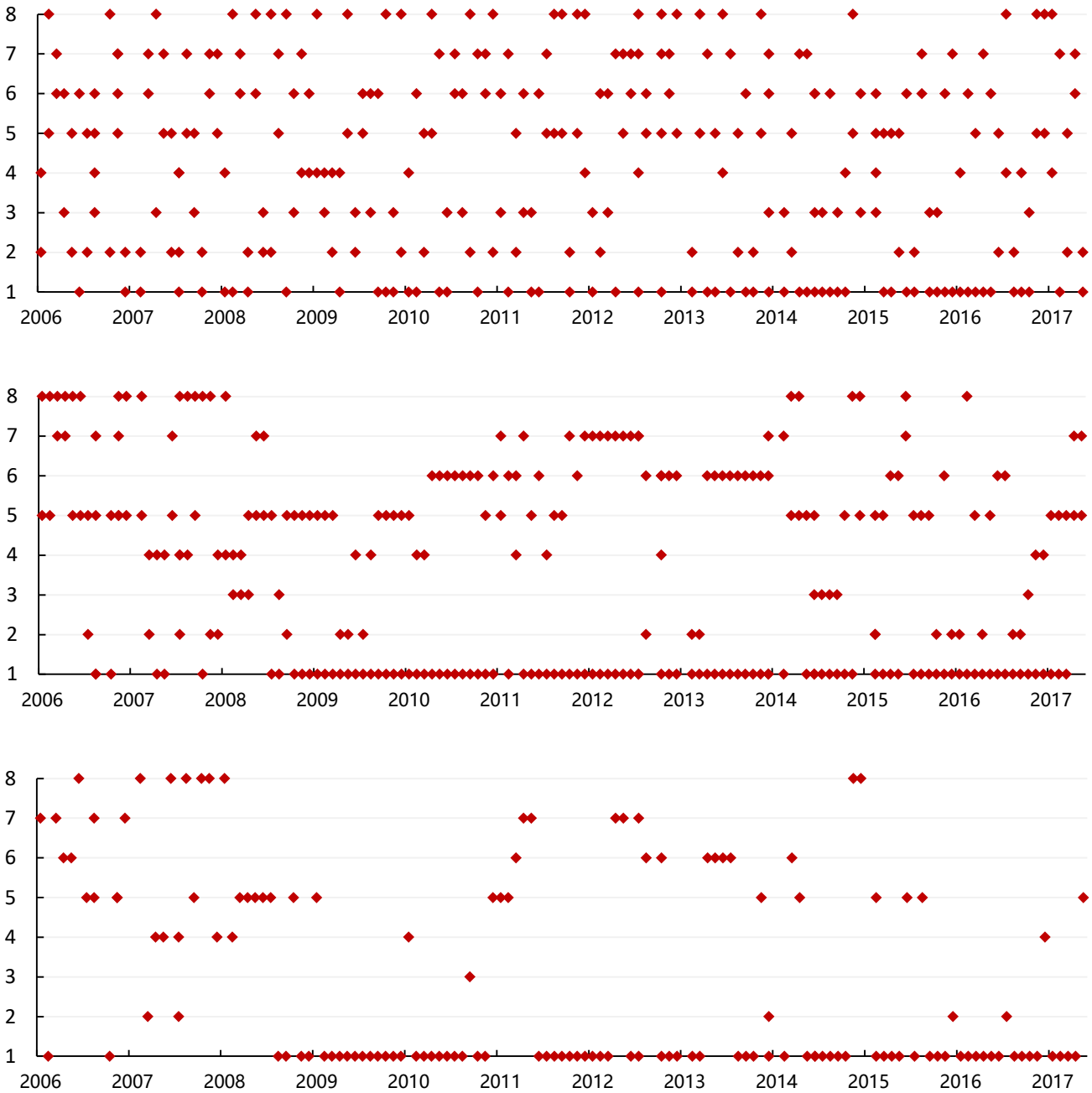
¹⁰ 单独选一种风格的绝对收益高于两种风格，但两种风格的稳健性高于一种风格，后文结果均基于仅选择一种风格

2.7 风格预测结果

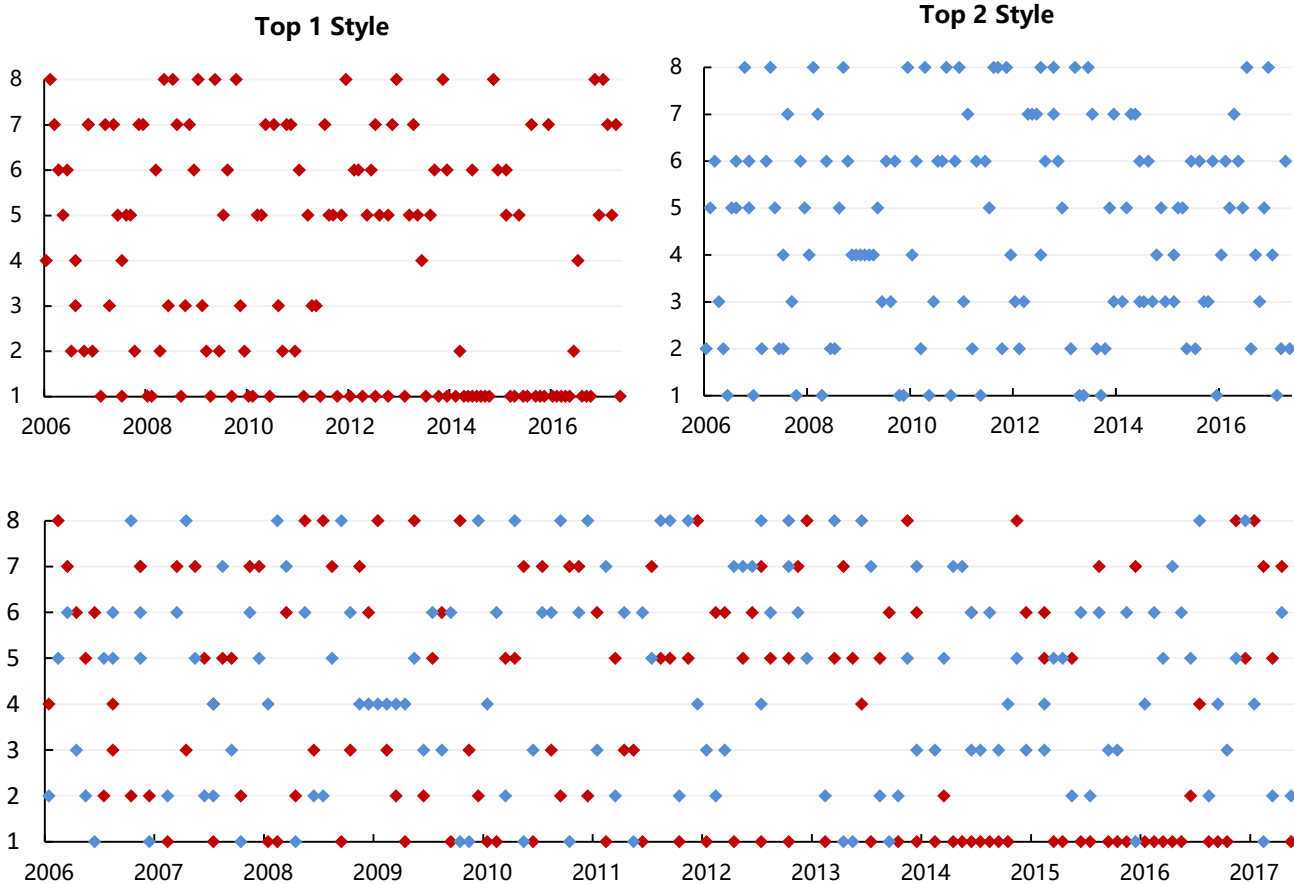
我们可以对比周频度模型预测结果和实际表现最好的两个市场风格，实际发生的市场风格虽然零散，转换也相对频繁。模型预测的结果延续性较强。下图分别依次为：实际发生的风格、预测 Top2 两种风格图、预测 Top1 一种风格图。



