

深度报告

金融工程

数量化投资

数量化投资系列之二十六

2010年5月5日

本报告的独到之处

■衡量因子强弱的两种测度：因子偏离度和因子贡献度

■FACTOR-SCORE 模型和多因子半衰期选股策略

■全样本多因子 Alpha 选股

■多因子增强的正 Alpha 行业配置策略

专题报告

多因子 Alpha 选股——将行业轮动落实到 Top 组合

在去年的报告《国信证券-数量化投资系列之七：基于 Alpha 的行业配置方法和投资策略——总是获得正 Alpha》中，我们通过正 Alpha 行业配置策略将行业分为长期 Alpha 距离最远的两类，且通过穷举验证，至今为止仍然战胜数十万种其余的行业组合分类方法。

从长期来看，Alpha 因子作为一种行业因子划分具有两端距离最远的特性。但投资者总是会关注不同的因子，因此如果单看一个较短的时期，则有可能存在动态变化的最强因子。如果能在每个时点都能通过选股捕捉到当时的最强因子组合，继而选择能从最强因子获得正向超额收益的股票组合，最终累计的超额收益应该能超过单纯的 Alpha 因子的效果。

正是基于以上的思路，我们决定尝试进行因子强度和多因子轮动的研究。为了衡量不同时期各个因子的强弱程度，构造了两种测度：**因子偏离度 DEV** 和 **因子贡献度 CON**。这是两个标准化的测度，没有量纲，取值在 $[-1, 1]$ 区间内。后文论证了两种度量方法的一致性。

利用因子贡献度，我们挑选了 MV、ROE 等 9 个因子，回溯了以中证 800 成分股为股票池的因子轮动历史路径，并得出相应的因子贡献度矩阵。回溯结果表明，因子轮动效应没有参数依赖性。

接下来我们构建了 **FACTOR-SCORE** 模型，对股票池的每只股票动态打分。根据打分排名结果，构建了**多因子半衰期选股策略**。历史样本检验表明即使选取不同的时间周期参数 T，该策略都能获得持续的超额收益。

进一步，我们将 **FACTOR-SCORE** 模型与**正 Alpha 行业配置策略**相结合，设计了多因子增强的新 Alpha 策略。在 4 年半的历史样本区间内，相对中证 800 只是 310.75% 的累计收益率，在 Alpha 行业配置和多因子选股（其中多因子增强权重为 15%）的双重增强下，新 Alpha 策略在相同区间获得了 413.21% 的累计收益率。

最后总结本篇报告的模型和衍生的两个量化策略，它们具有如下三大特点：

- (1) **良好的扩展性**：首先，不仅是文中所列的 9 个因子，对任意因子，甚至任意一种排名规则，都可以纳入我们的模型框架之内；其次，不仅是中证 800 成分股，对任意的股票池，都能计算对应的因子贡献度及其分布；最后，对任意已有的策略，只要是基于股票组合的操作，都能进行多因子增强。
- (2) **无参数依赖性**：我们的策略有效性并不依赖于时间周期参数 T，其本质是市场存在真实的因子驱动效应，而不是统计游戏的结果。
- (3) **非预测性**：我们的策略完全基于已有的数据，目标在于跟踪市场热点因子，不需要对因子的强度进行任何预测。

分析师 董艺婷

电话 021- 60933155

E-mail dongyt@guosen.com.cn

SAC 执业证书编号：S0980210030026

分析师 葛新元

电话 0755-82133332

E-mail: gexy@guosen.com.cn

SAC 执业证书编号：S0980200010107

独立性声明：

作者保证报告所采用的数据均来自合规渠道，分析逻辑基于本人的职业理解，通过合理判断并得出结论，力求客观、公正，结论不受任何第三方的授意、影响，特此声明。

内容目录

前言	4
研究框架	4
问题的提出	5
因子强度的度量	5
两种度量方法	5
量化模型	5
贡献度与偏离度	7
历史数据研究	7
数据来源	7
因子轮动历史规律	8
因子贡献度分布矩阵	10
不同参数周期下的因子轮动路径	12
因子轮动选股策略	13
FACTOR-SCORE 模型	13
多因子半衰期选股策略	13
策略实例	14
基于 Alpha 分类的多因子增强策略	16
因子贡献度模型的扩展应用	16
一二类行业的多因子增强	17
增强后的新正 Alpha 行业配置策略	18
结语	20
模型与量化策略的总结	20
后续研究方向	20

图表目录

图 1: 报告研究框架	4
表 1: 因子列表	7
图 2: 中证 800 成分股 MV 因子偏离度和贡献度比较 (T=60)	8
图 3: 中证 800 成分股 GX 因子偏离度和贡献度比较 (T=60)	9
图 4: 中证 800 成分股因子偏离度 (T=60)	9
图 5: 中证 800 成分股因子贡献度 (T=60)	10
表 2: 中证 800 成分股因子贡献度分布矩阵 (T=60)	10
表 3: 中证 800 成分股因子贡献度分布矩阵百分比形式 (T=60)	11
图 6: 中证 800 成分股因子贡献度 (T=20、30、40、50、60)	12
图 7: 多因子半衰期选股策略与中证 800 指数累积收益率对比	15
表 4: 多因子半衰期选股策略统计表	15
图 8: 多因子半衰期选股策略与中证 800 指数月度收益率对比	16
图 9: 一类行业、二类行业因子贡献度 (T=60)	17
图 10: 增强前后的一类行业收益率 (T=30, 15%配置)	18
图 11: 增强前后的二类行业收益率 (T=30, 15%配置)	18
图 12: 新一类行业和新二类行业 alpha 轮动图	18
图 13: 增强前后两类行业组合 Alpha 轨迹 (30 日)	18
图 14: 多因子增强后的正 Alpha 行业配置策略与中证 800 指数累计收 益率对比	19
图 15: 后续研究扩展示意图	21

前言

研究框架

本篇报告的研究框架如下图所示，首先在前言中提出我们研究的初衷：希望能找到不同时期市场上起主要作用的因子，即寻找因子轮动的最优路径。如果这一路径能够根据外生变量进行预测和选择，那么就是一个标准的投资时钟，反之，就需要一个不依赖外生变量和参数的组合构建方法。

为了衡量不同时期各个因子的强弱程度，我们构造了两种测度：因子偏离度 DEV 和因子贡献度 CON。这是两个标准化的测度，没有量纲，取值在 $[-1, 1]$ 区间内。最后论证了两种度量方法的一致性。

接下来，利用因子贡献度，以中证 800 成分股为样本，我们选择了 MV、ROE 等 9 个因子，检验了 2005 年至 2010 的因子轮动历史路径，并得到因子贡献度分布矩阵。从历史表现看，市值因子 MV 在 45% 的交易日中都是最强因子，而股息率因子 GX 则一直表现较弱。

在因子贡献度的基础上，我们设计了 FACTOR-SCORE 模型，用因子贡献度向量和股票因子排名向量的内积作为选定股票特定日期的分数。加上半衰期的换仓判定规则，便构成了我们的**多因子半衰期选股策略**。用中证 800 成分股作为股票池的历史回溯检验中，该策略从 07 年开始能较稳定的战胜中证 800 指数的收益率，且无参数依赖性。（06 年由于我们剔除了事件因子的冲击，导致股改收益未能获取）

在第二个应用实例中，我们构建了**多因子增强的正 Alpha 行业配置策略**，利用 Alpha 轮动选出两类行业中的超配类，并按照 15% 的比例增强，利用因子打分来赋予组合内股票新的权重。历史回溯检验表明相对中证 800 指数 310.75% 的区间累计收益，增强后的 Alpha 策略获得了约 100 个基点的超额收益。

图 1：报告研究框架



资料来源：国信证券经济研究所

最后的结语部分，我们总结了模型和策略在扩展性、无参数依赖性、非预测性等特点，并简单介绍了后续的研究计划。

问题的提出

在去年的报告《国信证券-量化投资系列之七：基于 Alpha 的行业配置方法和投资策略——总是获得正 Alpha》中，我们通过正 Alpha 行业配置策略将行业分为长期 Alpha 距离最远的两类，且通过穷举验证，至今为止仍然战胜数十万种其余的行业组合分类方法。

