




2017.06.15

# 基于短周期价量特征的多因子选股体系

## - - 数量化专题之九十三

	<b>李辰 (分析师)</b>	<b>刘富兵 (分析师)</b>
	021-38677309	021-38676673
	lichen@gtjas.com	liufubing008481@gtjas.com
证书编号	S0880516050003	S0880511010017

**本报告导读：**本篇报告为投资者详尽介绍了全新的交易型阿尔法体系的构建思路、策略特点及实证分析等研究成果。

### 摘要：

- 本篇报告中，我们将开创性的构建全新的多因子模型体系——短周期交易型多因子阿尔法选股体系。
- 通过交易型阿尔法策略的研究，我们发现在A股市场，与传统多因子模型所获取的股票价值阿尔法收益相比，交易型阿尔法收益的空间更大、收益稳定性也更强。
- 即便是最纯粹的价值投资者也不得不承认，交易行为在短期内对股票价格起着几乎是决定性的影响，而发掘这种交易型套利空间正是量化投资的优势所在，因此量化模型在这一领域内应有着广泛的运用空间。
- 在本篇报告所构建的交易体系中，我们总共构建、运用了将近200个短周期阿尔法因子，其中因子数据则均来自于个股日频率的价格与成交量数据。在此基础上，我们构建了基于短周期价量特征的风格中性多因子选股策略。
- 策略自2012年1月至2017年4月，在扣除所有交易成本后，较之中证500指数，实现年化超额收益率50.2%，最大回撤5.9%，信息比率4.67。
- 对交易成本的敏感性分析结果表明，由于对换手率与交易成本进行了最优平衡过程的处理，因此策略可容纳的极限交易成本接近双边1%，模型实战能力较强。
- 短周期交易型阿尔法体系既是对传统多因子体系的补充，也可以说是全新思路、独立设计的交易体系。在这其中，量化模型不再仅仅是低风险低收益的投资策略，同样也可获得高额收益回报，而尽管我们在这其中还只是看到了冰山一角。
- 短周期交易型阿尔法策略体系的构建，希望对投资者的研究思路有所拓展，在量化模型最擅长的领域发挥其真正的威力。

### 金融工程团队：

**刘富兵：(分析师)**  
 电话：021-38676673  
 邮箱：liufubing008481@gtjas.com  
 证书编号：S0880511010017

**陈奥林：(分析师)**  
 电话：021-38674835  
 邮箱：chenaolin@gtjas.com  
 证书编号：S0880516100001

**李辰：(分析师)**  
 电话：021-38677309  
 邮箱：lichen@gtjas.com  
 证书编号：S0880516050003

**孟繁雪：(分析师)**  
 电话：021-38675860  
 邮箱：mengfanxue@gtjas.com  
 证书编号：S0880517040005

**蔡旻昊：(研究助理)**  
 电话：021-38674743  
 邮箱：caiminhao@gtjas.com  
 证书编号：S0880117030051

**殷明：(研究助理)**  
 电话：021-38674637  
 邮箱：yinming@gtjas.com  
 证书编号：S0880116070042

**叶尔乐：(研究助理)**  
 电话：021-38032032  
 邮箱：yeerle@gtjas.com  
 证书编号：S0880116080361

### 相关报告

《基于短周期价量特征的多因子选股体系》  
 2017.06.01  
 《基于动态模分解的价格模式挖掘》  
 2017.05.15  
 《负基差已成常态，投机资金参与热情不高》  
 2017.05.05  
 《Timing the market-An strategy based on W-shaped Bottom》2017.04.11  
 《W型市场底部研究》2017.03.14

## 目录

1. 引言 .....	3
2. 阿尔法模型的评价标准 .....	4
2.1. 阿尔法模型的评价标准 .....	4
2.2. 阿尔法模型预测的一致性问题 .....	5
3. 基于短周期价量特征的多因子选股体系 .....	6
3.1. 短周期交易型阿尔法策略的构建思路 .....	6
3.2. 一些显著的价量特征举例 .....	8
3.3. 短周期价量因子体系构建 .....	10
3.4. 因子检验及相关特征 .....	17
3.4.1. 因子体系的显著性检验、相关性及有效周期 .....	17
3.4.2. 因子体系的预测能力 .....	20
3.4.3. 因子个数与模型预测能力的关系 .....	21
3.5. 组合换手率与超额收益 .....	22
3.6. 基于短周期价量特征的风格中性多因子选股策略 .....	23
3.6.1. 策略设计 .....	23
3.6.2. 实证分析 .....	24
3.6.3. 交易成本敏感性分析 .....	26
3.6.4. 策略可容纳规模估算 .....	27
4. 总结与展望 .....	28
5. 附录 .....	29
附录 1 大类风格因子定义明细 .....	29
附录 2 表 6 相关函数定义 .....	30

## 1. 引言

传统多因子模型在 A 股量化投资领域被广泛运用，在过去若干年内模型也获得较为稳健的超额收益。然而，由于市值效应在 A 股市场的影响过于明显，导致传统多因子模型或多或少都受其影响。

尤其是自 2017 年以来，随着市场风格发生的急剧变化，策略稳定性受到了一定冲击，其中主要原因不外乎以下 3 点：首先，阿尔法因子体系与市值因子相关性过高；其次，组合对风格的暴露控制力不足；最后，选股域与比较基准域之间的不匹配。总而言之，市场风格的变化对策略的影响远大于阿尔法因子本身所产生的波动，因此在当前市场环境下，对传统多因子模型的风格控制能力提出了更高的要求，即使没有显著的风格预判能力，那至少也需具备风格中性的控制能力。从目前市值因子的波动情况而言，小盘因子超额收益的黄金时代可能已经过去。

在本篇报告中，我们将开创性的构建全新的多因子模型体系——短周期交易型多因子阿尔法选股体系。这将对原有传统多因子模型的极大补充，我们将从收益来源、有效周期、因子体系特征、组合换手率、策略设计等多个方面来分析我们所构建的新体系。

通过交易型阿尔法策略的研究，我们发现在 A 股市场，与传统多因子模型所获取的股票价值阿尔法收益相比，交易型阿尔法收益的空间更大、收益稳定性也更强，这与 A 股市场的投资者结构、交易制度、投资者交易特征有着很大的关系。在 A 股市场中，即便是最纯粹的价值投资者也不得不承认，交易行为在短期内对股票价格起着几乎是决定性的影响，而挖掘这种交易型套利空间正是量化投资的优势所在，因此量化模型在这一领域内应有着广泛的运用空间。

在本篇报告中，我们总共构建、运用了将近 200 个短周期阿尔法因子，其中因子数据则均来自于个股日频率的价格与成交量数据。在此基础上，我们构建了基于短周期价量特征的风格中性多因子选股策略。模型自 2012 年 1 月至 2017 年 4 月，在扣除所有交易成本后，较之中证 500 指数，实现年化超额收益率 50.2%，最大回撤 5.9%，信息比率 4.67。策略设计以 T+2 日频率调仓，组合每期持股约为 50-60 只。

短周期交易型阿尔法体系既是对传统多因子体系的补充，也可以说是全新思路、独立设计的交易体系。在这其中，量化模型不再仅仅是低风险低收益的投资策略，同样也可获得高额的收益回报，而尽管我们在这其中还只是看到了冰山一角。

## 2. 阿尔法模型的评价标准

### 2.1. 阿尔法模型的评价标准

在介绍交易型阿尔法策略新体系之前，我们有必要对阿尔法模型的评价标准再做些许的梳理，而这样的标准既适用于交易型阿尔法体系，同样也适用于传统阿尔法模型。

毫无疑问，投资必然是以结果为导向的，因此策略最后实现的收益风险特征是衡量阿尔法模型好坏的最终标准。但是，阿尔法模型的本质即是预测收益，因此在观察策略实际收益率之前，我们可以通过一些定量的方法，计算阿尔法模型的收益预测是否精确、是否显著，这样的判断将更有利于我们了解自己阿尔法模型的特性，提高投资成功的概率。

在我们的因子模型体系中，收益率分解是基本的假设前提。我们认为任意股票在同一时刻暴露于多种不同的风险因素下，它们之间的共同作用形成了股票价格的波动。因此，我们致力于寻找对绝大部分股票价格波动都有影响共同因素，称其为风格因子，这部分收益则称为风格收益。而风格因子不可解释的部分，则认为是个股自身特有的属性，称其为特质因子，这部分收益则自然称为阿尔法收益。

即，股票收益率 = 风格收益 + 阿尔法收益。

其中，A股市场最为典型的风格收益包括行业风格、市值风格等。对于各大类风格因子（行业、贝塔、动量、市值、盈利、波动、成长、价值、杠杆、流动性）的定义，我们在之前的报告中已经做了详细的阐述，或详见附录1。

风格收益的特点在于显著但不可预测（或者说在我们的因子模型体系中不对风格收益进行预测），而阿尔法收益的特点是显著并且一定程度上可预测。

因此，阿尔法模型的目标应是针对阿尔法收益的预测，而不是针对股票整体收益率的预测。而评价阿尔法模型的标准也应是计算其对阿尔法收益的预测是否可靠。

从建模的角度而言，假设  $R$  是所有个股在某一期内的收益向量， $X_{industry}$

是定义的行业哑变量矩阵， $X_{style}$  为定义的风格因子载荷矩阵（其中包

括 *Beta*、*Momentum*、*Size*、*Earnings\_Yield*、*Volatility*、*Growth*、*Value*、*Leverage*、*Liquidity*），我们通过回归方程的计算，可以得到：

$$R_{t+1} = f_{industry} X_{industry} + f_{style} X_{style} + \varepsilon_{t+1}$$

其中， $f_{industry}$  和  $f_{style}$  分别为行业因子和风格因子的因子收益率

$factor\ return$  ,  $\varepsilon$  则为当期个股阿尔法收益率截面。

值得注意的是, 假设站在投资时点  $t$  时刻, 我们可获得的变量仅为  $X_{industry}$  和  $X_{style}$  , 因此未来一期的阿尔法收益率截面  $\varepsilon_{t+1}$  是需要预测的。