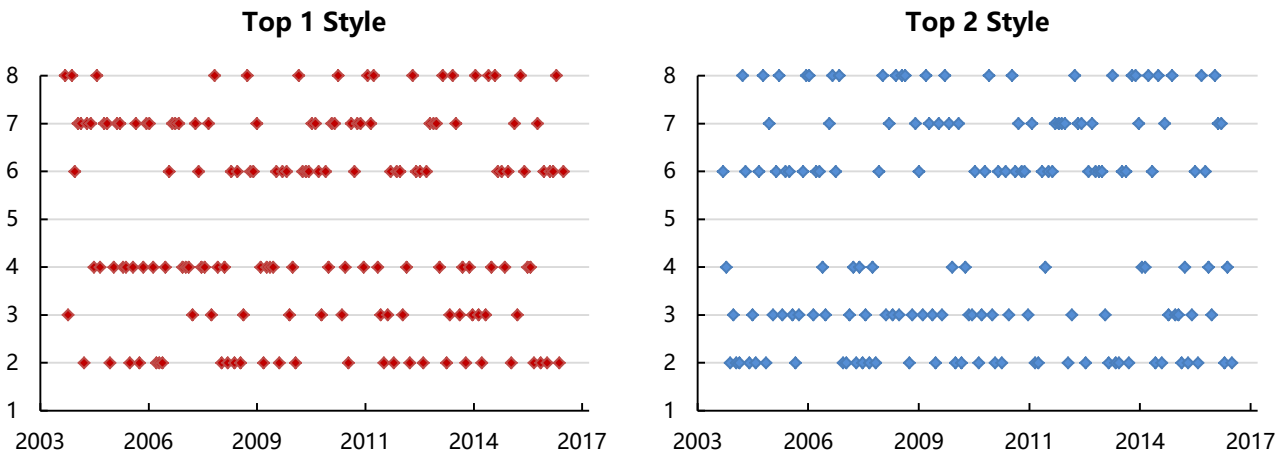
下图分别依次为：实际月度表现最好的两个风格、预测月度 Top2 两种风格图、预测月度 Top1 一种风格图。注意月度和周度模型结果并不简单等同于周和月之间的频度转换，而是两个不同的回归结果。

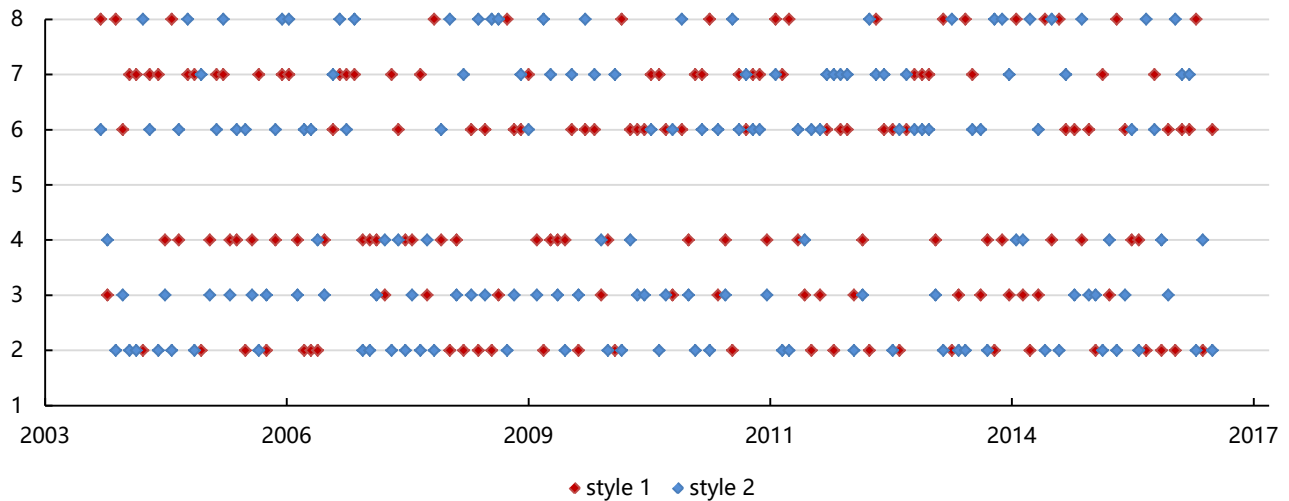


倘若分解上图实际发生的排名 1、2 的风格，我们可以发现 Size 风格在很长一段时间内都是持续排名第 1，趋势性非常强。风格 5、6、7 总体上也呈现出一定的趋势性。

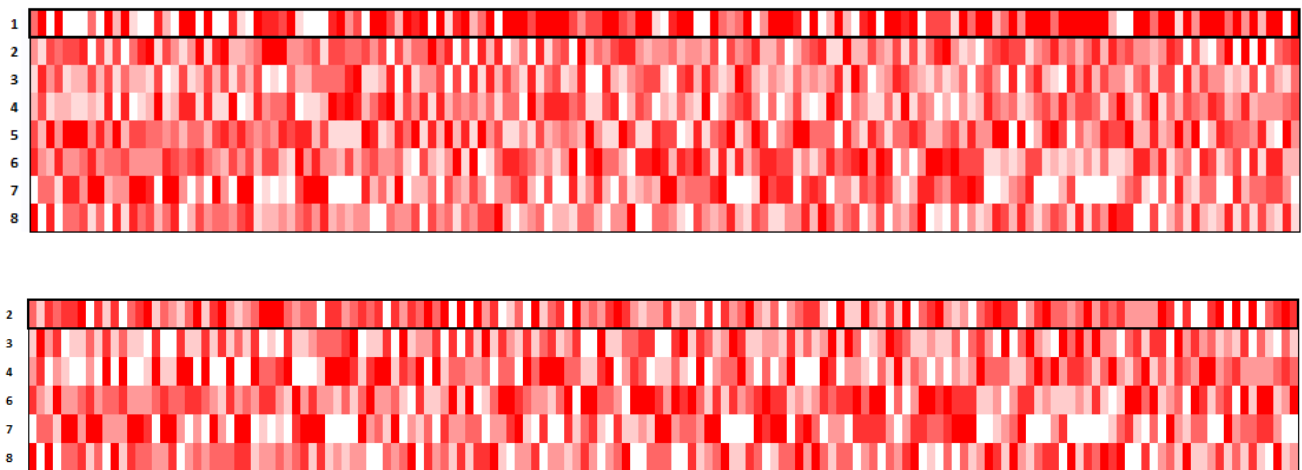


一个由市值因子所引发的担忧是：模型预测未来风格时趋势是一个重要输入变量，但在最强趋势项 size 缺席的情况下，市场其他风格是否还会存在趋势，模型中的趋势变量继续发挥作用？倘若剔除 8 种风格中趋势性较强的 1 (size) 和 5 (technical) ，是否整体风格会出现较大的噪音，趋势性变弱？剔除后重新排实际表现最好的第 1、2 名风格发现：趋势一定程度上依然存在于其他风格。