从长期来看，Alpha 因子作为一种行业因子划分具有两端距离最远的特性。但投资者总是会关注不同的因子，有时市场热点由市值因素催生，有时又一致关心财务指标；这个月行业特性最突出，下个月也许地域板块区分度最高。因此如果单看一个较短的时期，理论上总存在一个动态变化的最强因子。如果能在每个时点都捕捉到当时的最强因子，继而选择能从最强因子获得正向超额收益的股票组合，最终累计的超额收益应该能超过单纯的行业 alpha 轮动的效果。

进一步，通过对最强因子历史路径的分析总结，如果能找到驱动这种因子轮动效应的外生变量，则我们可以通过对外生变量的判断去预测未来一段时间的最强因子，从而提前布局。

正是基于以上的思路，我们决定尝试进行因子强度和多因子轮动的研究。那么一个很自然的问题就是：究竟用什么标准来判断一个因子的强弱？这种强弱程度最好能够用一个明确的量来表示，而且它既能满足因子横向比较的要求，又能满足时间纵向比较的需要。下面因子强度的度量，就是首先解决这个问题。

因子强度的度量

两种度量方法

在任一时间点，如何衡量某个因子的强度，我们有两种思路：

思路 1（收益率-分布）

首先将股票池中的股票按过去 T 天的对数收益率进行排序，选出收益率最高的 20% 和收益率最低的 20% 股票构成两个组合。然后看每个组合对某个因子的分布，分布相差越大，则说明该时点此因子的强度越高。

思路 2（分布-收益率）

将股票池中的股票按因子进行排名，分别选出排名靠前的 20% 和排名靠后的 20% 股票构成两个组合。两个组合的收益率相差越大，则说明该时点此因子的强度越高。

量化模型

按照上面的两种思路，下面给出具体的量化模型

DEV Model

设当前时间为 t ，考察周期为 T ，股票池总数为 N ， $n = \text{int}[N/5]$ 为 N 的 20%

取整值，股票 i 的对数收益率为 R_i ，过去 T 天收益率最高的 20% 股票组合为

SET_{high-R} ，收益率最低的 20% 股票组合为 SET_{low-R} 。

对于因子 F ，将 N 支股票依据 $t-T$ 时刻的因子值由小到大进行排名，第 i 只股票的排名记为 $rank(i)$ 。定义因子 F 的偏离度 DEV_F 为

$$DEV_F = \frac{\frac{\text{mean}[rank(i)] - \text{mean}[rank(j)]}{N - n}}{\frac{\sum_{i \in SET_{low-R}} rank(i) - \sum_{j \in SET_{high-R}} rank(j)}{n * (N - n)}}$$

DEV_F 位于区间 $[-1, 1]$ 内，其绝对值越大反映因子强度越高，当绝对值为 1 时表示收益率排名的两端恰好也是因子值 F 排名的两端； DEV_F 为正表示高收益率组合 SET_{high-R} 中的股票关于因子 F 的排名分布集中在较低区域，为负则表示关于因子 F 的排名分布集中在较高区域。

CON Model

设当前时间为 t ，考察周期为 T ，股票池总数为 N ， $n = \text{int}[N/5]$ 为 N 的 20% 取整值，股票 i 的对数收益率为 R_i ，过去 T 天收益率最高的 20% 股票组合为 SET_{high-R} ，收益率最低的 20% 股票组合为 SET_{low-R} 。

对于因子 F ，将 N 支股票依据 $t-T$ 时刻的因子值由小到大进行排名，因子值最大的 20% 股票组合为 SET_{high-F} ，因子值最小的 20% 股票组合为 SET_{low-F} 。定义因子 F 的贡献度 CON_F 为

$$CON_F = \frac{\frac{\text{mean}[R_i] - \text{mean}[R_j]}{\sum_{i \in SET_{low-F}} R_i - \sum_{j \in SET_{high-F}} R_j}}{\frac{\text{mean}[R_i] - \text{mean}[R_j]}{\sum_{i \in SET_{high-R}} R_i - \sum_{j \in SET_{low-R}} R_j}}$$

CON_F 位于区间 $[-1, 1]$ 内，分母表示过去 T 日最大的收益差，分子表示因子 F 带来的收益差，其绝对值越大反映因子强度越高，当绝对值为 1 时表示因子值 F 排名的两端恰好也是收益率排名的两端； CON_F 为正表示因子值 F 较小的股票组合的平均收益较高，为负则表示因子值 F 较大的股票组合平均收益较高。

贡献度与偏离度

从上面的模型可以看到，偏离度和贡献度都是经过标准化，没有量纲的量。进一步，从统计上而言，如果因子和收益差没有因果关系，两个度量值应该是非相关的。但是从后面的实证数据可以看出，两者正相关，说明因子确实是导致收益差的原因之一。

历史数据研究

数据来源

参数选择

股票池选择中证 800 指数历史成分股， $N=798\sim 800$ （2006 年 3 月 10 日~6 月 30 日为 799 只，2009 年 12 月 28 日~31 日为 798 只，其他时间段均为 800 只）。

时间周期 $T=60$ ，单位为交易日。时间区间为 2005 年 4 月 7 日~2010 年 4 月 9 日，共 1219 个交易日。

因子选择

本文选择了市值、换手率、机构持股、资产回报率、账面市值比、市销率、市现率、股息率和换手率变化等 9 个因子。

表 1：因子列表

因子名称	缩写	因子值计算方法
市值	MV	指定日总市值
换手率	TURN	60 日平均换手率
机构持股	INS	机构持股占流通股比
资产回报率	ROE	扣除非经常损益后的净资产收益率
账面市值比	BM	股东权益合计/总市值
市销率	PS	总市值/最近 12 个月主营收入
市现率	PCF	总市值/最近 12 个月经营活动产生的现金流量净额
股息率	GX	每股股利/每股收入
换手率变化	TURN2	30 日平均换手率/60 日平均换手率