因而, 无论我们选择何种因子定义、何种因子加权方式、何种收益率转化方式等, 最后产生的计算结果都是阿尔法收益率的预测值, 即

$$Alpha\_Model \rightarrow E\{\varepsilon_{t+1}\}$$

换言之, 我们在  $t$  时刻末通过阿尔法模型计算得到  $t+1$  期的阿尔法收益率预测值  $E\{\varepsilon_{t+1}\}$  , 而到了  $t+1$  时刻末, 我们通过回归建模得到  $t+1$  期

阿尔法收益率的实际值  $\varepsilon_{t+1}$  , 我们通过预测阿尔法截面与实际阿尔法截面之间的相关系数计算, 即可对阿尔法模型该期的预测效果得到定量的评估, 我们称其为阿尔法模型预测系数 ( $IC\ of\ Alpha\ Model$ ), 即

$$IC_{AlphaModel}^t = corr(E\{\varepsilon_{t+1}\}, \varepsilon_{t+1})$$

那么, 对于足够长的一段时间  $T$  , 我们可以计算得到每期阿尔法模型预测系数的时间序列, 接下来通过一些简单的统计检验就可以判断, 阿尔法模型的预测性是否显著, 例如  $T$  检验, 即

$$TStat_{AlphaModel} = \overline{IC_{AlphaModel}} / (\sigma_{IC_{AlphaModel}} / \sqrt{T})$$

显著性检验的结果即可视为对一个阿尔法模型的定量评价, 那么对于若干不同的阿尔法模型就可以有严格的好坏区分。并且, 从经验上而言, 预测能力较强的阿尔法模型, 其所对应的实际组合获取超额收益的概率也越高。当然, 如果阿尔法模型预测能力的检验非显著, 那么甚至都没有必要去尝试构造组合, 因为最终即便实现的超额收益, 若不是风格收益, 那么也只是运气而已。

遗憾的是, 传统阿尔法因子所构建的阿尔法模型体系, 其预测能力在剔除市值因子的效用之后究竟有多强的显著性, 值得深思。

## 2.2. 阿尔法模型预测的一致性问题的

至今为止, 没有任何一个投资模型可以解释市场运行的一切变化规律。投资不需要真理, 但是需要逻辑和依据。

在整个多因子模型体系中, 策略的构建分为选股和组合两个步骤。选股即为预测阿尔法收益, 组合则为实现阿尔法收益, 两者互相独立却又一脉相承。



所谓阿尔法模型预测的一致性问题的，指的是阿尔法因子的预测目标与组合实现的收益目标相一致。也就是说，阿尔法模型的预测过程与组合构造的实现过程需要有共同的目标。

举一些典型的例子来说明这个问题。我们经常利用分组计算收益的方式来考察因子的单调性，而这一过程其实是在对个股的整体收益率进行预测（风格收益+阿尔法收益），即选股目标为预测个股整体收益。而到了构建组合的时候，为了控制回撤降低波动率，我们加入了行业中性、市值中性等约束手段，即组合过程希望获取较为纯粹的阿尔法收益。显然，此时预测过程与组合实现过程出现了目标不一致的情况，在最后实现的超额收益中，因子的预测作用究竟体现了多少，便不得而知了。

另一种情况则恰恰相反，我们利用一些统计回归、截面正交等方式，在检验因子的时候就尽可能多的剔除各种风格的影响，最后得到因子对纯阿尔法收益的预测结果。但在组合构建过程中，没有对上述剔除的各种风格因子进行严格的中性约束，导致组合仍有一定的风格敞口暴露。此时预测过程与组合实现过程仍有不一致性的问题，因子预测的作用又大打折扣。

因此，在我们的多因子体系中，无论是传统多因子模型，还是本篇报告要介绍的交易型阿尔法体系，都尽可能的做到在因子检验、收益预测的过程中，剔除所有风格因素的影响，而在组合构建过程中，保持所有大类风格的中性化处理，最大程度的体现阿尔法模型的预测作用。

尽管，这种方式不一定是最优的，许多投资者或许认为承担风格收益的风险，获得的同样也是阿尔法收益。但正如我们上述提到的，投资需要的是逻辑和依据。对风格收益的预测就需要有独立的风格轮动模型，对市场涨跌的预测则需要有独立的择时模型。在我们的多因子体系中，我们大胆假设风格收益、市场收益难以预测，因此根据这样的逻辑，得到了因子模型和组合模型的推导结果，所有的策略设计过程做到逻辑严谨、有理可依、有据可靠。

阿尔法模型的定量评估和一致性问题是我们因子模型的理念基础，在此基础上，我们下面将介绍具体的交易型阿尔法新体系。

### 3. 基于短周期价量特征的多因子选股体系

#### 3.1. 短周期交易型阿尔法策略的构建思路

传统多因子模型的收益来源，往往集中于财务数据的挖掘、分析师一致预期数据的挖掘，而只有一小部分是中低频价量特征的数据，因此传统多因子模型亦可称为价值型阿尔法模型。

而在A股市场交易行为具有很强的随机性，机构投资者的交易效率较之成熟市场亦不十分有效，加之T+0、融券卖空等交易机制的缺乏，导致在短周期由于交易行为所产生的定价非有效十分常见。只要是投资者交

易产生的价格序列，就一定不是随机序列，而这一特征在 A 股更为明显。

因此，在这一领域内，超额收益有着更大的空间，量化模型也应该发挥其优势。我们将交易型阿尔法体系与传统阿尔法模型体系的特点做了简单的对比，具体为：

1. **超额收益来源：**传统多因子模型的超额收益来源多为股票内在价值，而交易型阿尔法体系的超额收益来源则是交易行为；
2. **因子数据来源：**传统多因子模型的因子数据来源集中于财务报表、分析师报告、及少部分中低频价量数据，而交易型阿尔法体系的因子数据则全部来自价格、成交数据（本篇报告中，我们仅采用日频率价量数据，在后续的研究中日内价量数据也将被广泛运用）；
3. **因子有效周期：**传统多因子模型的因子有效周期较长，财务类因子的预测周期超过 1 个月甚至更长；而交易型阿尔法体系的因子有效周期则较短，一般都在 1 周以内，但是因子预测的显著性更强；
4. **有效因子数量：**传统多因子模型的因子数量较少，并且更习惯于将多个有一定相关性的因子合成处理；而交易型阿尔法体系的因子数量则较多，不同因子贡献的边际收益较为独立；
5. **策略交易频率：**传统多因子模型的交易频率为季度、月频率、周频率等中低交易频率，而交易型阿尔法体系的交易频率则为日内、日度、隔日等中高频率（本篇报告中，我们将策略设计为 T+2 日换仓）；
6. **组合换手率：**传统多因子模型的整体换手率偏低，策略年化双边换手率约为 10 倍左右，而交易型阿尔法体系的整体换手率偏高，策略年化双边换手率约为 30-50 倍左右（换手率与客观交易成本呈负相关，我们将在后面的研究中具体介绍）；
7. **受市场风格影响：**传统多因子模型的收益与市场风格有一定的相关性，这不仅是由于因子选择、组合构建的偏差所导致，同样与市场环境、因子底层逻辑有关，而交易型阿尔法体系与市场风格的相关性则相对较低；
8. **策略内在收益率：**传统多因子模型的内在收益率较低，而交易型阿尔法体系的因子预测显著性更强，因而策略内在收益率较高。
9. **策略可容纳规模：**传统多因子模型的容量规模较大，而交易型阿尔法体系的策略容量相对较低，对交易效率的要求相对较高。

表 1：交易型阿尔法体系与传统多因子模型特征对比

	传统多因子模型	交易型阿尔法体系
超额收益来源	股票价值	交易行为
因子数据来源	财务报表、分析师研报	价格、成交量
因子有效周期	较长	较短
有效因子数量	较少	较多
策略交易频率	月频、周频	日频、隔日频
组合换手率	较低	较高
受市场风格影响	明显	不明显
策略内在收益率	较低	较高
策略可容纳规模	较大	较低

数据来源：国泰君安证券研究

### 3.2. 一些显著的价格特征举例

本节中，我们将对一些显著的、便于直观观察的价格特征进行举例说明。在后期的因子构建中，我们将对这些 底层因子进行一些衍生、变化、合成，使得策略更加适应市场的实战环境。

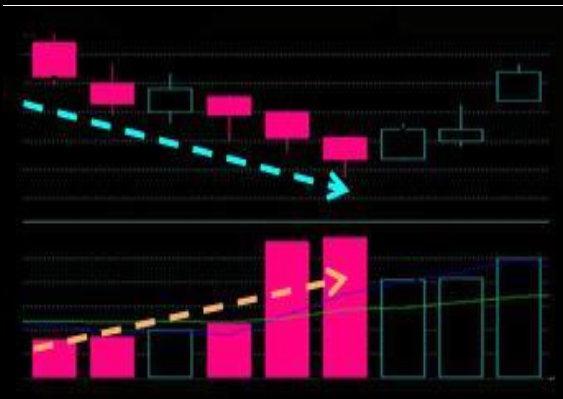
我们在每一期的股票截面上，计算全部个股的价格特征因子值，随后进行统计显著性检验。具体的，我们将检验风格因子中性后的因子收益率、因子收益率  $IR$ 、因子相关系数和因子相关系数  $IR$  ( $Factor Return$ 、 $Factor Return IR$ 、 $IC$ 、 $ICIR$ )，其中检验时间均为 2010 年 1 月至 2017 年 4 月。

#### 举例 1：价量背离

**经验观察：**短周期内成交量逐步提升，价格不断下降；或成交量逐步下降，价格不断提升。

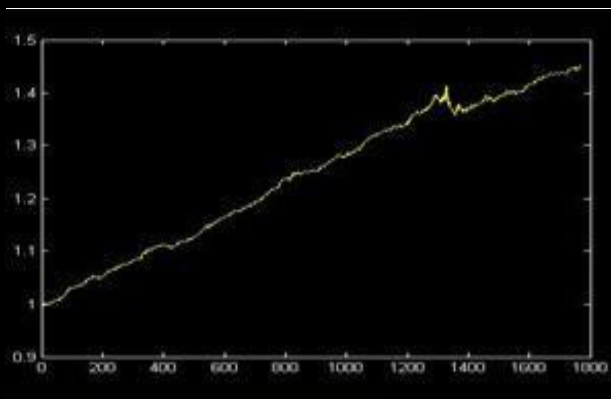
**因子定义：**  $Alpha_t^i = -1 * corr(vwap_{t-d:t}^i, volume_{t-d:t}^i)$