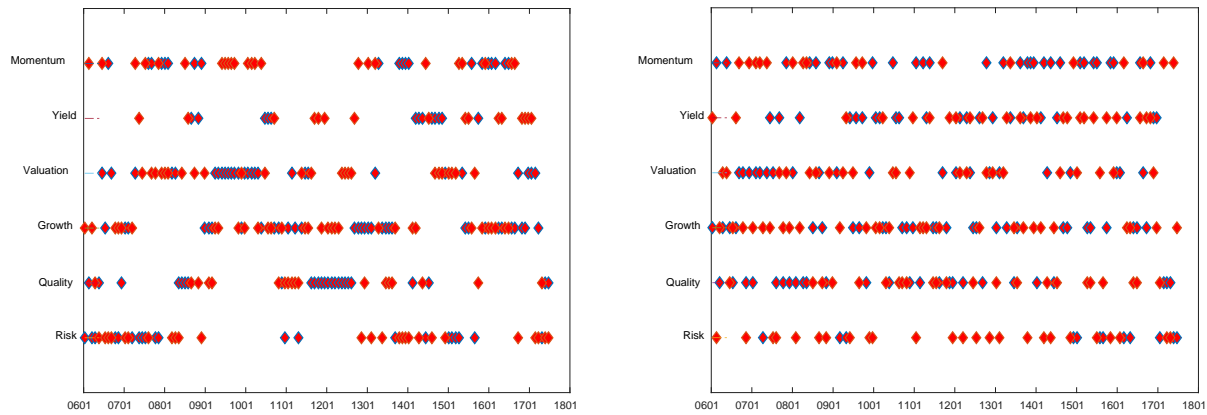




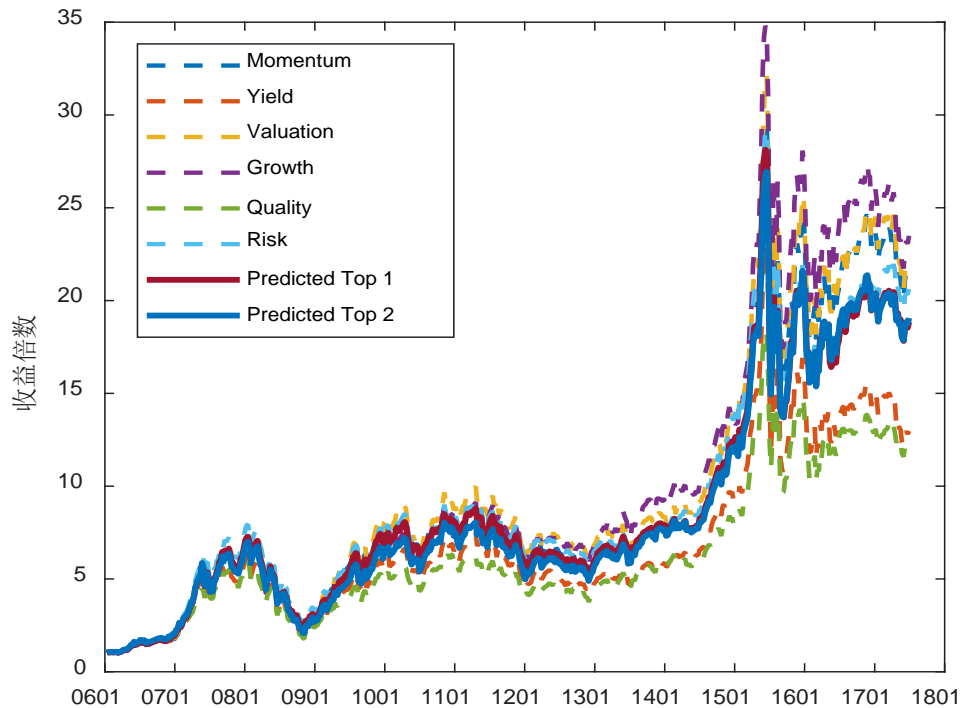
换一个角度，观察排名热力图（排名第 1 为红色，第 8 为白色，红色随排名靠后而变淡）发现：虽然前后两张图红白相间居多，但局部依然不乏红色连贯延续的风格。



剔除趋势性最强的 Size 和 Technique 风格，用最近 12 个月的历史样本信息做混合面板回归的预测风格，罗列所有双参数的排列组合，回测收益率并选取最近 6 个月累计收益率最高的组合 作为下一期的参数。两个参数的范围均以 1 个月为间隔选取 1 至 12 个月中的 12 个参数。下左图带蓝色边框的为排名预测未来收益率第一的风格和实际风格（下右图）之间依然存在一定的预测效力：

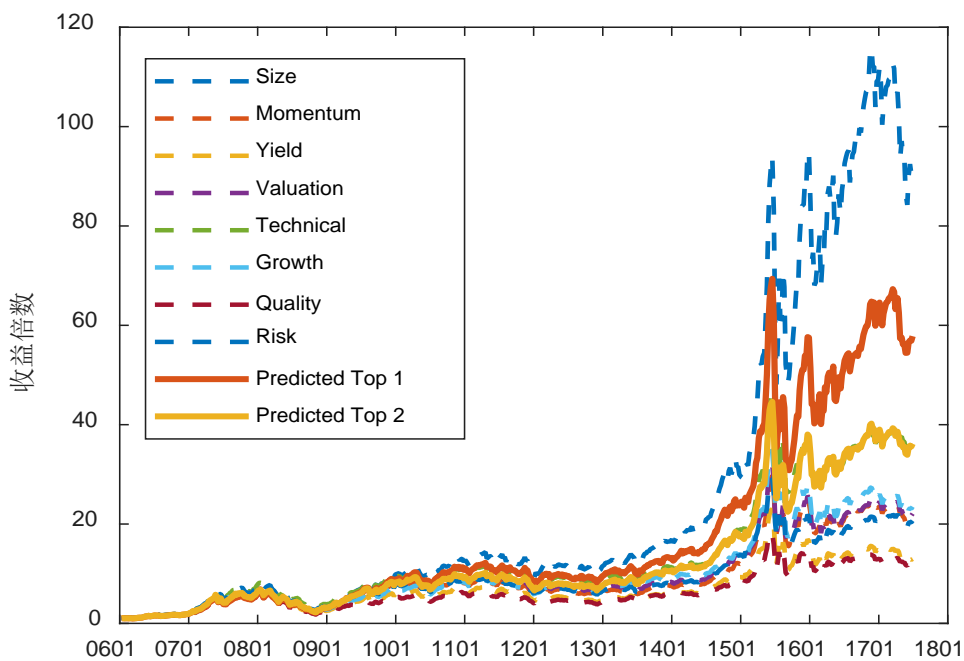


回测¹¹结果显示：缺少了强趋势性的风格 1 和 5 后，实际风格相对更为混沌、转换频繁，用模型预测有很大困难。模型的预测效果相比原来 8 种风格下不仅绝对值上大幅下降，押注排名第一的风格和同时押注排名前两位的风格的净值曲线仅好于两个单独风格。整体效果差强人意，这也说明后续需要进一步深入挖掘风格预测模型，有逻辑地提高预测正确率、寻找有效的因子分类器是整个策略的核心之一。