资料来源：国信证券经济研究所

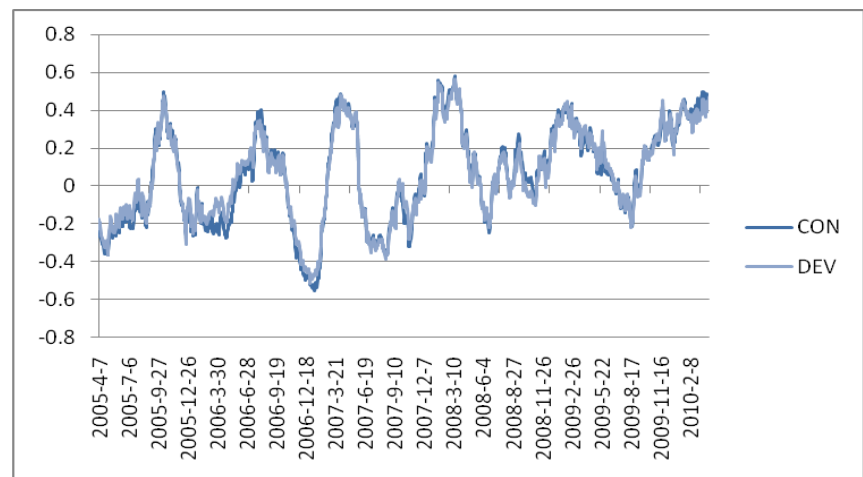
因子轮动历史规律

MV 因子和 GX 因子的例子

首先我们来看两种度量方法在历史样本中的表现。下面两张图分别是 MV 因子和 GX 因子的偏离度和贡献度比较。从图形上可以得到如下结论：

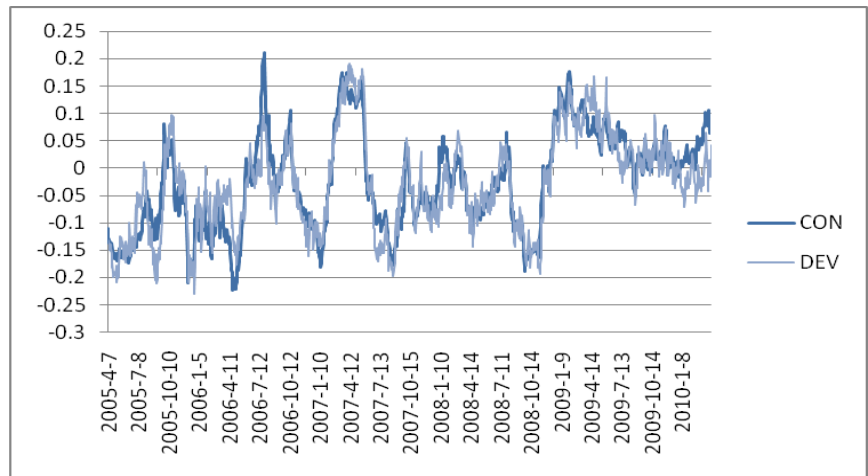
- (1) 偏离度 DEV 和贡献度 CON 的值明显正相关。其中较强的 MV 因子从图形上来看两种度量基本重合，相关系数为 0.9898，一阶变差后两者的相关系数为 0.8069；较弱的 GX 因子的两种度量也有相同的趋势，相关系数为 0.9078，而一阶变差后两者的相关系数为 0.5214。这说明我们定义的两度度量方法互相印证，衡量因子的强弱具有一致性；因子越强，则两种度量方法的结果越一致。
- (2) MV 的因子强度要显著高于 GX 因子，这表示历史上市场对大小盘的关注度远远高于对股息率的关注度。虽然从理论上而言，股息率是衡量企业是否具有投资价值的重要标尺之一，但是 A 股市场存在大量“铁公鸡”式的上市公司，多年一毛不拔依然受到投资者的追捧。
- (3) 从 MV 因子的偏离度和贡献度走势来看，具有一定的周期性，小市值股票相对大市值股票能定期取得正的收益差。而大市值股票只在 2006 年底和 2007 年三季度有明显的正收益差。此外自 2009 年 8 月以来 CON 和 DEV 值一直为正，即小市值股票持续强于大市值股票，且时间长度超过了历史上的任何一次。
- (4) 从直观上来看，MV 因子的强弱周期与股市周期以及宏观经济周期的关系并不显著，而是独立存在的周期。我们试图寻找驱动因子的外生变量，但是没能得到有效解释。

图 2：中证 800 成分股 MV 因子偏离度和贡献度比较 (T=60)



资料来源：国信证券经济研究所

图 3：中证 800 成分股 GX 因子偏离度和贡献度比较 (T=60)



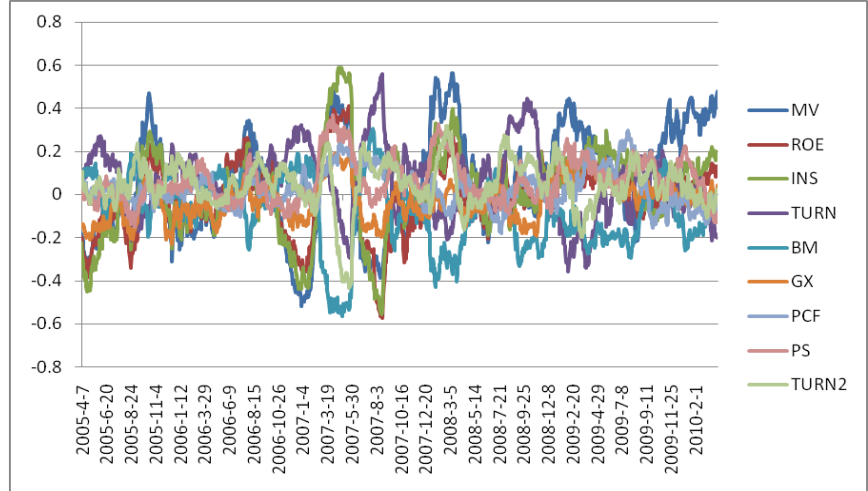
资料来源：国信证券经济研究所

因子轮动表现

看完 MV 因子和 GX 因子的例子之后，我们研究全部 9 个因子的历史表现，如下面两张图所示：

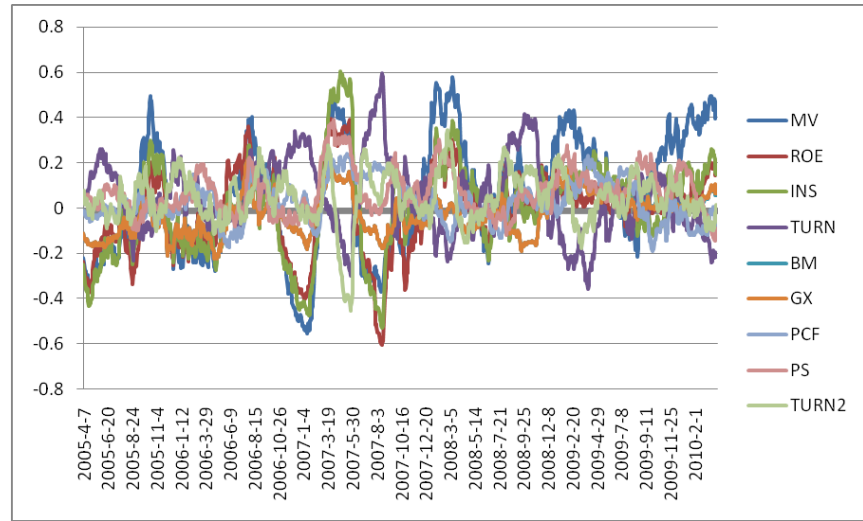
- (1) 同一个因子的偏离度和贡献度的趋势基本相同，因此后面的研究中，我们都只用计算贡献度 CON，后续的量化策略也是基于贡献度的结果。
- (2) 所选的因子集合具有一定的轮动性，即在不同的时间阶段，因子的相对强度排名会发生变化；另一方面，因子表现又具有一定的持续性，所以我们可以利用这两点来寻找特定区间内较强的因子，由此构建对应的量化策略。
- (3) 从因子集合的总体表现来看，也不是任何时候都有显著的因子，比如 2006 年上半年，2007 年四季度，2008 年 6~7 月，2009 年 5~9 月，因子强度就较低，反映当时所选的 9 个因子在市场上作用都不明显，或者说当时投资者不参考这些因子来选择股票。

图 4：中证 800 成分股因子偏离度 (T=60)



资料来源：国信证券经济研究所

图 5：中证 800 成分股因子贡献度 (T=60)



资料来源：国信证券经济研究所

- (4) 如上所提的 4 个区间,并不是说这些时间段市场一定就没有热点或者没有强势因子,比如 2006 年股改、2009 年新能源,可能是政策因子或者行业因子在起作用;又比如 2007 年、2008 年的两个区间,可以看到都发生在大盘的趋势拐点之前,那个时间段市场情绪极度乐观或者极度悲观,确实几乎所有因子都失去作用,投资者不依据理性判断选择股票,因子对股票的区分度降到最低。
- (5) 虽然在两次市场拐点前都出现了低因子强度区间,但从因子贡献度本身无法判断是有集合外因子起作用还是所有因子都失效,所以从这个角度直接构建择时策略是不合适的。在后续的研究中,我们会尝试结合其他指标考虑择时的问题。

因子贡献度分布矩阵

我们对因子贡献度的绝对值进行排名,1 表示该因子相对强度最大,9 则表示该因子相对强度最小,由此得到因子贡献度的分布矩阵。

表 2：中证 800 成分股因子贡献度分布矩阵 (T=60)

因子 CON 排名	MV	ROE	INS	TURN	BM	GX	PCF	PS	TURN2
1	557	150	161	144	26	0	54	75	52
2	231	91	336	167	55	26	79	88	146
3	129	261	173	153	98	55	110	143	97
4	111	178	142	199	3	254	95	124	113
5	50	109	118	136	172	217	140	137	140
6	34	116	79	136	240	205	142	130	137
7	41	92	81	75	386	48	201	157	138
8	28	111	40	86	216	218	187	125	208
9	38	111	89	123	23	196	211	240	188

资料来源：国信证券经济研究所

表 3：中证 800 成分股因子贡献度分布矩阵百分比形式 (T=60)

因子 CON 排名	MV	ROE	INS	TURN	BM	GX	PCF	PS	TURN2
1	45.69%	12.31%	13.21%	11.81%	2.13%	0.00%	4.43%	6.15%	4.27%
2	18.95%	7.47%	27.56%	13.70%	4.51%	2.13%	6.48%	7.22%	11.98%
3	10.58%	21.41%	14.19%	12.55%	8.04%	4.51%	9.02%	11.73%	7.96%
4	9.11%	14.60%	11.65%	16.32%	0.25%	20.84%	7.79%	10.17%	9.27%
5	4.10%	8.94%	9.68%	11.16%	14.11%	17.80%	11.48%	11.24%	11.48%
6	2.79%	9.52%	6.48%	11.16%	19.69%	16.82%	11.65%	10.66%	11.24%
7	3.36%	7.55%	6.64%	6.15%	31.67%	3.94%	16.49%	12.88%	11.32%
8	2.30%	9.11%	3.28%	7.05%	17.72%	17.88%	15.34%	10.25%	17.06%
9	3.12%	9.11%	7.30%	10.09%	1.89%	16.08%	17.31%	19.69%	15.42%