图 1 价量背离



数据来源：国泰君安证券研究

图 2 累计因子收益率



数据来源：国泰君安证券研究

以短周期内个股均价序列与成交量序列的相关系数衡量价量背离的程度，相关系数越低，后期实现超额收益的概率越高，累计因子收益率仅在股灾期间发生波动，其余时间段均较为显著。

表 2：因子显著性检验

因子定义	$Factor Return$	$Factor Return IR$	$IC$	$IC IR$
价量背离	5.27%	3.78	0.003	2.65

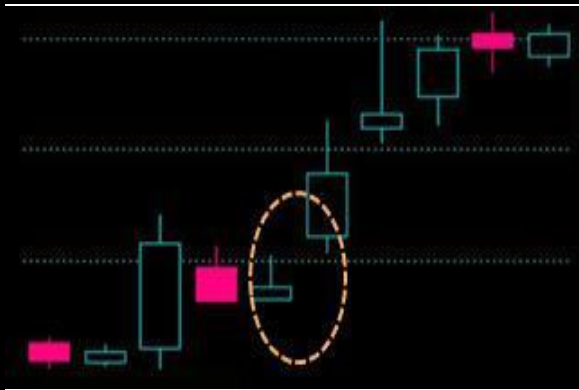
#### 举例 2：开盘缺口

**经验观察：**当日个股跳空高开或低开。

**因子定义：**  $Alpha_t^i = Open\_price_t^i / Close\_price_{t-1}^i$

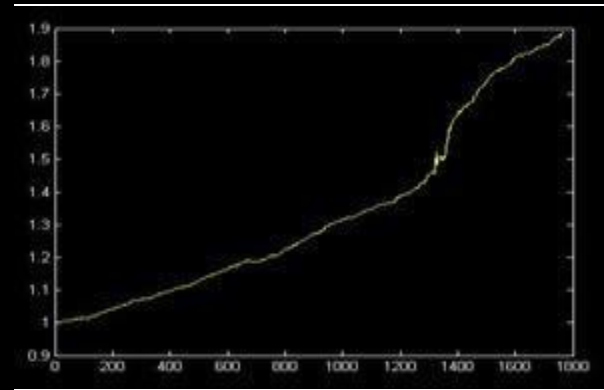


图 3 开盘缺口



数据来源：国泰君安证券研究

图 4 累计因子收益率



数据来源：国泰君安证券研究

当日个股开盘缺口，短周期内有较强的动量效应，累计因子收益率较为显著。

表 3: 因子显著性检验

因子定义	Factor Return	Factor Return IR	IC	IC IR
开盘缺口	9.00%	4.83	0.005	2.94

### 举例 3: 异常成交量

经验观察：当日成交量较短周期均值异常放大、减小。

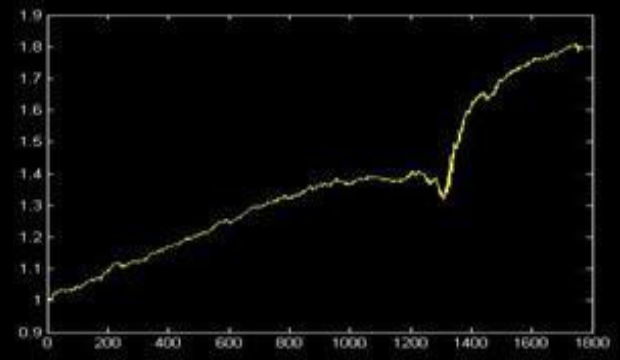
因子定义： $\text{Alpha}_t^i = -1 * \text{Volume}_t^i / \text{mean}(\text{Volume}_{t-d:t}^i)$

图 5 异常成交量



数据来源：国泰君安证券研究

图 6 累计因子收益率



数据来源：国泰君安证券研究

个股成交量的异常变化往往预示着阶段性反转，尤其是异常缩量，是阶段底部的特征。成交量的因子处理过程中应注意区分由于涨跌停导致的成交量异常变化。

表 4: 因子显著性检验

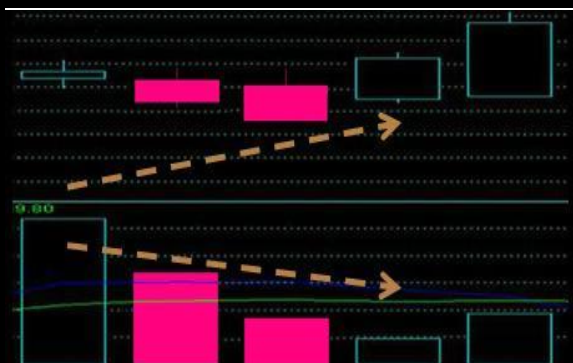
因子定义	Factor Return	Factor Return IR	IC	IC IR
异常成交量	8.35%	2.01	0.005	1.71

#### 举例 4：量幅背离

经验观察：短周期内成交量逐步提升，振幅不断下降；或成交量逐步下降，振幅不断提升。

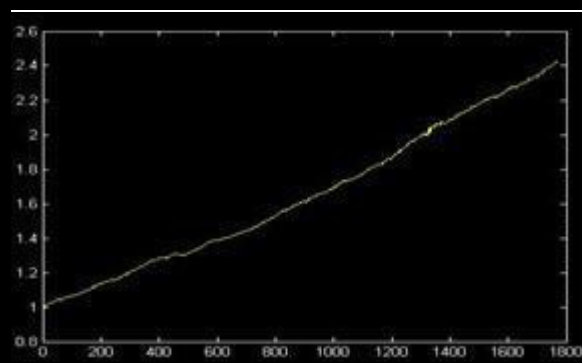
因子定义： $\text{Alpha}_t^i = -1 * \text{corr}(\text{high}_{t-d:t}^i / \text{low}_{t-d:t}^i, \text{volume}_{t-d:t}^i)$

图 7 量幅背离



数据来源：国泰君安证券研究

图 8 累计因子收益率



数据来源：国泰君安证券研究

以短周期内个股振幅序列（最高价比最低价）与成交量序列的相关系数衡量价量背离的程度，相关系数越低，后期实现超额收益的概率越高，累计因子收益率是所有因子中稳定性最强的。

表 5：因子显著性检验

因子定义	Factor Return	Factor Return IR	IC	IC IR
量幅背离	12.52%	8.39	0.007	4.26

对价量特征进行因子化处理的最大优势在于，避免了利用单一模式在时间序列上进行买点和卖点的选择，因为这牵涉到开平仓参数的敏感性问题等，会带来很大的不确定性和参数过拟合的问题。

价量形态的精髓在于，以当前市场的运行特征，寻找价格运行的规律。如果说传统多因子模型更加重视因子背后的价值投资逻辑，那么交易型阿尔法体系则更加重视交易行为背后的规律显著性，从某种意义上而言，这恰恰是最为直白朴素的投资逻辑。

### 3.3. 短周期价量因子体系构建

本篇报告中，我们总共构建了 191 个基于价格和成交量数据的短周期阿尔法因子，其中数据维度均为日频率交易数据。本节中，我们对所有因子的构建给出了详细的定义方式，而在之后的章节中，我们将对因子体系的特征进行分析。因子明细详见下表，部分函数表达详见附录 2：

表 6: 因子明细:

因子顺序	因子构建方式
Alpha1	$(-1 * CORR(RANK(DELTA(LOG(VOLUME), 1)), RANK(((CLOSE - OPEN) / OPEN)), 6))$
Alpha2	$(-1 * DELTA((((CLOSE - LOW) - (HIGH - CLOSE)) / (HIGH - LOW)), 1))$
Alpha3	$SUM((CLOSE=DELAY(CLOSE,1)?0:CLOSE-(CLOSE>DELAY(CLOSE,1)?MIN(LOW,DELAY(CLOSE,1)):MAX(HIGH,DELAY(CLOSE,1))))),6)$
Alpha4	$((((SUM(CLOSE, 8) / 8) + STD(CLOSE, 8)) < (SUM(CLOSE, 2) / 2)) ? (-1 * 1) : (((SUM(CLOSE, 2) / 2) < ((SUM(CLOSE, 8) / 8) - STD(CLOSE, 8))) ? 1 : (((1 < (VOLUME / MEAN(VOLUME,20)))    ((VOLUME / MEAN(VOLUME,20)) == 1)) ? 1 : (-1 * 1))))$
Alpha5	$(-1 * TSMAX(CORR(TSRANK(VOLUME, 5), TSRANK(HIGH, 5), 5), 3))$
Alpha6	$(RANK(SIGN(DELTA((((OPEN * 0.85) + (HIGH * 0.15))), 4)))) * -1)$
Alpha7	$((RANK(MAX((VWAP - CLOSE), 3)) + RANK(MIN((VWAP - CLOSE), 3))) * RANK(DELTA(VOLUME, 3)))$
Alpha8	$RANK(DELTA((((HIGH + LOW) / 2) * 0.2) + (VWAP * 0.8)), 4) * -1)$
Alpha9	$SMA(((HIGH+LOW)/2-(DELAY(HIGH,1)+DELAY(LOW,1))/2)*(HIGH-LOW)/VOLUME,7,2)$
Alpha10	$(RANK(MAX(((RET < 0) ? STD(RET, 20) : CLOSE)^2),5))$
Alpha11	$SUM((((CLOSE-LOW)-(HIGH-CLOSE))/HIGH-LOW)*VOLUME,6)$
Alpha12	$(RANK((OPEN - (SUM(VWAP, 10) / 10)))) * (-1 * (RANK(ABS((CLOSE - VWAP))))))$
Alpha13	$((HIGH * LOW)^{0.5} - VWAP)$
Alpha14	$CLOSE-DELAY(CLOSE,5)$
Alpha15	$OPEN/DELAY(CLOSE,1)-1$
Alpha16	$(-1 * TSMAX(RANK(CORR(RANK(VOLUME), RANK(VWAP), 5)), 5))$
Alpha17	$RANK((VWAP - MAX(VWAP, 15)))^DELTA(CLOSE, 5)$
Alpha18	$CLOSE/DELAY(CLOSE,5)$
Alpha19	$(CLOSE<DELAY(CLOSE,5)?(CLOSE-DELAY(CLOSE,5))/DELAY(CLOSE,5):(CLOSE=DELAY(CLOSE,5)?0:(CLOSE-DELAY(CLOSE,5))/CLOSE))$
Alpha20	$(CLOSE-DELAY(CLOSE,6))/DELAY(CLOSE,6)*100$
Alpha21	$REGBETA(MEAN(CLOSE,6),SEQUENCE(6))$
Alpha22	$SMEAN(((CLOSE-MEAN(CLOSE,6))/MEAN(CLOSE,6)-DELAY((CLOSE-MEAN(CLOSE,6))/MEAN(CLOSE,6),3)),12,1)$
Alpha23	$SMA((CLOSE>DELAY(CLOSE,1)?STD(CLOSE,20):0),20,1)/(SMA((CLOSE>DELAY(CLOSE,1)?STD(CLOSE,20):0),20,1)+SMA((CLOSE<=DELAY(CLOSE,1)?STD(CLOSE,20):0),20,1))*100$
Alpha24	$SMA(CLOSE-DELAY(CLOSE,5),5,1)$
Alpha25	$((-1 * RANK((DELTA(CLOSE, 7) * (1 - RANK(DECAYLINEAR((VOLUME / MEAN(VOLUME,20)), 9)))))) * (1 + RANK(SUM(RET, 250))))$
Alpha26	$((((SUM(CLOSE, 7) / 7) - CLOSE)) + ((CORR(VWAP, DELAY(CLOSE, 5), 230))))$
Alpha27	$WMA((CLOSE-DELAY(CLOSE,3))/DELAY(CLOSE,3)*100+(CLOSE-DELAY(CLOSE,6))/DELAY(CLOSE,6)*100,12)$
Alpha28	$3*SMA((CLOSE-TSMIN(LOW,9))/(TSMAX(HIGH,9)-TSMIN(LOW,9))*100,3,1)-2*SMA(SMA((CLOSE-TSMIN(LOW,9))/(MAX(HIGH,9)-TSMAX(LOW,9))*100,3,1),3,1)$
Alpha29	$(CLOSE-DELAY(CLOSE,6))/DELAY(CLOSE,6)*VOLUME$
Alpha30	$WMA((REGRESI(CLOSE/DELAY(CLOSE)-1,MKT,SMB,HML, 60))^2,20)$
Alpha31	$(CLOSE-MEAN(CLOSE,12))/MEAN(CLOSE,12)*100$
Alpha32	$(-1 * SUM(RANK(CORR(RANK(HIGH), RANK(VOLUME), 3)), 3))$
Alpha33	$(((-1 * TSMIN(LOW, 5)) + DELAY(TSMIN(LOW, 5), 5)) * RANK(((SUM(RET, 240) - SUM(RET, 20)) / 220))) * TSRANK(VOLUME, 5))$
Alpha34	$MEAN(CLOSE,12)/CLOSE$
Alpha35	$(MIN(RANK(DECAYLINEAR(DELTA(OPEN, 1), 15)), RANK(DECAYLINEAR(CORR((VOLUME), ((OPEN * 0.65) + (OPEN * 0.35)), 17),7)))) * -1)$

Alpha36	RANK(SUM(CORR(RANK(VOLUME), RANK(VWAP)), 6), 2)
Alpha37	(-1 * RANK(((SUM(OPEN, 5) * SUM(RET, 5)) - DELAY((SUM(OPEN, 5) * SUM(RET, 5)), 10))))
Alpha38	((SUM(HIGH, 20) / 20) < HIGH) ? (-1 * DELTA(HIGH, 2)) : 0)
Alpha39	((RANK(DECAYLINEAR(DELTA((CLOSE), 2), 8)) - RANK(DECAYLINEAR(CORR(((VWAP * 0.3) + (OPEN * 0.7)), SUM(MEAN(VOLUME, 180), 37), 14), 12))) * -1)
Alpha40	SUM((CLOSE > DELAY(CLOSE, 1) ? VOLUME : 0), 26) / SUM((CLOSE <= DELAY(CLOSE, 1) ? VOLUME : 0), 26) * 100
Alpha41	(RANK(MAX(DELTA((VWAP), 3), 5)) * -1)
Alpha42	((-1 * RANK(STD(HIGH, 10))) * CORR(HIGH, VOLUME, 10))
Alpha43	SUM((CLOSE > DELAY(CLOSE, 1) ? VOLUME : (CLOSE < DELAY(CLOSE, 1) ? -VOLUME : 0)), 6)
Alpha44	(TSRANK(DECAYLINEAR(CORR(((LOW)), MEAN(VOLUME, 10), 7), 6), 4) + TSRANK(DECAYLINEAR(DELTA((VWAP), 3), 10), 15))
Alpha45	(RANK(DELTA((((CLOSE * 0.6) + (OPEN * 0.4))), 1)) * RANK(CORR(VWAP, MEAN(VOLUME, 150), 15)))
Alpha46	(MEAN(CLOSE, 3) + MEAN(CLOSE, 6) + MEAN(CLOSE, 12) + MEAN(CLOSE, 24)) / (4 * CLOSE)
Alpha47	SMA((TSMAX(HIGH, 6) - CLOSE) / (TSMAX(HIGH, 6) - TSMIN(LOW, 6)) * 100, 9, 1)
Alpha48	(-1 * ((RANK(((SIGN((CLOSE - DELAY(CLOSE, 1))) + SIGN((DELAY(CLOSE, 1) - DELAY(CLOSE, 2)))) + SIGN((DELAY(CLOSE, 2) - DELAY(CLOSE, 3)))))) * SUM(VOLUME, 5)) / SUM(VOLUME, 20))
Alpha49	SUM(((HIGH + LOW) >= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12) / (SUM(((HIGH + LOW) >= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12) + SUM(((HIGH + LOW) <= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12))
Alpha50	SUM(((HIGH + LOW) <= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12) / (SUM(((HIGH + LOW) <= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12) + SUM(((HIGH + LOW) >= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12) - SUM(((HIGH + LOW) >= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12) / (SUM(((HIGH + LOW) >= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12) + SUM(((HIGH + LOW) <= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12))
Alpha51	SUM(((HIGH + LOW) <= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12) / (SUM(((HIGH + LOW) <= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12) + SUM(((HIGH + LOW) >= (DELAY(HIGH, 1) + DELAY(LOW, 1))) ? 0 : MAX(ABS(HIGH - DELAY(HIGH, 1)), ABS(LOW - DELAY(LOW, 1))), 12))
Alpha52	SUM(MAX(0, HIGH - DELAY((HIGH + LOW + CLOSE) / 3, 1)), 26) / SUM(MAX(0, DELAY((HIGH + LOW + CLOSE) / 3, 1) - LOW, 26) * 100
Alpha53	COUNT(CLOSE > DELAY(CLOSE, 1), 12) / 12 * 100
Alpha54	(-1 * RANK((STD(ABS(CLOSE - OPEN)) + (CLOSE - OPEN)) + CORR(CLOSE, OPEN, 10)))
Alpha55	SUM(16 * (CLOSE - DELAY(CLOSE, 1)) + (CLOSE - OPEN) / 2 + DELAY(CLOSE, 1) - DELAY(OPEN, 1)) / ((ABS(HIGH - DELAY(CLOSE, 1)) > ABS(LOW - DELAY(CLOSE, 1))) & ABS(HIGH - DELAY(CLOSE, 1)) > ABS(HIGH - DELAY(LOW, 1))) ? ABS(HIGH - DELAY(CLOSE, 1)) + ABS(LOW - DELAY(CLOSE, 1)) : 2 * ABS(DELAY(CLOSE, 1) - DELAY(OPEN, 1)) / 4 : (ABS(LOW - DELAY(CLOSE, 1)) > ABS(HIGH - DELAY(LOW, 1))) & ABS(LOW - DELAY(CLOSE, 1)) > ABS(HIGH - DELAY(CLOSE, 1))) ? ABS(LOW - DELAY(CLOSE, 1)) + ABS(HIGH - DELAY(CLOSE, 1)) : 2 * ABS(DELAY(CLOSE, 1) - DELAY(OPEN, 1)) / 4 : ABS(HIGH - DELAY(LOW, 1)) + ABS(DELAY(CLOSE, 1) - DELAY(OPEN, 1)) / 4))) * MAX(ABS(HIGH - DELAY(CLOSE, 1)), ABS(LOW - DELAY(CLOSE, 1))), 20)
Alpha56	(RANK((OPEN - TSMIN(OPEN, 12))) < RANK((RANK(CORR(SUM(((HIGH + LOW) / 2), 19), SUM(MEAN(VOLUME, 40), 19), 13))^5)))
Alpha57	SMA((CLOSE - TSMIN(LOW, 9)) / (TSMAX(HIGH, 9) - TSMIN(LOW, 9)) * 100, 3, 1)
Alpha58	COUNT(CLOSE > DELAY(CLOSE, 1), 20) / 20 * 100