¹¹ 此处沿用上文所述的方法动态选择最优双参数押注排名前两位的风格，交易假设无手续费，且仅用 Top20%组平均收益率作为风格的收益率累乘得到，尚未涉及选股模型。不同于一般的回测，此处只是快速检验风格轮动的有效性。

原来 8 种风格下轮动长期领先除 Size 以外的其他单风格曲线。落后于 Size 风格的原因如前文所述，主要是限制了双参数上限为 12 个月，以提高模型的灵敏度。倘若采用扩大的历史窗口进行参数选择，则回测曲线能非常贴近甚至好于 Size 风格，但这样的结果不仅存在过拟问题，而且模型反应迟缓。



3. 选股模型

在计算出下月所要押注的风格后，需要进行选股，在选股前通用的规则为：

- (1) 剔除 ST 和 ST*；
- (2) 选股 30 个，再选股 70 个作为缓冲区间，共计 100 个；
- (3) 全部进行行业中性化选股，首先计算所选股票数量股票在全市场的占比（例如全市场当时有 N 个股票则比例为 $X=50/N$ ），接着在每个行业股票数上乘以 X ，计算出每个行业所需选出的股票数，每个行业至少选一支股票。加总所有行业股票，按照因子值从大到小取排名前 100 的股票，再按照因子排名大小取前 30 个股票，组合内可能存在有行业未被选中和行业比例失调的情况；
- (4) 除所有正常交易费用¹²外，每次交易单边滑点 0.3%，风格按周轮动，股票按周换仓；

然后，根据以下选股模型选股回测：

3.1 模型一：IC 打分法选股，等权组合

考虑到增加容错率，押注排名前两位的风格稳健性好于仅押注一种风格。因此我们选取排名前两名风格下的因子，按各因子周 IC 的移动 6 个月平均值的绝对值作为因子权重，归一化，对各股票在相应因子上的排名分值（越好分值越高）乘上相应权重，行业中性化选股后等权组合。注意，周 IC 的移动 6 个月平均值为负的因子在计算加权得分时，因子的分值应当倒过来¹³，得出综合得分。

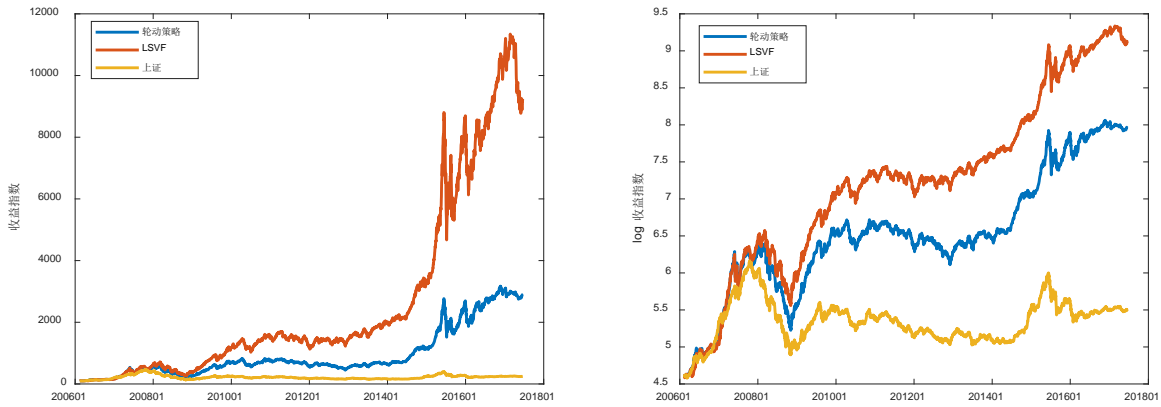
3.1.1 策略回测结果

组合的净值曲线如下，单边做多的曲线还是出现了 70%左右的巨大回撤。模型在因子艰难的 2011-2013 年期间和 2004-2006 期间未能产生较好的超额收益。