资料来源：国信证券经济研究所

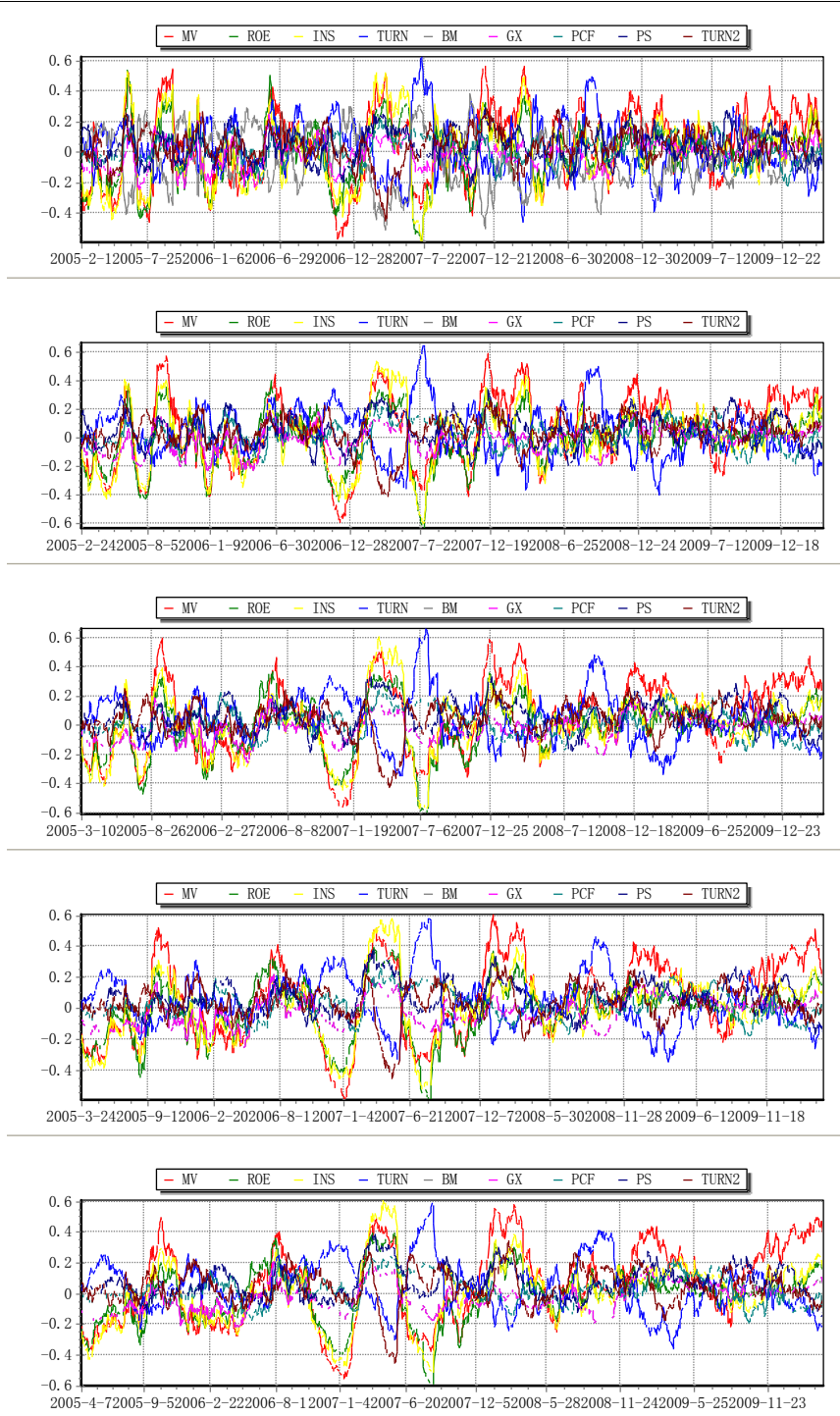
根据因子贡献度分布矩阵，我们对结果分析如下：

- (1) 首先需要强调的是，上面的因子贡献度分布矩阵，从概率意义上来说是一个条件分布，它和所选取的样本池有关，此处我们用的中证 800 成分股，因子上表实际上是中证 800 的因子贡献度分布。如果换成沪深 300 或者全 A 股票，会得到不同的因子分布。这从逻辑上也好理解，以沪深 300 为例，其本身就是大市值股票的集合，因此它的市值因子对收益差的影响，会弱于中证 800。
- (2) 从百分比形式的分布矩阵可以看出，MV 市值因子最为显著，在 1219 个历史交易日中 45.69% 的时间贡献度最高。换言之，从 2005 年至今的 5 年里，大小盘轮动效应确实显著存在。
- (3) 接下来较强的三个因子是 INS、ROE 和 TURN，排第一百分比都在 12% 左右。INS 机构持股因子，在 2008 年之前表现较为突出，且和 MV 因子的趋势接近，这说明两点：第一，在 2008 年之前，大市值的股票机构持股比例要高于小市值股票，但之后机构股票池的覆盖面更广，因子强度降低；第二，INS 因子贡献度的值在正负分布的较为均衡，说明机构持股比例高的股票并不能带来持续的正收益，因此没有必要迷信机构扎堆的股票。
- (4) ROE 因子反映的是市场投资者对公司盈利基本面的选择。其最显著的时期是在 2007 年 6-7 月，即“5.30”之后，市场在突然调整后大家纷纷选择 ROE 高、基本面好、安全边际高的股票。从 2009 年开始 ROE 因子不再显著，一是反映市场热点转移至“概念”“板块”上，二是由于经历全球金融危机，年报所反映的 ROE 可能没有完全真实的反映一个公司长期的盈利能力。
- (5) TURN 因子反映了市场的流动性偏好，它和 MV 因子有一定的负相关性。除了换手率的绝对水平外，我们还考察了换手率的变化程度。结果 TURN2 因子的强度弱于 TURN 因子，证明单只股票成交量的放大或者缩小并不能说明股票未来会有正的超额收益，即成交量变化包含很多无效干扰信息。
- (6) BM、PCF 和 PS 因子都可以看做是估值因子，它们的贡献度较弱，说明简单的看估值指标对指导选股意义不大。

不同参数周期下的因子轮动路径

前面时间周期T都是选择的60日，研究过程中我们也计算了T=20、30、40、50的情形。

图6：中证800成分股因子贡献度（T=20、30、40、50、60）



资料来源：Tinysoft，国信证券经济研究所

上面 5 张图分别是 T 取 20~60 日参数的因子轮动路径，从图上可以看出：首先，因子贡献度的大体趋势是一致的，某个因子强的时候在不同的参数 T 下都能表现出来，而且因子贡献度的峰值对参数 T 不依赖。其次，周期参数 T 越小，因子贡献度的波动就越频繁； T 越大，因子的周期性趋势就越强。最后周期参数 T 越小，滞后程度就越小，拐点在图上反映的就越及时。

因子轮动的外生变量驱动

前面提过，我们的最初想法是找到引起因子轮动的外生变量，比如说现实中发生了什么导致小盘股能跑赢大盘股？如果能够找到这种规律，则可以根据宏观变量或者其他外生变量来预测因子未来的强弱趋势，从而提前布局。

但对比了包括工业增加值、货币供应量、贷款数据、行业数据、开户数等变量后，并没有成功找到这种规律。至少我们没有找到简单的对应关系，而大多数外生变量数据都不具备足够多的数据点来进行复杂分析。因此从过去有限的时间周期内来看，因子的轮动，似乎与这些外生变量独立，是纯粹的市场现象。

没能找到预期中的“因子时钟”，虽然让人沮丧，但也启发我们从另一条途径去构造策略，即不去预测因子的轮动趋势，只根据当前的因子强弱，选择最强的因子来构建组合。

因子轮动选股策略

FACTOR-SCORE 模型

本文前面系统介绍了因子强度的度量方法，构建了偏离度和贡献度两种测度，然后用中证 800 成分股回溯了因子轮动的历史路径。接下来很自然的想法就是如何利用这种因子轮动来构建量化策略。