Alpha59	$SUM((CLOSE=DELAY(CLOSE,1)?0:CLOSE-(CLOSE>DELAY(CLOSE,1)?MIN(LOW,DELAY(CLOSE,1)):MAX(HIGH,DELAY(CLOSE,1))))),20)$
Alpha60	$SUM(((CLOSE-LOW)-(HIGH-CLOSE))/(HIGH-LOW).*VOLUME,20)$
Alpha61	$(MAX(RANK(DECAYLINEAR(DELTA(VWAP, 1), 12)), RANK(DECAYLINEAR(RANK(CORR((LOW),MEAN(VOLUME,80), 8)), 17))) * -1)$
Alpha62	$(-1 * CORR(HIGH, RANK(VOLUME), 5))$
Alpha63	$SMA(MAX(CLOSE-DELAY(CLOSE,1),0),6,1)/SMA(ABS(CLOSE-DELAY(CLOSE,1)),6,1)*100$
Alpha64	$(MAX(RANK(DECAYLINEAR(CORR(RANK(VWAP), RANK(VOLUME), 4), 4)), RANK(DECAYLINEAR(MAX(CORR(RANK(CLOSE), RANK(MEAN(VOLUME,60)), 4), 13), 14))) * -1)$
Alpha65	$MEAN(CLOSE,6)/CLOSE$
Alpha66	$(CLOSE-MEAN(CLOSE,6))/MEAN(CLOSE,6)*100$
Alpha67	$SMA(MAX(CLOSE-DELAY(CLOSE,1),0),24,1)/SMA(ABS(CLOSE-DELAY(CLOSE,1)),24,1)*100$
Alpha68	$SMA(((HIGH+LOW)/2-(DELAY(HIGH,1)+DELAY(LOW,1))/2)*(HIGH-LOW)/VOLUME,15,2)$
Alpha69	$(SUM(DTM,20)>SUM(DBM,20) ? (SUM(DTM,20)-SUM(DBM,20))/SUM(DTM,20) : (SUM(DTM,20)=SUM(DBM,20) ? 0 : (SUM(DTM,20)-SUM(DBM,20))/SUM(DBM,20)))$
Alpha70	$STD(AMOUNT,6)$
Alpha71	$(CLOSE-MEAN(CLOSE,24))/MEAN(CLOSE,24)*100$
Alpha72	$SMA((TSMAX(HIGH,6)-CLOSE)/(TSMAX(HIGH,6)-TSMIN(LOW,6))*100,15,1)$
Alpha73	$((TSRANK(DECAYLINEAR(DECAYLINEAR(CORR((CLOSE), VOLUME, 10), 16), 4), 5) - RANK(DECAYLINEAR(CORR(VWAP, MEAN(VOLUME,30), 4),3))) * -1)$
Alpha74	$(RANK(CORR(SUM(((LOW * 0.35) + (VWAP * 0.65)), 20), SUM(MEAN(VOLUME,40), 20), 7))) + RANK(CORR(RANK(VWAP), RANK(VOLUME), 6)))$
Alpha75	$COUNT(CLOSE>OPEN \& BANCHMARKINDEXCLOSE<BANCHMARKINDEXOPEN,50)/COUNT(BANCHMARKINDEXCLOSE<BANCHMARKINDEXOPEN,50)$
Alpha76	$STD(ABS((CLOSE/DELAY(CLOSE,1)-1))/VOLUME,20)/MEAN(ABS((CLOSE/DELAY(CLOSE,1)-1))/VOLUME,20)$
Alpha77	$MIN(RANK(DECAYLINEAR((((HIGH + LOW) / 2) + HIGH) - (VWAP + HIGH)), 20)), RANK(DECAYLINEAR(CORR(((HIGH + LOW) / 2), MEAN(VOLUME,40), 3), 6)))$
Alpha78	$((HIGH+LOW+CLOSE)/3-MA((HIGH+LOW+CLOSE)/3,12))/(0.015*MEAN(ABS(CLOSE-MEAN((HIGH+LOW+CLOSE)/3,12)),12))$
Alpha79	$SMA(MAX(CLOSE-DELAY(CLOSE,1),0),12,1)/SMA(ABS(CLOSE-DELAY(CLOSE,1)),12,1)*100$
Alpha80	$(VOLUME-DELAY(VOLUME,5))/DELAY(VOLUME,5)*100$
Alpha81	$SMA(VOLUME,21,2)$
Alpha82	$SMA((TSMAX(HIGH,6)-CLOSE)/(TSMAX(HIGH,6)-TSMIN(LOW,6))*100,20,1)$
Alpha83	$(-1 * RANK(COVIANCE(RANK(HIGH), RANK(VOLUME), 5)))$
Alpha84	$SUM((CLOSE>DELAY(CLOSE,1)?VOLUME:(CLOSE<DELAY(CLOSE,1)?-VOLUME:0)),20)$
Alpha85	$(TSRANK((VOLUME / MEAN(VOLUME,20)), 20) * TSRANK((-1 * DELTA(CLOSE, 7)), 8))$
Alpha86	$((0.25 < (((DELAY(CLOSE, 20) - DELAY(CLOSE, 10)) / 10) - ((DELAY(CLOSE, 10) - CLOSE) / 10))) ? (-1 * 1) : (((((DELAY(CLOSE, 20) - DELAY(CLOSE, 10)) / 10) - ((DELAY(CLOSE, 10) - CLOSE) / 10)) < 0) ? 1 : ((-1 * 1) * (CLOSE - DELAY(CLOSE, 1)))))$
Alpha87	$((RANK(DECAYLINEAR(DELTA(VWAP, 4), 7)) + TSRANK(DECAYLINEAR((((LOW * 0.9) + (LOW * 0.1)) - VWAP) / (OPEN - ((HIGH + LOW) / 2))), 11), 7)) * -1)$
Alpha88	$(CLOSE-DELAY(CLOSE,20))/DELAY(CLOSE,20)*100$
Alpha89	$2*(SMA(CLOSE,13,2)-SMA(CLOSE,27,2)-SMA(SMA(CLOSE,13,2)-SMA(CLOSE,27,2),10,2))$
Alpha90	$( RANK(CORR(RANK(VWAP), RANK(VOLUME), 5)) * -1)$
Alpha91	$((RANK((CLOSE - MAX(CLOSE, 5)))*RANK(CORR((MEAN(VOLUME,40)), LOW, 5))) * -1)$





Alpha92	(MAX(RANK(DECAYLINEAR(DELTA(((CLOSE * 0.35) + (VWAP * 0.65)), 2), 3)), TSRANK(DECAYLINEAR(ABS(CORR((MEAN(VOLUME,180)), CLOSE, 13)), 5), 15)) * -1)
Alpha93	SUM((OPEN >= DELAY(OPEN, 1) ? 0 : MAX((OPEN - LOW), (OPEN - DELAY(OPEN, 1)))) , 20)
Alpha94	SUM((CLOSE > DELAY(CLOSE, 1) ? VOLUME : (CLOSE < DELAY(CLOSE, 1) ? -VOLUME : 0)), 30)
Alpha95	STD(AMOUNT, 20)
Alpha96	SMA(SMA((CLOSE - TSMIN(LOW, 9)) / (TSMAX(HIGH, 9) - TSMIN(LOW, 9)) * 100, 3, 1), 3, 1)
Alpha97	STD(VOLUME, 10)
Alpha98	(((((DELTA((SUM(CLOSE, 100) / 100), 100) / DELAY(CLOSE, 100)) < 0.05)    ((DELTA((SUM(CLOSE, 100) / 100), 100) / DELAY(CLOSE, 100)) == 0.05)) ? (-1 * (CLOSE - TSMIN(CLOSE, 100))) : (-1 * DELTA(CLOSE, 3))))
Alpha99	(-1 * RANK(COVIANCE(RANK(CLOSE), RANK(VOLUME), 5)))
Alpha100	STD(VOLUME, 20)
Alpha101	((RANK(CORR(CLOSE, SUM(MEAN(VOLUME, 30), 37), 15)) < RANK(CORR(RANK(((HIGH * 0.1) + (VWAP * 0.9))), RANK(VOLUME), 11))) * -1)
Alpha102	SMA(MAX(VOLUME - DELAY(VOLUME, 1), 0), 6, 1) / SMA(ABS(VOLUME - DELAY(VOLUME, 1)), 6, 1) * 100
Alpha103	((20 - LOWDAY(LOW, 20)) / 20) * 100
Alpha104	(-1 * (DELTA(CORR(HIGH, VOLUME, 5), 5) * RANK(STD(CLOSE, 20))))
Alpha105	(-1 * CORR(RANK(OPEN), RANK(VOLUME), 10))
Alpha106	CLOSE - DELAY(CLOSE, 20)
Alpha107	(((-1 * RANK((OPEN - DELAY(HIGH, 1)))) * RANK((OPEN - DELAY(CLOSE, 1)))) * RANK((OPEN - DELAY(LOW, 1))))
Alpha108	((RANK((HIGH - MIN(HIGH, 2))) ^ RANK(CORR((VWAP), (MEAN(VOLUME, 120)), 6))) * -1)
Alpha109	SMA(HIGH - LOW, 10, 2) / SMA(SMA(HIGH - LOW, 10, 2), 10, 2)
Alpha110	SUM(MAX(0, HIGH - DELAY(CLOSE, 1)), 20) / SUM(MAX(0, DELAY(CLOSE, 1) - LOW), 20) * 100
Alpha111	SMA(VOL * ((CLOSE - LOW) - (HIGH - CLOSE)) / (HIGH - LOW), 11, 2) - SMA(VOL * ((CLOSE - LOW) - (HIGH - CLOSE)) / (HIGH - L OW), 4, 2)
Alpha112	(SUM((CLOSE - DELAY(CLOSE, 1) > 0 ? CLOSE - DELAY(CLOSE, 1) : 0), 12) - SUM((CLOSE - DELAY(CLOSE, 1) < 0 ? ABS(CLOS E - DELAY(CLOSE, 1)) : 0), 12)) / (SUM((CLOSE - DELAY(CLOSE, 1) > 0 ? CLOSE - DELAY(CLOSE, 1) : 0), 12) + SUM((CLOSE - DE LAY(CLOSE, 1) < 0 ? ABS(CLOSE - DELAY(CLOSE, 1)) : 0), 12)) * 100
Alpha113	(-1 * ((RANK((SUM(DELAY(CLOSE, 5), 20) / 20)) * CORR(CLOSE, VOLUME, 2)) * RANK(CORR(SUM(CLOSE, 5), SUM(CLOSE, 20), 2))))
Alpha114	((RANK(DELAY(((HIGH - LOW) / (SUM(CLOSE, 5) / 5)), 2)) * RANK(RANK(VOLUME))) / (((HIGH - LOW) / (SUM(CLOSE, 5) / 5)) / (VWAP - CLOSE)))
Alpha115	(RANK(CORR(((HIGH * 0.9) + (CLOSE * 0.1)), MEAN(VOLUME, 30), 10)) ^ RANK(CORR(TSRANK(((HIGH + LOW) / 2), 4), TSRANK(VOLUME, 10), 7)))
Alpha116	REGBETA(CLOSE, SEQUENCE, 20)
Alpha117	((TSRANK(VOLUME, 32) * (1 - TSRANK(((CLOSE + HIGH) - LOW), 16))) * (1 - TSRANK(RET, 32)))
Alpha118	SUM(HIGH - OPEN, 20) / SUM(OPEN - LOW, 20) * 100
Alpha119	(RANK(DECAYLINEAR(CORR(VWAP, SUM(MEAN(VOLUME, 5), 26), 5), 7)) - RANK(DECAYLINEAR(TSRANK(MIN(CORR(RANK(OPEN), RANK(MEAN(VOLUME, 15))), 21), 9), 7), 8)))
Alpha120	(RANK((VWAP - CLOSE)) / RANK((VWAP + CLOSE)))
Alpha121	((RANK((VWAP - MIN(VWAP, 12))) ^ TSRANK(CORR(TSRANK(VWAP, 20), TSRANK(MEAN(VOLUME, 60), 2), 18), 3)) * -1)
Alpha122	(SMA(SMA(SMA(LOG(CLOSE), 13, 2), 13, 2), 13, 2) - DELAY(SMA(SMA(SMA(LOG(CLOSE), 13, 2), 13, 2), 13, 2), 1)) / DELAY(SM A(SMA(SMA(LOG(CLOSE), 13, 2), 13, 2), 13, 2), 1)
Alpha123	((RANK(CORR(SUM(((HIGH + LOW) / 2), 20), SUM(MEAN(VOLUME, 60), 20), 9)) < RANK(CORR(LOW, VOLUME, 6))) * -1)
Alpha124	(CLOSE - VWAP) / DECAYLINEAR(RANK(TSMAX(CLOSE, 30)), 2)