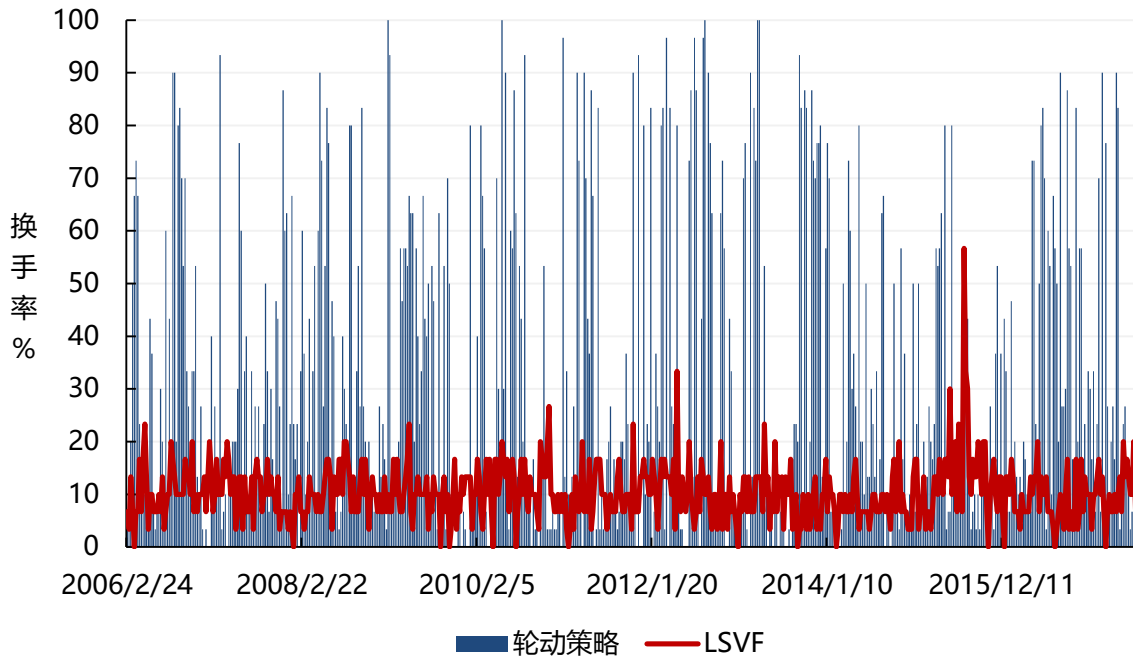
¹² 券商手续费设置为万三。

¹³ 例如原来市值越小 zscore 越大，市值 IC 为负时计算 zscore 时应该倒过来，市值越大 zscore 越大。

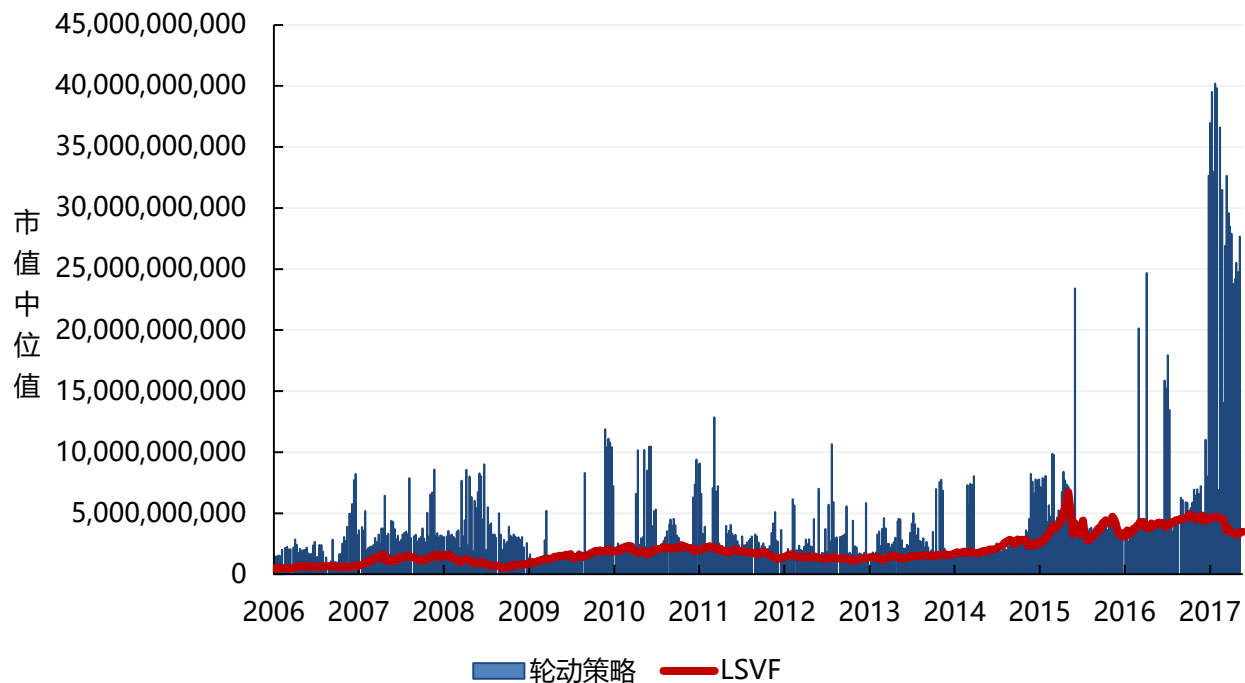
定点回测的绝对收益大幅落后小盘壳股为主的 LSVF（选股数目同为 30/100），由于曾持续押注小盘股，净值曲线和 LSVF 呈现强相关性。



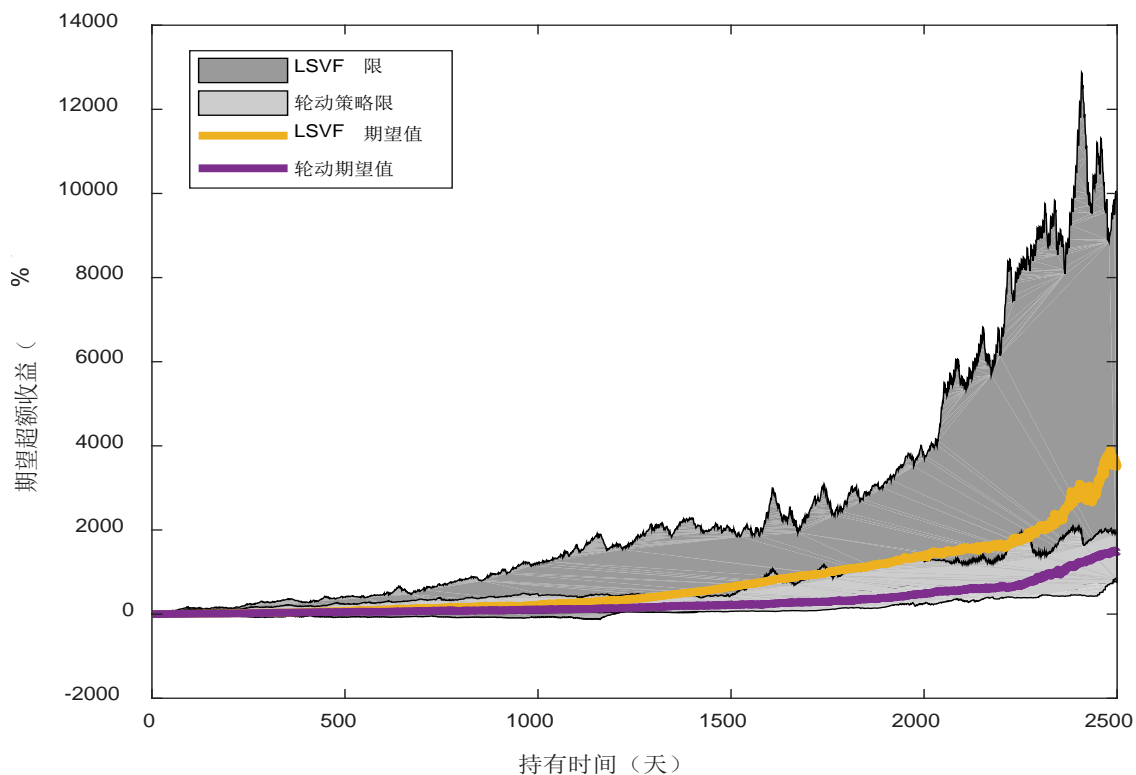
换手率方面由于风格轮换中往往出现较高的换手，平均周单边换手率高达 28%，因此策略的潜在风险点是风格持续快速变化，模型无法押注准确，徒增换手，收益被交易费用拖累。



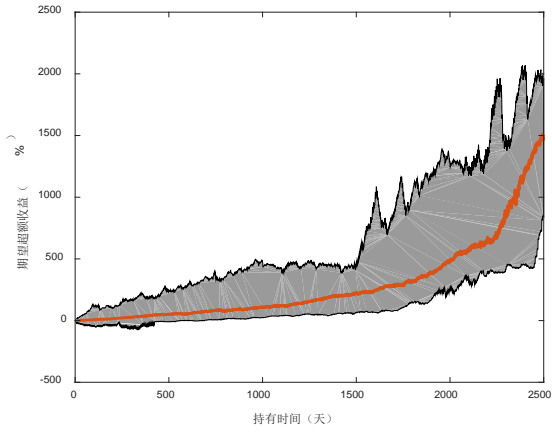
组合的市值方面自然也是随风格而变化，历史上的大部分时间由于选中的风格包含 size，因此市值中位数近似于小市值风格的 LSVF。另一方面由于同时押注两种风格，市值偏中盘股居多。