那么则样将因子的强弱反应到股票上呢？我们开发了一个因子打分模型，利用多因子的排名向量和贡献度向量的内积作为股票的因子分数。

设因子集合为 Γ ，选定任意股票 i 和时间 t ，对于因子 $F \in \Gamma$ ，设股票 i 按照 F 因子值正序排序法的排名为 $rank_t(F, i)$ ，因子 F 的贡献度为 $CON_t(F)$ ，则 t 时刻股票 i 的打分为

$$SCORE_t(i) = \sum_{F \in \Gamma} rank_t(F, i) * CON_t(F)$$

根据前面对贡献度 CON 的定义，一只股票的分数越低，则受因子正向影响的程度越大。

多因子半衰期选股策略

根据 FACTOR-SCORE 模型，就能得到每一天的股票 SCORE 截面。我们可以每天选取打分前 20% 和后 20% 的股票作为多头和空头，接下来的问题就是多空组合分别持有多长时间。

常见的方法是依据因子的走势来换仓，但这涉及到两个问题：首先，如果

需要对因子的走势进行判断，预测因子的难度不小于预测股价的难度；其次，我们面对的是一个 9 维的因子向量，很难找到合适的方法去处理 9 个因子间的关系。

因此构建策略的时候，从另一角度出发，既然在数学上 SCORE 作为内积是因子排名向量在贡献度扩展平面的投影，可以只对这个投影的变化设定准则而不考虑原向量的位移。最终我们设定 SCORE 排名的半衰期作为组合的持有期。以多头组合为例，策略具体步骤为：

- (1) 利用 FACTOR-SCORE 模型，对股票池中全部 N 只股票进行打分，得到任意股票 i 在时刻 t 的分数 $SCORE_t(i)$ 。
- (2) 每日都将 N 只股票按照 SCORE 以正序法排名，得到前 20% 的 n 只股票组合 $SET_{low-S}(t)$ （后 20% 的 n 只股票组合类似的记为 $SET_{high-S}(t)$ ）。
- (3) 设策略起始点为 t_0 ，初始多头组合 $PF_{long}(t_0) = SET_{low-S}(t_0)$ 。
- (4) 对因子贡献度绝对值的最大值进行判定，当最大值小于 0.2 时，持有组合不变。当贡献度绝对值小于 0.2 时，一方面此时因子带来的超额收益率已经非常低，不能覆盖换仓的冲击成本；另一方面当贡献度绝对值很小时，会出现因子失效现象，此时因子集的排名会变得相当不稳定。
- (5) 当因子贡献度最大值超过 0.2 时，比较当前持有的多头组合与当天打分前 20% 的股票组合。当持有的多头组合中首次超过一半数量的股票落在打分前 20% 的组合外时，进行换仓。用公式表示则为：

$$t_{k+1} = \min_{s > t_k} (|PF_{long}(t_k) \cap SET_{low-S}(s)| \leq n/2 \text{ 且 } \max(|CON_s(F)|) > 0.2)$$

其中 t_k 为第 k 次换仓点， $|PF_{long}(t_k) \cap SET_{low-S}(s)|$ 表示两个集合交集的

元素个数。 $t_{k+1} - t_k$ 则为第 k 个多头因子的半衰期。

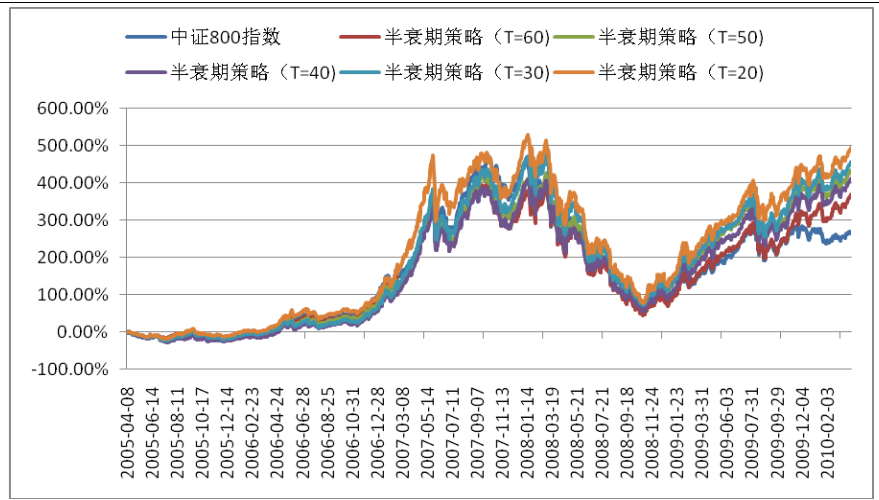
策略实例

在历史模拟中，我们选择中证 800 成分股为股票池，以中证 800 指数为比较基准。多头组合选取 160 只（特殊日期为 159 只）股票，等资金权重建仓。

模拟中单边交易成本取 0.15%，初始资金为 10 亿。成交价格为组合中个股第二天的 TWAP 成交价格。

模拟区间为 2005 年 4 月 7 日至 2010 年 4 月 9 日，参数 T 分别取 20、30、40、50、60。下图是多因子半衰期选股策略与中证 800 指数的累计收益率比较。

图 7：多因子半衰期选股策略与中证 800 指数累积收益率对比



资料来源：国信证券经济研究所

为了逼近最真实的操作效果，我们设计了严格的测试流程：每次换仓时点，利用当日收盘后数据进行判定，如果判断需要换仓，下一个交易日才进行操作；保留5%现金仓位，每次的操作根据TWAP成交均价计算冲击成本，考虑单边0.15%的股票交易成本，并对停牌、涨跌停等异常交易条件下的交易做严格的剔除；分红配送计入组合收入并进行再投资。

对于停牌的股票，若没有发生交易，则不需要处理；若发生交易，则分股票考虑。若为应加入的股票，则不买入该股票；若为应剔除的股票，则直接按上一收盘价卖出；若为继续持有的股票，则维持上一日持仓手数，不作变动。

这样，股票交易的管理成本可控，同时我们得到的是一个在极为接近现情形的场景下所实现的股票组合收益。

表 4：多因子半衰期选股策略统计表

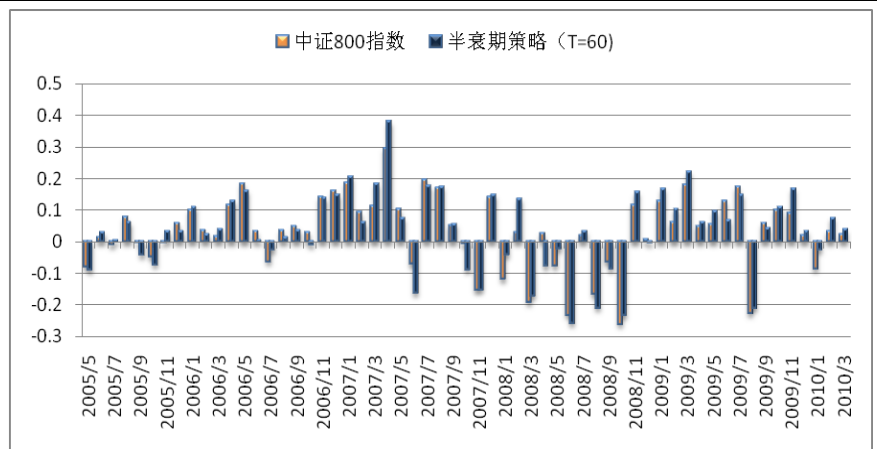
	中证 800 指数	多因子半衰期选股策略				
		T=60	T=50	T=40	T=30	T=20
累积收益率（全区间）	269.23%	368.32%	435.62%	412.22%	455.99%	492.97%
累积收益率（07 年起）	85.28%	157.69%	204.01%	224.97%	237.23%	205.08%
换仓次数（全区间）		32	35	43	49	67
月度收益率跑赢基准次数（全区间）	59	33	33	31	31	34
月度收益率跑赢基准次数（07 年起）	39	26	28	27	27	26