Alpha125	$(\text{RANK}(\text{DECAYLINEAR}(\text{CORR}(\text{VWAP}), \text{MEAN}(\text{VOLUME}, 80), 17), 20)) / \text{RANK}(\text{DECAYLINEAR}(\text{DELTA}(((\text{CLOSE} * 0.5) + (\text{VWAP} * 0.5))), 3), 16)))$
Alpha126	$(\text{CLOSE} + \text{HIGH} + \text{LOW}) / 3$
Alpha127	$(\text{MEAN}(((100 * (\text{CLOSE} - \text{MAX}(\text{CLOSE}, 12)) / (\text{MAX}(\text{CLOSE}, 12)))^2))^{\wedge}(1/2) - 100 - (100 / (1 + \text{SUM}(((\text{HIGH} + \text{LOW} + \text{CLOSE}) / 3 > \text{DELAY}((\text{HIGH} + \text{LOW} + \text{CLOSE}) / 3, 1) ? (\text{HIGH} + \text{LOW} + \text{CLOSE}) / 3 * \text{VOLUME} : 0), 14)) / \text{SUM}(((\text{HIGH} + \text{LOW} + \text{CLOSE}) / 3 < \text{DELAY}((\text{HIGH} + \text{LOW} + \text{CLOSE}) / 3, 1) ? (\text{HIGH} + \text{LOW} + \text{CLOSE}) / 3 * \text{VOLUME} : 0), 14)))$
Alpha129	$\text{SUM}((\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 1) < 0 ? \text{ABS}(\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 1)) : 0), 12)$
Alpha130	$(\text{RANK}(\text{DECAYLINEAR}(\text{CORR}(((\text{HIGH} + \text{LOW}) / 2), \text{MEAN}(\text{VOLUME}, 40), 9), 10)) / \text{RANK}(\text{DECAYLINEAR}(\text{CORR}(\text{RANK}(\text{VWAP}), \text{RANK}(\text{VOLUME}), 7), 3)))$
Alpha131	$(\text{RANK}(\text{DELTA}(\text{VWAP}, 1))^{\wedge} \text{TSRANK}(\text{CORR}(\text{CLOSE}, \text{MEAN}(\text{VOLUME}, 50), 18), 18))$
Alpha132	$\text{MEAN}(\text{AMOUNT}, 20)$
Alpha133	$((20 - \text{HIGHDAY}(\text{HIGH}, 20)) / 20) * 100 - ((20 - \text{LOWDAY}(\text{LOW}, 20)) / 20) * 100$
Alpha134	$(\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 12)) / \text{DELAY}(\text{CLOSE}, 12) * \text{VOLUME}$
Alpha135	$\text{SMA}(\text{DELAY}(\text{CLOSE} / \text{DELAY}(\text{CLOSE}, 20), 1), 20, 1)$
Alpha136	$((-1 * \text{RANK}(\text{DELTA}(\text{RET}, 3))) * \text{CORR}(\text{OPEN}, \text{VOLUME}, 10))$ $16 * (\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 1) + (\text{CLOSE} - \text{OPEN}) / 2 + \text{DELAY}(\text{CLOSE}, 1) - \text{DELAY}(\text{OPEN}, 1)) / ((\text{ABS}(\text{HIGH} - \text{DELAY}(\text{CLOSE}, 1)) > \text{ABS}(\text{LOW} - \text{DELAY}(\text{CLOSE}, 1))) \&$ $\text{ABS}(\text{HIGH} - \text{DELAY}(\text{CLOSE}, 1)) > \text{ABS}(\text{HIGH} - \text{DELAY}(\text{LOW}, 1)) ? \text{ABS}(\text{HIGH} - \text{DELAY}(\text{CLOSE}, 1)) + \text{ABS}(\text{LOW} - \text{DELAY}(\text{CLOSE}, 1)) / 2 + \text{ABS}(\text{DELAY}(\text{CLOSE}, 1) - \text{DELAY}(\text{OPEN}, 1)) / 4 : (\text{ABS}(\text{LOW} - \text{DELAY}(\text{CLOSE}, 1)) > \text{ABS}(\text{HIGH} - \text{DELAY}(\text{LOW}, 1)) \&$ $\text{ABS}(\text{LOW} - \text{DELAY}(\text{CLOSE}, 1)) > \text{ABS}(\text{HIGH} - \text{DELAY}(\text{CLOSE}, 1)) ? \text{ABS}(\text{LOW} - \text{DELAY}(\text{CLOSE}, 1)) + \text{ABS}(\text{HIGH} - \text{DELAY}(\text{CLOSE}, 1)) / 2 + \text{ABS}(\text{DELAY}(\text{CLOSE}, 1) - \text{DELAY}(\text{OPEN}, 1)) / 4 : \text{ABS}(\text{HIGH} - \text{DELAY}(\text{LOW}, 1)) + \text{ABS}(\text{DELAY}(\text{CLOSE}, 1) - \text{DELAY}(\text{OPEN}, 1)) / 4))) * \text{MAX}(\text{ABS}(\text{HIGH} - \text{DELAY}(\text{CLOSE}, 1)), \text{ABS}(\text{LOW} - \text{DELAY}(\text{CLOSE}, 1)))$
Alpha138	$((\text{RANK}(\text{DECAYLINEAR}(\text{DELTA}(((\text{LOW} * 0.7) + (\text{VWAP} * 0.3))), 3), 20)) - \text{TSRANK}(\text{DECAYLINEAR}(\text{TSRANK}(\text{CORR}(\text{TSRANK}(\text{LOW}, 8), \text{TSRANK}(\text{MEAN}(\text{VOLUME}, 60), 17), 5), 19), 16), 7)) * -1)$
Alpha139	$(-1 * \text{CORR}(\text{OPEN}, \text{VOLUME}, 10))$
Alpha140	$\text{MIN}(\text{RANK}(\text{DECAYLINEAR}(((\text{RANK}(\text{OPEN}) + \text{RANK}(\text{LOW})) - (\text{RANK}(\text{HIGH}) + \text{RANK}(\text{CLOSE}))), 8)), \text{TSRANK}(\text{DECAYLINEAR}(\text{CORR}(\text{TSRANK}(\text{CLOSE}, 8), \text{TSRANK}(\text{MEAN}(\text{VOLUME}, 60), 20), 8), 7), 3))$
Alpha141	$(\text{RANK}(\text{CORR}(\text{RANK}(\text{HIGH}), \text{RANK}(\text{MEAN}(\text{VOLUME}, 15))), 9)) * -1)$
Alpha142	$(((-1 * \text{RANK}(\text{TSRANK}(\text{CLOSE}, 10))) * \text{RANK}(\text{DELTA}(\text{DELTA}(\text{CLOSE}, 1), 1))) * \text{RANK}(\text{TSRANK}((\text{VOLUME} / \text{MEAN}(\text{VOLUME}, 20)), 5)))$
Alpha143	$\text{CLOSE} > \text{DELAY}(\text{CLOSE}, 1) ? (\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 1)) / \text{DELAY}(\text{CLOSE}, 1) * \text{SELF} : \text{SELF}$
Alpha144	$\text{SUMIF}(\text{ABS}(\text{CLOSE} / \text{DELAY}(\text{CLOSE}, 1) - 1) / \text{AMOUNT}, 20, \text{CLOSE} < \text{DELAY}(\text{CLOSE}, 1)) / \text{COUNT}(\text{CLOSE} < \text{DELAY}(\text{CLOSE}, 1), 20)$
Alpha145	$(\text{MEAN}(\text{VOLUME}, 9) - \text{MEAN}(\text{VOLUME}, 26)) / \text{MEAN}(\text{VOLUME}, 12) * 100$
Alpha146	$\text{MEAN}((\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 1)) / \text{DELAY}(\text{CLOSE}, 1) - \text{SMA}((\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 1)) / \text{DELAY}(\text{CLOSE}, 1), 61, 2), 20) * ((\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 1)) / \text{DELAY}(\text{CLOSE}, 1) - \text{SMA}((\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 1)) / \text{DELAY}(\text{CLOSE}, 1), 61, 2)) / \text{SMA}(((\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 1)) / \text{DELAY}(\text{CLOSE}, 1) - ((\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 1)) / \text{DELAY}(\text{CLOSE}, 1) - \text{SMA}((\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 1)) / \text{DELAY}(\text{CLOSE}, 1), 61, 2)))^2, 60);$
Alpha147	$\text{REGBETA}(\text{MEAN}(\text{CLOSE}, 12), \text{SEQUENCE}(12))$
Alpha148	$((\text{RANK}(\text{CORR}((\text{OPEN}), \text{SUM}(\text{MEAN}(\text{VOLUME}, 60), 9), 6)) < \text{RANK}((\text{OPEN} - \text{TSMIN}(\text{OPEN}, 14)))) * -1)$ $\text{REGBETA}(\text{FILTER}(\text{CLOSE} / \text{DELAY}(\text{CLOSE}, 1) - 1, \text{BANCHMARKINDEXCLOSE} < \text{DELAY}(\text{BANCHMARKINDEXCLOSE}, 1)), \text{FILTER}(\text{BANCHMARKINDEXCLOSE} / \text{DELAY}(\text{BANCHMARKINDEXCLOSE}, 1) - 1, \text{BANCHMARKINDEXCLOSE} < \text{DELAY}(\text{BANCHMARKINDEXCLOSE}, 1)), 252)$
Alpha150	$(\text{CLOSE} + \text{HIGH} + \text{LOW}) / 3 * \text{VOLUME}$
Alpha151	$\text{SMA}(\text{CLOSE} - \text{DELAY}(\text{CLOSE}, 20), 20, 1)$