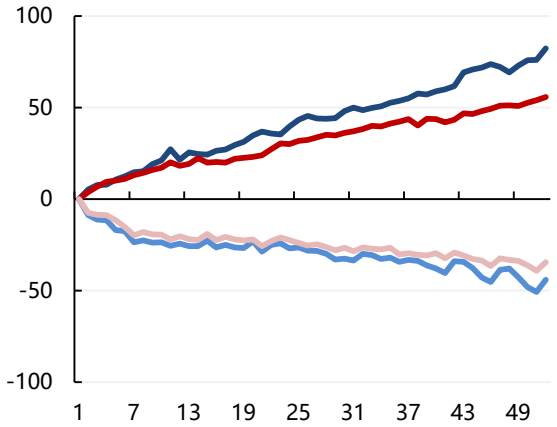
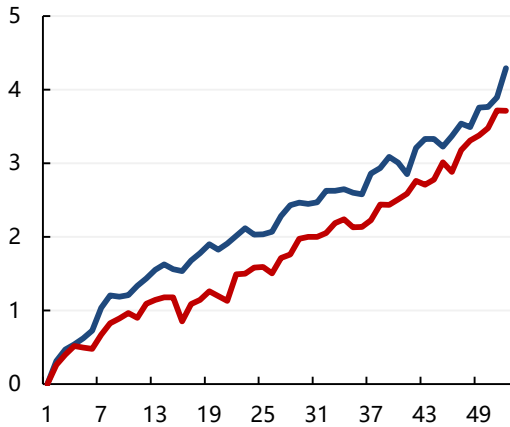
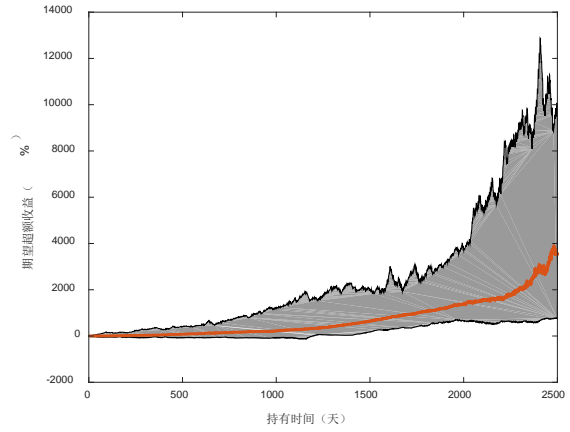
策略超额收益的期望值从长期角度远不及 LSVF。超额收益期望值的上下界来看 MFSR 是牺牲了上界来换取较高的下界，是如何平衡风险和收益的问题。策略日收益率分布图也印证了同样的事实。