资料来源：国信证券经济研究所

对模拟的结果，我们做了如下分析：

- 多因子半衰期选股策略在 07 年之前收益率不如作为基准的中证 800 指数，我们认为这是股改的特殊性造成的：首先在 05~06 年股改政策使得投资者淡化对其他因子的关注，纷纷选择股改的股票；其次在我们严格的模拟条件中，剔除了涨停板和停牌的股票，但是这些股票的收益包含在中证 800 指数中。
- 如果从 2007 年开始统计，多因子半衰期选股策略则明显优于中证 800 指数，在 39 个月中有 26 个月的月度收益率超越基准，即胜率为 66%。

图 8：多因子半衰期选股策略与中证 800 指数月度收益率对比



资料来源：国信证券经济研究所

- (3) 对于参数 T ，可以看到策略的有效性并不依赖于 T 。从理论上而言， T 越小，则因子强度变化的反应速度就越快，滞后性越弱。从模拟结果中可以看到，换仓次数随着 T 的减小而增多， $T=60$ 时平均每年换仓约 6 次，而 $T=20$ 时平均每年换仓约 13 次，大致每个月一次。
- (4) 我们也模拟了空头组合的效果（即打分排名最高的 160 只股票，换仓规则也是半衰期方法），但是结果并不理想，参数依赖比较严重，只有 $T=20$ 时在 07 年前和 07 年后都弱于中证 800 指数。不过空头组合的收益率低于多头组合具有稳定性，与参数无关。事实上，因子对于多头和空头的作用效果是不对称的，对于因子的正向作用往往存在反应不足，产生动量效应；而对因子的反向作用存在反应过度，产生反转效应。
- (5) 多因子半衰期选股策略，其逻辑并不是要选出全部最好的股票，而是从因子的角度出发，追踪市场上普遍关注并正在起作用的因子，并选出受益于这些因子的股票集来，希望通过概率去战胜基准。

基于 Alpha 分类的多因子增强策略

因子贡献度模型的扩展应用

上文在介绍因子贡献度分布矩阵时曾经提过，从概率意义上来说是一个条件分布。因此我们可以对任何一个股票池和任意一个因子集合，计算其分部矩阵。结合 FACTOR-SCORE 模型，在构建量化策略时，不仅仅能够设计像多因子半衰期选股策略这样在中证 800 里的选股模型，而且能够对任意一个已有的股票组合进行增强。

下面将介绍如何将贡献度模型和 FACTOR-SCORE 模型用在正 Alpha 行业配置策略的增强上。正 Alpha 行业配置策略将行业分为长期 Alpha 距离最远的两类，通过穷举验证，我们至今为止仍然战胜数十万种其余的行业组合分类方法。这一分类的结论详见《国信证券-数量化投资系列之七：基于 Alpha 的行业配置方法和投资策略——总是获得正 Alpha》。运用上述思路，将中证 800 指数成份股按

照行业的属性分成一类和二类，根据一类和二类组合资产相对于中证800的Alpha值轮动，进行超配和低配操作，从而获得正的Alpha。

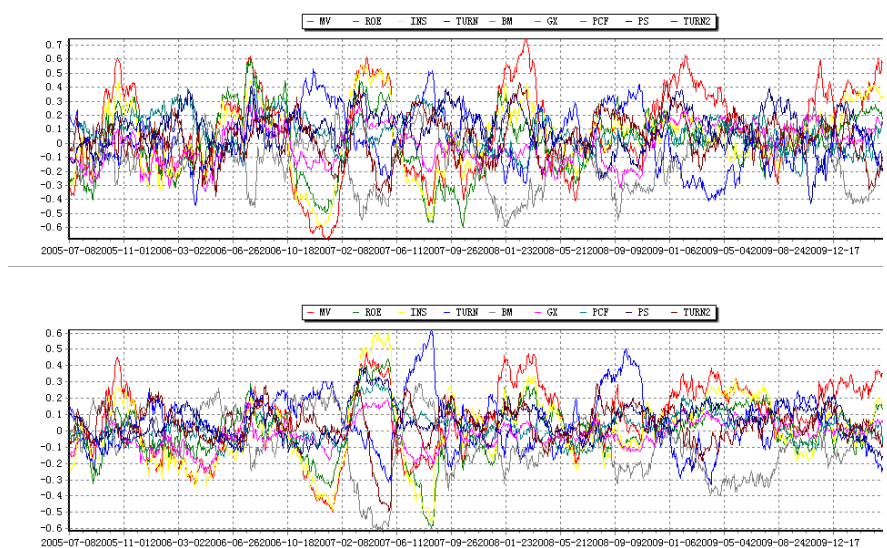
我们的增强方法并不改变分类，只是利用其他的因子对每一类组合的股票权重进行调整，可以视作多因子对原来 Alpha 行业因子的补充。

一二类行业的多因子增强

增强步骤：

- (1) 首先按照 alpha 因子和行业属性，将中证 800 成分股分为一类行业和二类行业两个组合，将这两个组合分别作为股票池，计算 9 个因子的因子贡献度向量。
- (2) 利用 FACTOR-SCORE 模型，对每日一二类行业的组合分别打分，选出分数最低的 20%股票。
- (3) 对一类行业和二类行业，85%的比例按照自由流动股本权重配置该分类全部股票，剩下 15%的比例等权重配置增强的组合，这样便得到新一类行业和新二类行业两个组合，与原来的组合相比股票构成并不变，发生变化的只是股票权重。

图 9：一类行业、二类行业因子贡献度 (T=60)

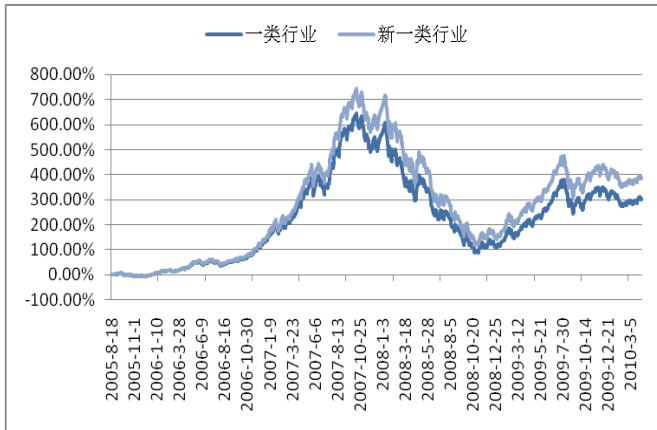


资料来源：Tinysoft，国信证券经济研究所

从上图可以看出，一类行业和二类行业的因子贡献度的历史路径存在差异，与中证 800 全样本下的因子贡献度表现也不相同。有些因子如 MV 因子在行业分类后的贡献度降低了，而如 BM 因子在分类后贡献度却升高了。

按照原始的一二类行业划分，从 2005 年 8 月 19 日至 2010 年 4 月 9 日，两个组合的区间收益率分别为 304.53%和 367.46%。以 15%比例增强后，新的两个组合的区间收益率分别为 388.72%和 405.21%，均有明显的增强效果，其中对一类行业的增强效果更为明显。多因子增强的极端情形是将增强的比例从 15%提高至 100%，此时即一二类行业均只选取其 Top 组合。

图 10: 增强前后的一类行业收益率 (T=30, 15%配置)



资料来源: Tinysoft, 国信证券经济研究所

图 11: 增强前后的二类行业收益率 (T=30, 15%配置)



资料来源: Tinysoft, 国信证券经济研究所

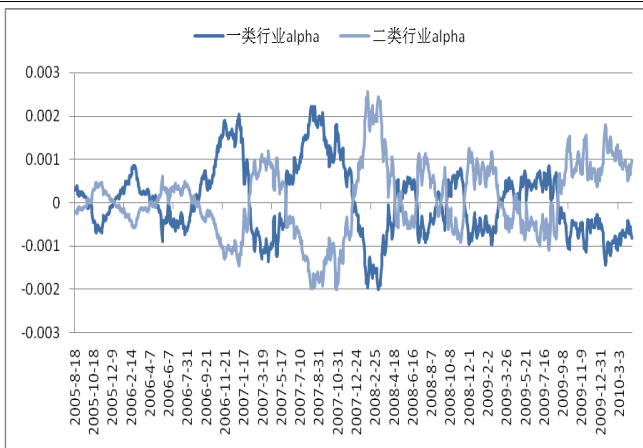
增强后的新正 Alpha 行业配置策略

有了新的一类行业 and 二类行业组合, 我们将两个组合根据市值加权合并到一起作为一个新的计算 Alpha 值的基准组合, 重新计算两类新行业组合的 Alpha 值, 两者始终存在较为稳定的轮动关系, 并且在每一轮轮动中能够完成高于一定水平的 Alpha 偏离, 这是轮动策略能够获取超额收益的关键。本报告中我们仍然采用 Alpha 动量作为选择组合的方法, 总是选择新 Alpha 值为正的新行业组合。实际上由于我们的量化研究体系在方法论上不断创新, 目前我们已经拥有四种可以比较切换效果的方法, 这些方法的原理和目标各不相同, 并不存在自反馈的问题, 因此可以在未来的应用中持续比较和辅助参考。在随后的研究报告中我们将进一步对比它们的效果和试用周期。