Alpha152	$SMA(MEAN(DELAY(SMA(DELAY(CLOSE/DELAY(CLOSE,9),1),9,1),1),12)-MEAN(DELAY(SMA(DELAY(CLOSE/DELAY(CLOSE,9),1),9,1),1),26),9,1)$
Alpha153	$(MEAN(CLOSE,3)+MEAN(CLOSE,6)+MEAN(CLOSE,12)+MEAN(CLOSE,24))/4$
Alpha154	$((VWAP - MIN(VWAP, 16))) < (CORR(VWAP, MEAN(VOLUME,180), 18)))$
Alpha155	$SMA(VOLUME,13,2)-SMA(VOLUME,27,2)-SMA(SMA(VOLUME,13,2)-SMA(VOLUME,27,2),10,2)$
Alpha156	$(MAX(RANK(DECAYLINEAR(DELTA(VWAP, 5), 3)), RANK(DECAYLINEAR(((DELTA(((OPEN * 0.15) + (LOW * 0.85)), 2) / ((OPEN * 0.15) + (LOW * 0.85))) * -1), 3))) * -1)$
Alpha157	$(MIN(PROD(RANK(RANK(LOG(SUM(TSMIN(RANK(RANK((-1 * RANK(DELTA((CLOSE - 1), 5))))), 2), 1))))), 1), 5) + TSRANK(DELAY((-1 * RET), 6), 5))$
Alpha158	$((HIGH-SMA(CLOSE,15,2))-(LOW-SMA(CLOSE,15,2)))/CLOSE$
Alpha159	$((CLOSE-SUM(MIN(LOW,DELAY(CLOSE,1)),6))/SUM(MAX(HIGH,DELAY(CLOSE,1))-MIN(LOW,DELAY(CLOSE,1)),6) * 12 * 24 + (CLOSE-SUM(MIN(LOW,DELAY(CLOSE,1)),12))/SUM(MAX(HIGH,DELAY(CLOSE,1))-MIN(LOW,DELAY(CLOSE,1)),12) * 6 * 24 + (CLOSE-SUM(MIN(LOW,DELAY(CLOSE,1)),24))/SUM(MAX(HIGH,DELAY(CLOSE,1))-MIN(LOW,DELAY(CLOSE,1)),24) * 6 * 24) * 100 / (6 * 12 + 6 * 24 + 12 * 24)$
Alpha160	$SMA((CLOSE <= DELAY(CLOSE,1)) ? STD(CLOSE,20) : 0), 20, 1)$
Alpha161	$MEAN(MAX(MAX((HIGH-LOW),ABS(DELAY(CLOSE,1)-HIGH)),ABS(DELAY(CLOSE,1)-LOW))),12)$
Alpha162	$(SMA(MAX(CLOSE-DELAY(CLOSE,1),0),12,1)/SMA(ABS(CLOSE-DELAY(CLOSE,1)),12,1)*100-MIN(SMA(MAX(CLOSE-DELAY(CLOSE,1),0),12,1)/SMA(ABS(CLOSE-DELAY(CLOSE,1)),12,1)*100,12))/(MAX(SMA(MAX(CLOSE-DELAY(CLOSE,1),0),12,1)/SMA(ABS(CLOSE-DELAY(CLOSE,1)),12,1)*100,12)-MIN(SMA(MAX(CLOSE-DELAY(CLOSE,1),0),12,1)/SMA(ABS(CLOSE-DELAY(CLOSE,1)),12,1)*100,12))$
Alpha163	$RANK(((((-1 * RET) * MEAN(VOLUME,20)) * VWAP) * (HIGH - CLOSE)))$
Alpha164	$SMA((((CLOSE > DELAY(CLOSE,1)) ? 1 / (CLOSE-DELAY(CLOSE,1)) : 1) - MIN((((CLOSE > DELAY(CLOSE,1)) ? 1 / (CLOSE-DELAY(CLOSE,1)) : 1), 12)) / (HIGH-LOW) * 100, 13, 2)$
Alpha165	$MAX(SUMAC(CLOSE-MEAN(CLOSE,48)))-MIN(SUMAC(CLOSE-MEAN(CLOSE,48)))/STD(CLOSE,48)$
Alpha166	$-20 * ((CLOSE/DELAY(CLOSE,1) - MEAN(CLOSE/DELAY(CLOSE,1),20)) / ((20-1) * (20-2) * (SUM((CLOSE/DELAY(CLOSE,1),20)^2,20))^1.5))$
Alpha167	$SUM((CLOSE-DELAY(CLOSE,1)) > 0 ? CLOSE-DELAY(CLOSE,1) : 0), 12)$
Alpha168	$(-1 * VOLUME / MEAN(VOLUME, 20))$
Alpha169	$SMA(MEAN(DELAY(SMA(CLOSE-DELAY(CLOSE,1),9,1),1),12)-MEAN(DELAY(SMA(CLOSE-DELAY(CLOSE,1),9,1),1),26),10,1)$
Alpha170	$((((RANK((1 / CLOSE)) * VOLUME) / MEAN(VOLUME,20)) * ((HIGH * RANK((HIGH - CLOSE))) / (SUM(HIGH, 5) / 5))) - RANK((VWAP - DELAY(VWAP, 5))))$
Alpha171	$((-1 * ((LOW - CLOSE) * (OPEN^5))) / ((CLOSE - HIGH) * (CLOSE^5)))$
Alpha172	$MEAN(ABS(SUM((LD > 0 & LD > HD) ? LD : 0, 14) * 100 / SUM(TR, 14) - SUM((HD > 0 & HD > LD) ? HD : 0, 14) * 100 / SUM(TR, 14)) / (SUM((LD > 0 & LD > HD) ? LD : 0, 14) * 100 / SUM(TR, 14) + SUM((HD > 0 & HD > LD) ? HD : 0, 14) * 100 / SUM(TR, 14)) * 100, 6)$
Alpha173	$3 * SMA(CLOSE, 13, 2) - 2 * SMA(SMA(CLOSE, 13, 2), 13, 2) + SMA(SMA(SMA(LOG(CLOSE), 13, 2), 13, 2), 13, 2);$
Alpha174	$SMA((CLOSE > DELAY(CLOSE,1)) ? STD(CLOSE,20) : 0), 20, 1)$
Alpha175	$MEAN(MAX(MAX((HIGH-LOW),ABS(DELAY(CLOSE,1)-HIGH)),ABS(DELAY(CLOSE,1)-LOW))),6)$
Alpha176	$CORR(RANK(((CLOSE - TSMIN(LOW, 12)) / (TSMAX(HIGH, 12) - TSMIN(LOW,12))))), RANK(VOLUME), 6)$
Alpha177	$((20-HIGHDAY(HIGH,20))/20)*100$
Alpha178	$(CLOSE-DELAY(CLOSE,1))/DELAY(CLOSE,1)*VOLUME$
Alpha179	$(RANK(CORR(VWAP, VOLUME, 4)) * RANK(CORR(RANK(LOW), RANK(MEAN(VOLUME,50)), 12)))$
Alpha180	$((MEAN(VOLUME,20) < VOLUME) ? ((-1 * TSRANK(ABS(DELTA(CLOSE, 7)), 60)) * SIGN(DELTA(CLOSE, 7)) : (-1 * VOLUME)))$