MFSR



LSVF

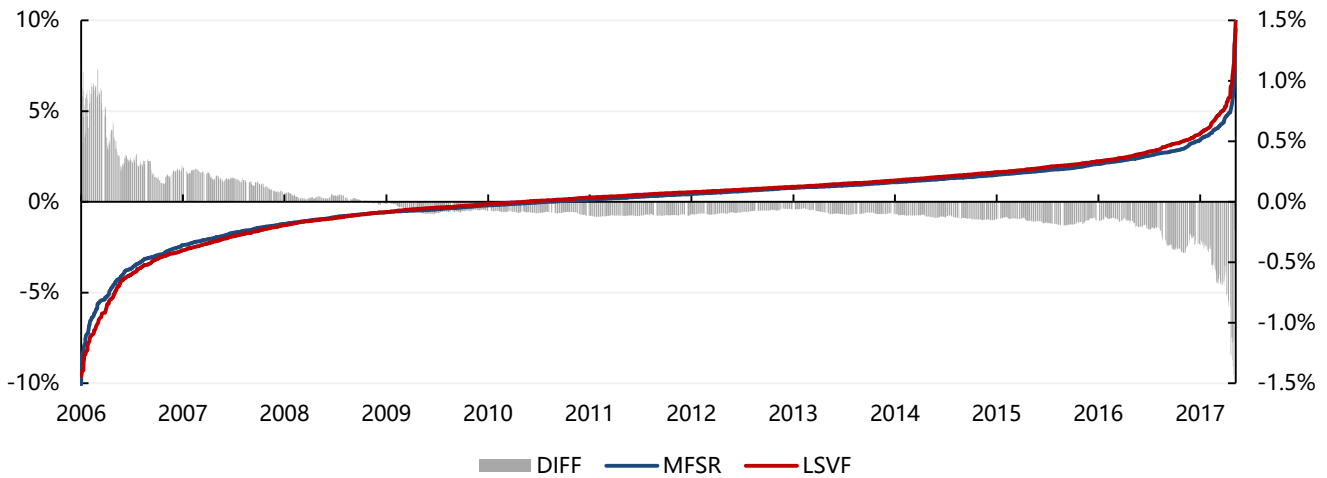


— LSVF超额收益期望值 — MFSR超额收益期望值

— LSVF-超额收益期望值上界 — MFSR-超额收益期望值上界

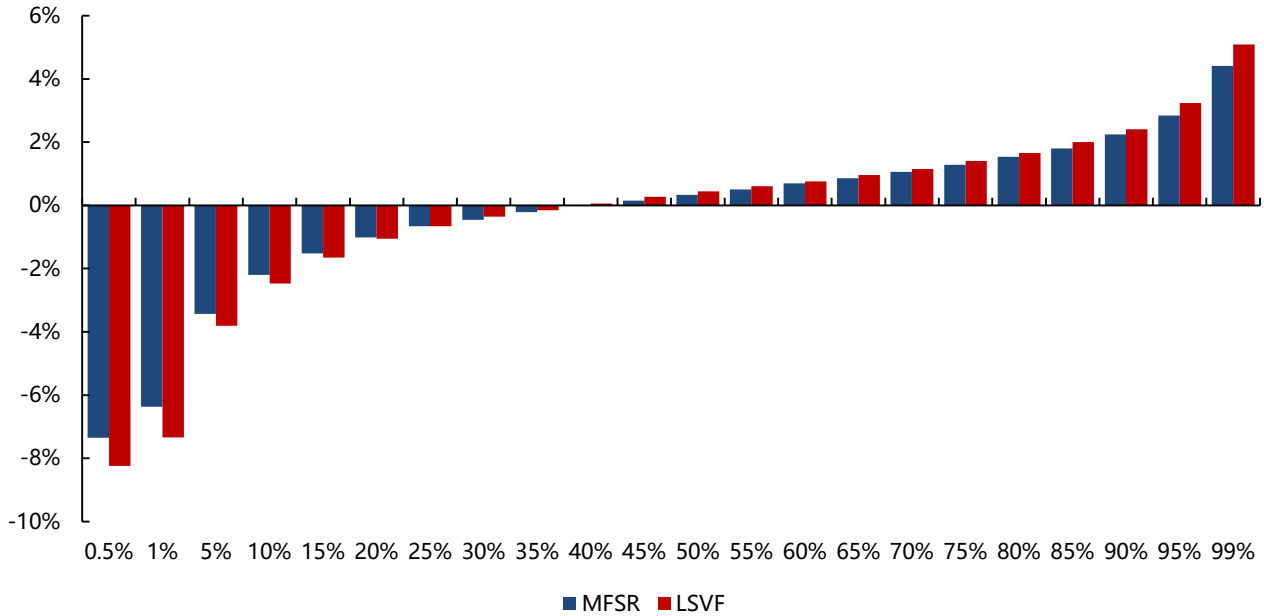
— LSVF-超额收益期望值下界 — MFSR-超额收益期望值下界

策略日收益率分布图



— DIFF — MFSR — LSVF

策略日收益率分位值



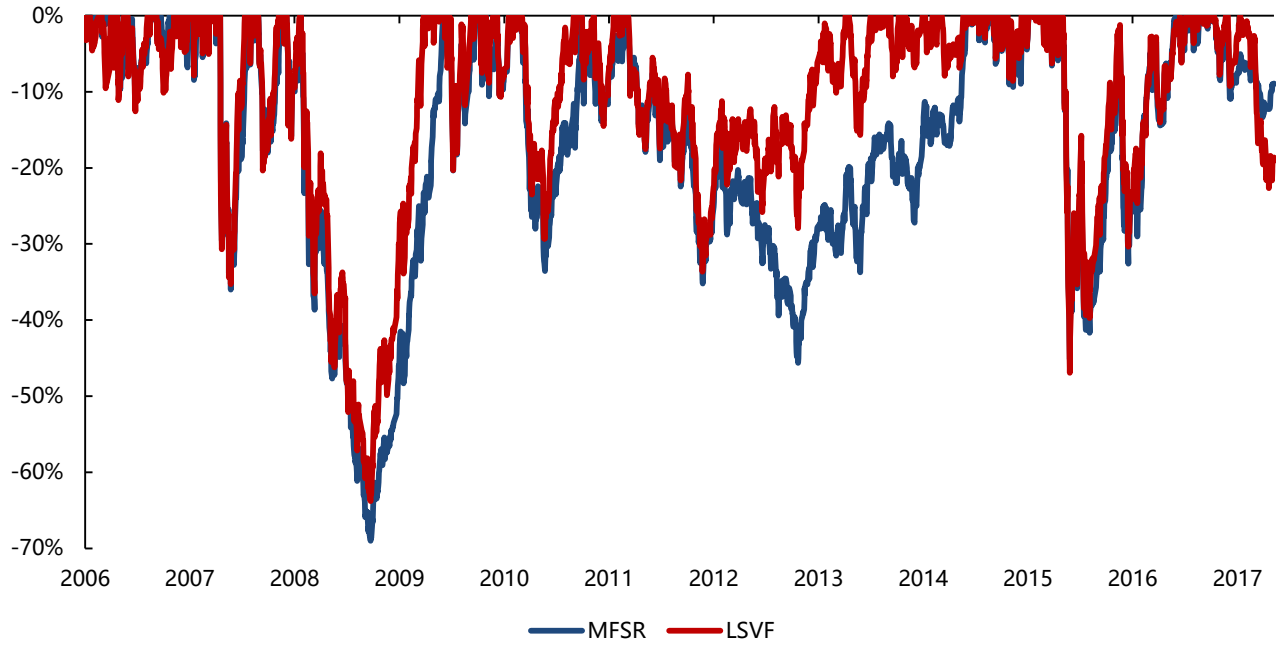
由于历史上的大部分时间由于选中的是 size，因此历史回测近似于小市值风格的 LSVF，相关系数较高。但近期小市值因子实效，MFSR 和 LSVF 走势的相关性开始变弱。

	MFSR 策略	LSVF 策略
收益相关系数 (2006-2016.10)		0.93
超额收益相关系数 (2006-2016.10)		0.78
收益相关系数 (2016.11-2017.7)		0.75
超额收益相关系数 (2016.11-2017.7)		0.49
第一大回撤	69.0%	63.8%
夏普比率	1.14	1.37
Beta	0.93	1.03
下跌波动率 ¹⁴	1.7%	2.1%

MFSR 的最大回撤也较 LSVF 更大。综合以上历史回测结果的比较，MFSR 完败 LSVF。我们做的过程中力求从逻辑本身出发，不刻意提高回测结果，牺牲一定的收益来换取模型的灵敏度和控制下跌风险。但也要思考：各种指标都差的情况下我们是否应该使用更动态，逻辑可能更好些，但历史回测结果明显更差的策略。

¹⁴ 策略下跌日涨跌幅的标准差

最大回撤



4. 择时策略

4.1 多空信号

由于 MFSR 多因子风格轮动策略 V 1.0 的选股涵盖面较广，因此使用单一指数择时的代表性较差应当综合中短期各类指数，从各个角度进行多空判断。

信号一：首先取全市场 A 股，剔除退市、停牌的、上市日期小于 20 日的，其次用全市场流通市值最小的 20% 股票的日涨跌幅等权生成一个小市值指数的日涨跌幅，以 5 日指数移动平均（EMA5），和 EMA10 的交叉作为多空信号。EMA5>EMA10 时多，否则为空。

信号二：用沪深 300 指数 EMA10 作为多空信号，当日指数>EMA10 时多，否则为空。

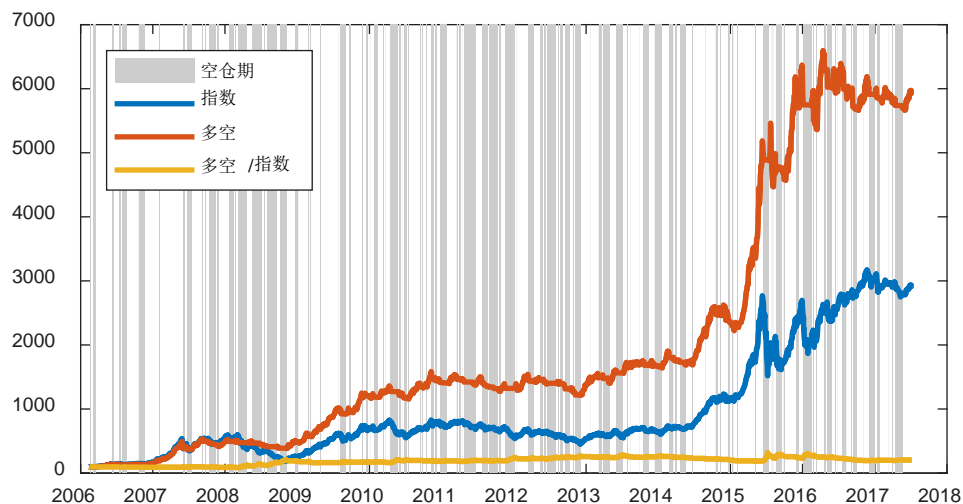
信号三：用中证 500 指数 MA20 作为多空信号，当日指数>MA20 时多，否则为空。