在历史回溯模拟中, 我们选择 2005 年 8 月 19 日至 2010 年 4 月 9 日为样本区间, 交易条件与多因子半衰期选股策略中的设定相同, 尽量接近真实的交易环境。

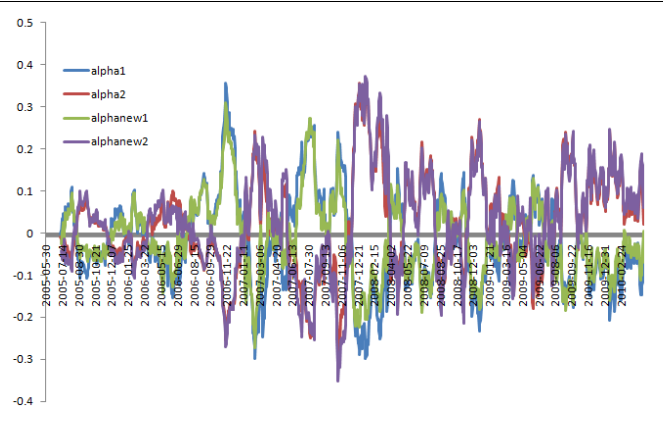
对于参数 T, 由于程序计算量过大, 只检验了 T=30 的情形。

图 12: 新一类行业和新二类行业 alpha 轮动图



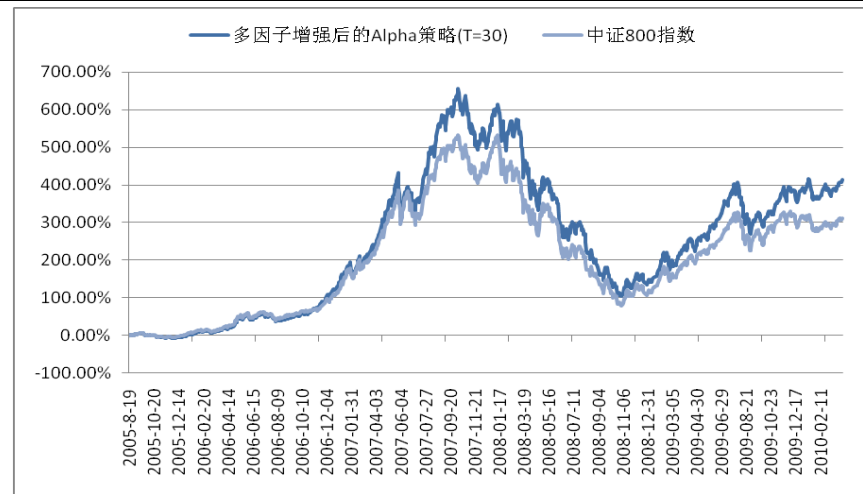
资料来源: Tinysoft, 国信证券经济研究所

图 13: 增强前后两类行业组合 Alpha 轨迹 (30 日)



资料来源: Tinysoft, 国信证券经济研究所

图 14：多因子增强后的正 Alpha 行业配置策略与中证 800 指数累计收益率对比



资料来源：国信证券经济研究所

上面第一张图可以明显的看出，新一类行业和新二类行业的 alpha 值具有强负相关性，相关系数为-0.9757，在能够正确选择基准组合及加权方式的基础上，这两个行业组合的 Alpha 强负相关是必然的，我们更关注的是二者之间的 Alpha 差和轮动频率。

为了使 Alpha 的来源集中在配置和选股，而不是加权方式上，我们在新一类和新二类的 alpha 轮动中，采用多因子打分更改加权方式后的中证 800 重新拟合了基准，因此图 12 中的 Alpha 与我们在周报中公布的中证 800 两类行业的 Alpha 并不能简单比较。

图 13 是我们根据一类、二类行业以 15%权重配置多因子选股的 top 组合后的新组合与传统中证 800 指数回归得到的 Alpha，对比旧组合的 Alpha。从图上简单地看，新组合并未从 Alpha 轨迹上偏离原有组合，也未显著改变两个组合的收益差和交叉次数。因此我们认为最终累计出来的超额收益大多来自选股而不是对配置策略的更改。

图 14 为增强的新 Alpha 行业配置策略和中证 800 指数的累计收益率对比图，可以看到相对于样本区间内中证 800 指数 310.75%的收益率，新策略为 413.21%，获得了约 100 个基点的超额收益，其来源于原 Alpha 策略的行业增强和多因子的选股增强两部分。

与多因子半衰期选股策略相同，由于在设定交易条件时无法买入涉及股改的股票，因此在 2006 年的特殊阶段，增强后的 Alpha 策略也不尽如人意。但随着中国股票市场建设的逐步完善，投资者群体的日趋理性，这两个量化策略会有更稳定的表现。这也为我们提出了一个新的课题，就是能否将事件冲击的影响量化，在不对事件发生做预测和不采用非公开信息的情况下，利用事件公布后停牌前的时间进行组合修正，平滑单个股事件冲击对组合市值稳定性的影响。

结语

模型与量化策略的总结

良好的扩展性

不论是在研究过程中，还是在相应的程序编写中，我们都尽量保证模型的扩展性。事实上，影响市场的因子很有多，宏观因子、财务因子、股指因子、市场因子等等，不但是文中所列的 9 个因子，对任意因子，甚至任意一种排名规则，都可以纳入我们的因子贡献度模型和 FACTOR-SCORE 模型框架之内。

其次，文中多次强调，贡献度因子分布实质上中一种条件分布，因此不仅是中证 800 成分股，对任意的股票池，都能计算对应的因子贡献度及其分布。换言之，多因子半衰期选股策略在沪深 300、沪市、深市甚至全部 A 股的股票池都能应用，也能直接移植到各种指数基金的增强上

最后，对任意已有的量化策略，只要该策略是基于股票组合的操作，类似于正 Alpha 行业配置策略，我们都能对其进行多因子增强。

无参数依赖性

量化策略良莠不齐，很多历史回溯有效的策略在未来失效，除了策略钝化和市场结构变化之外，很多时候是因为策略具有参数依赖性，历史超额收益的获得只不过是取了特定的参数。但我们的策略有效性并不依赖于参数，对于参数时间周期 T，结果已经体现不同的 T 值都能获得相对一致的超额收益，其本质是市场存在真实的因子驱动效应，而不是统计游戏的结果。在组合构建的逻辑上，也没有依赖任何从历史中提炼的规律（其实规律不可怕，可怕的是在有限的样本内提取的不是规律，而是巧合现象）。

非预测性

有些量化策略需要对未来进行预测，比如之前我们所做的宏观多因子模型，就是将对众多股票的预测转换为对四个核心宏观因子的预测。这样的策略虽然简化了问题，但仍然存在主观判断的步骤。而本报告的策略完全基于已有的数据，目标在于有效地跟踪市场热点因子，不需要对因子的强度变化或趋势进行任何预测。

后续研究方向

我们的多因子研究框架只是完成了一个初步的底层框架，并且用我们所能拿到的最大样本尽可能地做了回溯。目前还有几个关键问题没有解决，后续的研究也会在以下四个方向继续前进：

外生变量的问题

我们的模型解决了因子强度的度量和跟踪问题，但因子效应究竟是什么外生变量驱动的？因子轮动又有什么规律，这种规律是否稳定？这两个问题始终是因子分析的最高目标，貌似没有最优解。

最优周期问题

在固定的时间周期参数 T 下，我们选出了每个时期的最强因子。但在不同时期，因子轮动的速度是不同的，即应该存在动态的最优周期。如果能构建一种方法，消化周期变化对策略的负面影响，还将大幅提升策略的效果。

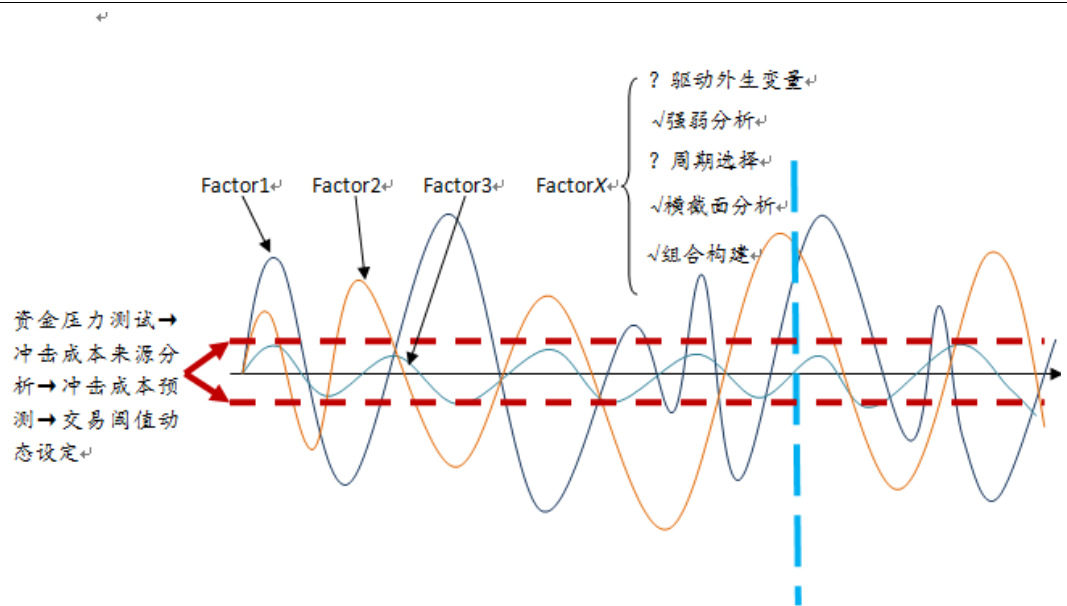
因子集选择问题

本文只选择了 9 个常见的因子，且不包含宏观因子。在后续的研究中，我们会尝试加入更多的因子，将多因子轮动模型和宏观多因子模型结合起来，更准确的捕捉可测的最强因子。

交易优化问题

对于换仓的因子贡献度阈值 0.2, 主要是考虑因子带来的正收益差必须覆盖交易成本和冲击成本。严格的说，交易成本和冲击成本是动态变化的，因此最合理的阈值应该取决于动态的交易成本和冲击成本估计，这将和未来的交易优化研究结合起来。

图 15：后续研究扩展示意图



资料来源：国信证券经济研究所

国信证券投资评级

类别	级别	定义
股票 投资评级	推荐	预计 6 个月内，股价表现优于市场指数 20%以上
	谨慎推荐	预计 6 个月内，股价表现优于市场指数 10%-20%之间
	中性	预计 6 个月内，股价表现介于市场指数±10%之间
	回避	预计 6 个月内，股价表现弱于市场指数 10%以上
行业 投资评级	推荐	预计 6 个月内，行业指数表现优于市场指数 10%以上
	谨慎推荐	预计 6 个月内，行业指数表现优于市场指数 5%-10%之间
	中性	预计 6 个月内，行业指数表现介于市场指数±5%之间
	回避	预计 6 个月内，行业指数表现弱于市场指数 5%以上

免责声明

本报告信息均来源于公开资料，我公司对这些信息的准确性和完整性不作任何保证。报告中的内容和意见仅供参考，并不构成对所述证券买卖的出价或询价。我公司及其雇员对使用本报告及其内容所引发的任何直接或间接损失概不负责。我公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行业务服务。本报告版权归国信证券所有，未经书面许可任何机构和个人不得以任何形式翻版、复制、刊登。

国信证券经济研究所研究团队(含联系人)

宏观			策略			交通运输		
周炳林	0755-82133339		赵 谦	021-60933153		郑 武	0755- 82130422	
林松立	010-82254212		崔 嵘	021-60933159		陈建生	0755- 82130422	
			廖 喆	021-60933162		岳 鑫	0755- 82130422	
			黄学军	021-60933142		高 健	0755-82130678	
银行			房地产			机械		
邱志承	021-68864597		方 焱	0755-82130648		余爱斌	0755-82133400	
黄 飙	0755-82133476		区瑞明	0755-82130678		黄海培	021-60933150	
谈 煜	010- 82254212		黄道立	0755-82130833		陈 玲	0755-82133400	
王一峰	010- 66026320					杨 森	0755-82133343	
						李筱筠	010-82254205	
汽车及零配件			钢铁			商业贸易		
李 君	021-60933156		郑 东	010-82254160		孙菲菲	0755-82133400	
左 涛	021-60933164		秦 波	010-66026317		吴美玉	010-82252911	
						祝 彬	0755-82131528	
基础化工			医药			石油与石化		
张栋梁	0755-82130532		贺平鸽	0755-82133396		李 晨	021-60875160	
陈爱华	0755-82133397		丁 丹	0755-82130678		严蓓娜	021-60933165	
邱 斌	0755-82130532		陈 栋	021-60933147				
电力设备与新能源			传媒			有色金属		
皮家银	021-60933160		陈财茂	021-60933163		彭 波	0755-82133909	
						谢鸿鹤	0755-82130646	
电力与公用事业			非银行金融			通信		
徐颖真	021-60875162		邵子钦	0755- 82130468		严 平	021-60875165	
谢达成	021-60933161		田 良	0755-82130513		程 峰	021-60933167	
			童成敦	0755-82130513				
造纸			家电			计算机		
李世新	0755-82130565		王念春	0755-82130407		段迎晟	0755- 82130761	
邵 达	0755-82132098							
电子元器件			纺织服装			农业		
段迎晟	0755- 82130761		方军平	021-60933158		张 如	021-60933151	
旅游			食品饮料			建材		
廖绪发	021-60875168		黄 茂	0755-82133476		杨 昕	021-60933168	
刘智景	021-60933148							
煤炭			建筑			固定收益		
李 然	010-66026322		邱 波	0755-82133390		李怀定	021-60933152	
陈 健	010-66215566		李遵庆	0755-82133343		张 旭	010-82254210	
苏绍许	021-60933144					高 宇	0755-82133528	
						蔺晓熠	021-60933146	
指数与产品设计			投资基金			量化投资		
焦 健	0755-82131822		杨 涛	0755-82133339		葛新元	0755-82133332	
赵学昂	0755-82131822		黄志文	0755-82133928		董艺婷	021-60933155	
王军清	0755-82133297		刘舒宇	0755-82131822		林晓明	021-60933154	
阳 瑾	0755-82131822		康 亢	010-66026337				
周 琦	0755-82131822		彭怡萍					
彭甘霖								

国信证券机构销售团队

华南区			华东区			华北区		
万成水	0755-82133147		盛建平	021-60875169		王立法	010-82252236	
	13923401205			15821778133			13910524551	
	wancs@guosen.com.cn			shengjp@guosen.com.cn			wanglf@guosen.com.cn	
邵燕芳	0755-82133148		马小丹	021-60875172		王晓建	010-82252615	
	13480668226			13801832154			13701099132	
	shaoyf@guosen.com.cn			maxd@guosen.com.cn			wangxj@guosen.com.cn	
林 莉	0755- 82133197		郑 毅	021-60875171		谭春元	010-82254209	
	13824397011			13795229060			13810118116	
	Linli2@guosen.com.cn			zhengyi@guosen.com.cn			tancy@guosen.com.cn	
王昊文	0755-82130818		黄胜蓝	021-60875173		焦 骥	010-82254202	
	18925287888			13761873797			13601094018	
	wanghaow@guosen.com.cn			huangsl@guosen.com.cn			jiaojian@guosen.com.cn	
甘 墨	0755-82133456		刘 塑	021-60875177		李锐	010-82254212	
	15013851021			13817906789			13691229417	
	ganmo@ guosen.com.cn			liusu@guosen.com.cn			lirui2@guosen.com.cn	
			叶琳菲	021-60875178		徐文琪	010-82254210	
				13817758288			13811271758	
				yelf@guosen.com.cn			xuwq@guosen.com.cn	
			孔华强	021-60875170				
				13681669123				