信号四：用中证 1000 指数 EMA10 作为多空信号，当日指数>EMA10 时多，否则为空。

信号五：用上证综指 MA20 作为多空信号，当日指数>MA20 时多，否则为空。

信号六：用策略本身净值的 LLT 作为多空信号，当日策略净值>LLT 时多，否则为空。

当以上 6 个信号有 2 个以上为多时，整体为多信号，否则为空。也就是有 0/1/2 个多信号时策略整体为空仓信号。



	MFSR 策略 - 无择时	MFSR 策略 - 择时
胜率	54.6%	N/A
赔率 ¹⁵	1.95	N/A
频率	19.4%	N/A
闪烁信号占比 ¹⁶	43.4%	N/A
空信号次数	143	N/A
最大回撤	27.9%	69.0%
夏普比率	1.80	1.14
下跌波动率 ¹⁷	1.5%	1.7%

¹⁵ 可以行生成空仓阶段净值与倘若不空仓同期净值之比，这里是所有日期比值的中位值来表示

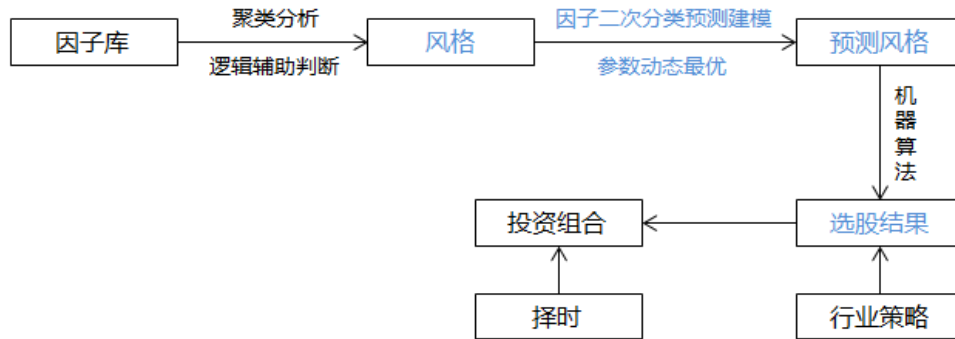
¹⁶ 持续时间小于 3 天的信号占总信号数量之比

¹⁷ 策略下跌日涨跌幅的标准差

5. 总结

5.1 策略的基本逻辑是什么？

通过归类出簇内相关性较大、群间相关性较小的因子簇，然后运用多因子模型预测未来可能起作用的因子簇，最后根据预测的因子簇选股。多因子模型分别前后应用于因子簇预测层面，和选股层面。



5.2 策略的逻辑漏洞和风险点？

当多因子预测模型中的解释变量不能涵盖未来因子变动的主因时，预测效力较差。当市场风格混沌，频繁变动时模型失效，在频繁错误轮动风格中突增换手率和交易费用。当因子本身和因子之间的相关性发生突变时，基于历史信息的聚类可能出错，导致错在起点。最后，择时和行业策略可能影响策略本身的有效性。

5.3 期望收益从策略风格和风险因子角度如何归因？是否有合适的分散？

目前多因子预测模型中的解释变量主要涵盖价值和趋势这两个策略风格，这取决于预测模型中的解释变量。因子类内的因子都是风险因子，但并没有静态押注风险因子组合，模型至少是动态的。体现在：聚类、预测模型、参数优化、机器学习配置因子权重等诸多方面。静态主要曝光在行业 and 择时策略。

策略风格上目前在价值和趋势两个常规方向上分散，因子类上的分散体现在取排名靠前的两个因子类，因子层面的分散体现在每个因子类下有多个因子，股票层面分散体现为持股数较多。

5.4 策略在极端情况下的表现如何？

从历史的极端情况来看，最大回撤控制在 30%以内，多数情况下在 20%以内。

5.5 超额收益来自于哪里？

押注的因子的 alpha 和押注对风格因子所带来的 alpha，还有择时所带来的超额收益。

5.6 未来出现何种情况应果断停止策略的使用？

当业绩归因时未能解释部分的损失超过一定阈值时应当重新审视模型，当个体效应引起的业绩影响超过一定阈值时选股模型可能存在问题。

6. 待优化项目

1. 嵌套行业策略，具体参看 6.14 说明文档。
2. 尝试做下 PLS，WLS，固定效应，随机效应等方法，系数归 0 不归 0 两个选项。
3. 换股调仓频率改月度，以降低换手率。
4. 缩短时间窗口然后计算，查看下 06 年以前结果。
5. 神经网络/Adaboost 因子归类选股。
6. 神经网络选风格。
7. 根据条件动态筛选每一期的有效因子集合，在集合内做聚类划分（过去每个时间点上做二维聚类，因子 IC 和 IC 波动率，然后计算时间序列上的相关系数来聚类？根据历史 IC 时间序列的相关性来划分？具体我不太懂，有待研究）这样设计的最终目的是避免人为静态地定义风格，挑选因子风格和因子本身具有一定的动态性。

尝试以下因子：

- (1) IC 偏离度： $R = (\text{因子 Top20\%组合的 20 累计收益率} - \text{大盘 20 日累计收益率})$ ， $DEV = R$ 的 12 个月移动窗口 zscore，用 IC 偏离度来排名，然后做 IC 图
- (2) 自回归滞后 n 阶项，即因子 t-n 期排名，按照排名予以 1-8 的值或排名前两位予以 1 值，否则为 0 值。
- (3) $IR = \text{各因子的 IC 平均} / \text{IC 的标准差}$ ，IC 用周的，平均和标准差窗口用 12 个月
- (4) 涨跌幅类：沪深 300 指数、中证 500 指数、全 A 指数前 1 个月/3 个月涨跌幅
- (5) 波动率类：沪深 300 指数、中证 500 指数、全 A 指数前 1 个月/3 个月日收益波动率、全 A 指数与沪深 300 指数 1 个月/3 个月波动之差、中证 500 指数与沪深 300 指数 1 个月/3 个月波动之差；
- (6) 换手率类：沪深 300 指数前 1 个月/3 个月日均换手、全 A 前 1 个月/3 个月日均换手；
- (7) 利率类：SHIBOR 1M

8. 选股方法打分时拟合组内非线性，考虑因子的非线性特征，表格 D 中黑色代码的因子具有非线性。
9. 合成一个指数，全市场所有股票周平均涨跌幅是否小于-5%，-5%到 5%，5%以上为虚拟变量。这样做主要是：全市场更具有代表性，防止例如目前这样指数不跌，个股普跌的情况。其次-5%，5%这样能划分出 3 个区间，衡量市场处于较严重下跌，正常，和普涨的 3 种情况。
10. 用上下 20%分位的收益率之差作为 Y，选风格，然后结合判断收益来源到底是 top 还是 bottom，因子 IC 加权、反向排序选股。
11. 尝试 A-C 的因子库