Alpha181	$SUM(((CLOSE/DELAY(CLOSE,1)-1)-MEAN((CLOSE/DELAY(CLOSE,1)-1),20))-(BANCHMARKINDEXCLOSE-MEAN(BANCHMARKINDEXCLOSE,20))^2,20)/SUM((BANCHMARKINDEXCLOSE-MEAN(BANCHMARKINDEXCLOSE,20))^3)$
Alpha182	$COUNT((CLOSE>OPEN \& BANCHMARKINDEXCLOSE>BANCHMARKINDEXOPEN)OR(CLOSE<OPEN \& BANCHMARKINDEXCLOSE<BANCHMARKINDEXOPEN),20)/20$
Alpha183	$MAX(SUMAC(CLOSE-MEAN(CLOSE,24)))-MIN(SUMAC(CLOSE-MEAN(CLOSE,24)))/STD(CLOSE,24)$
Alpha184	$(RANK(CORR(DELAY((OPEN - CLOSE), 1), CLOSE, 200)) + RANK((OPEN - CLOSE)))$
Alpha185	$RANK((-1 * ((1 - (OPEN / CLOSE))^2)))$
Alpha186	$(MEAN(ABS(SUM((LD>0 \& LD>HD)?LD:0,14)*100/SUM(TR,14)-SUM((HD>0 \& HD>LD)?HD:0,14)*100/SUM(TR,14))/(SUM((LD>0 \& LD>HD)?LD:0,14)*100/SUM(TR,14)+SUM((HD>0 \& HD>LD)?HD:0,14)*100/SUM(TR,14)))*100,6)+DELAY(MEAN(ABS(SUM((LD>0 \& LD>HD)?LD:0,14)*100/SUM(TR,14)-SUM((HD>0 \& HD>LD)?HD:0,14)*100/SUM(TR,14))/(SUM((LD>0 \& LD>HD)?LD:0,14)*100/SUM(TR,14)+SUM((HD>0 \& HD>LD)?HD:0,14)*100/SUM(TR,14))*100,6))/2$
Alpha187	$SUM((OPEN<=DELAY(OPEN,1)?0:MAX((HIGH-OPEN),(OPEN-DELAY(OPEN,1))))),20)$
Alpha188	$((HIGH-LOW-SMA(HIGH-LOW,11,2))/SMA(HIGH-LOW,11,2))*100$
Alpha189	$MEAN(ABS(CLOSE-MEAN(CLOSE,6)),6)$
Alpha190	$LOG((COUNT(CLOSE/DELAY(CLOSE)-1>((CLOSE/DELAY(CLOSE,19))^(1/20)-1),20)-1)*(SUMIF(((CLOSE/DELAY(CLOSE)-1-(CLOSE/DELAY(CLOSE,19))^(1/20)-1))^2,20,CLOSE/DELAY(CLOSE)-1<(CLOSE/DELAY(CLOSE,19))^(1/20)-1))/((COUNT((CLOSE/DELAY(CLOSE)-1<(CLOSE/DELAY(CLOSE,19))^(1/20)-1),20)*(SUMIF((CLOSE/DELAY(CLOSE)-1-((CLOSE/DELAY(CLOSE,19))^(1/20)-1))^2,20,CLOSE/DELAY(CLOSE)-1>(CLOSE/DELAY(CLOSE,19))^(1/20)-1))))$
Alpha191	$((CORR(MEAN(VOLUME,20), LOW, 5) + ((HIGH + LOW) / 2)) - CLOSE)$

数据来源：国泰君安证券研究

### 3.4. 因子检验及相关特征

本节，我们将通过一些特征分析，使得投资者对我们的短周期交易型因子体系有更深入的认识。

#### 3.4.1. 因子体系的显著性检验、相关性及有效周期

首先，我们对所有单因子进行有效性及预测周期检验，进而判断整个因子体系的可靠程度以及可预测的未来极限周期。其中，单因子显著性及有效周期检验的具体步骤如下：

**Step1:** 正交标准化，取目标因子残差截面  $\varepsilon_K^t$ ，即

$$X_K^t = \beta_{industry} X_{industry} + \beta_{style} X_{style} + \varepsilon_K^t$$

**Step2:** 针对给定预测周期  $d$ ，通过回归方程计算单期因子收益率  $f_k$ ；

$$R_{t+d} = f_{industry} X_{industry} + f_{style} X_{style} + f_k \varepsilon_K + \varepsilon_{t+d}$$

**Step3:** 计算因子收益率序列的年化收益  $E(f_k)$  及信息比率  $IR(f_k)$ ；

$$E(f_k) = 252 * (f_k / d) \quad IR(f_k) = \sqrt{252 * (f_k / d) / \sigma(f_k / d)}$$

**Step4:** 对于不同的收益预测周期  $d'$ ，重复第 2、3 步。

$$R_{t+d'} = f_{industry} X_{industry} + f_{style} X_{style} + f'_k \varepsilon_K + \varepsilon_{t+d'}$$

$$E(f'_k) = 252 * (f'_k / d') \quad IR(f'_k) = \sqrt{252} * (f'_k / d') / \sigma(f'_k / d')$$

我们统计了全部因子分别在预测周期为 1 天、2 天、3 天、4 天、5 天的情况下，年化因子收益率及信息比率。

其中  $FR\_DAY1$  表示预测未来 1 天收益率的所有因子收益率分布，预测周期以此类推； $IR\_DAY1$  表示预测未来 1 天收益率的所以因子收率信息比率分布，预测周期同样以此类推，具体分布及统计情况如下：

图 9  $FR\_DAY1$

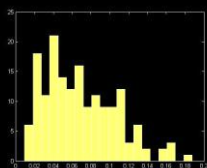


图 10  $FR\_DAY2$

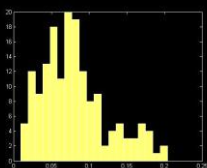


图 11  $FR\_DAY3$

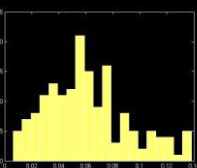


图 12  $FR\_DAY4$

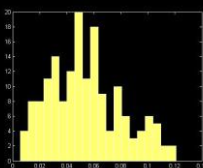
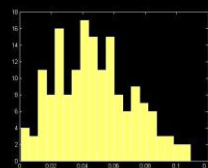


图 13  $FR\_DAY5$



数据来源：国泰君安证券研究

表 7：不同预测周期下所有因子收益率均值

	DAY1	DAY2	DAY3	DAY4	DAY5
年化因子 收益率均值	6.85%	8.08%	6.09%	5.42%	4.78%

数据来源：国泰君安证券研究

图 14  $IR\_DAY1$

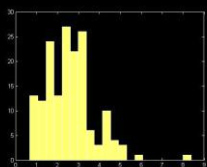


图 15  $IR\_DAY2$

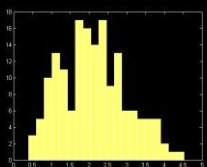


图 16  $IR\_DAY3$

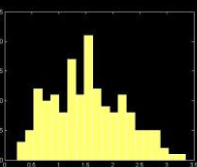


图 17  $IR\_DAY4$

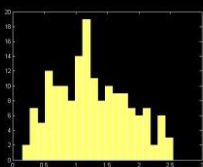
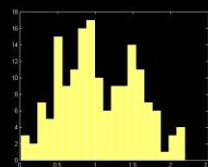


图 18  $IR\_DAY5$



数据来源：国泰君安证券研究

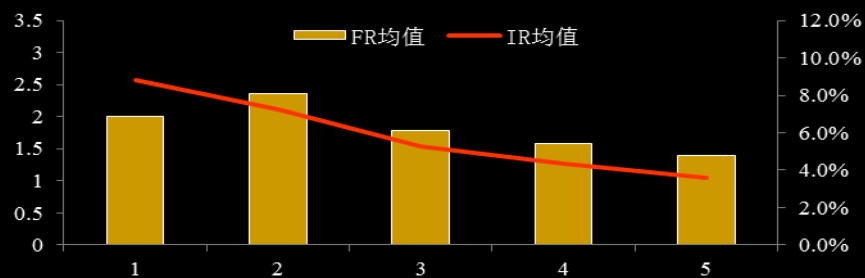
表 8：不同预测周期下所有因子收益率 IR 均值

	DAY1	DAY2	DAY3	DAY4	DAY5
因子收益率 IR 均值	2.57	2.11	1.73	1.26	1.05

数据来源：国泰君安证券研究



图 19 不同周期下因子收益率均值及 IR 均值变化情况



数据来源：国泰君安证券研究

从上述因子收益率的分布情况而言，大部分因子的年化因子收益率在 2%-10% 左右，极少部分因子收益率超过 15%。从平均收益情况而言，因子体系对未来 2 天的收益预测平均收益最高，T+3 日之后收益情况递减。

而从上述因子收益率的显著性水平情况而言，平均来说因子体系对越短周期的预测显著性越强，呈现明显的单调变化。而在 T+4 日之后，整体因子收益率 IR 均值未达到 95% 统计显著性水平。换言之，因子体系的预测周期极限即为 T+4 日。

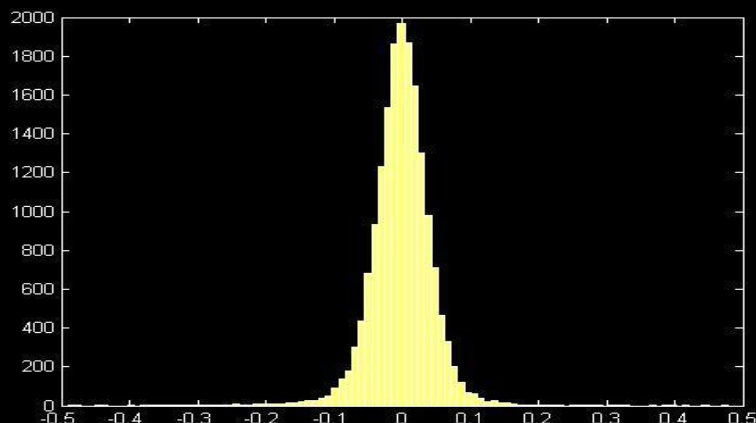
对于因子体系的相关性问题，由于短周期因子体系的因子数量较多，我们则不再对任意两两因子做正交处理，而选择在构建完毕所有因子后统一处理，即在我们得到所有因子收益率  $f$  之后，我们分别计算任意两个因子  $k$  和  $k'$  的因子收益率相关系数，即

$$IC_{k:k'} = corr(f_k, f_{k'})$$

若  $IC_{k:k'}$  大于某一给定阈值  $\xi$ ，则剔除因子收益率 IR 较低的因子，保留

IR 较高的因子。上一节中给出的因子体系，则是我们完成上述计算筛选后的结果。我们对所有因子收益率的相关系数分布情况及均值进行了计算，具体结果如下：

图 20 因子收益率相关系数分布



数据来源：国泰君安证券研究

经计算，因子收益率相关系数分布均值仅为-0.0057，并且94%的因子收益率相关系数绝对值小于0.1。

上述因子收益率显著性、相关性、及预测周期的分析表明，在短周期内（T<4 天）因子体系呈现了整体高显著、低相关的特征，这样的特点十分有助于因子模型在收益预测和实现组合收益的过程中，获得较高的较稳定的超额收益。

### 3.4.2. 因子体系的预测能力

我们进一步考察因子体系对收益的预测能力，其中统计方法与我们在第二章中提到的阿尔法模型的评价标准相一致。

我们首先利用一定的转化方式，将全部因子转化为对个股阿尔法收益截面的预测值，并进而统计预测值与事后计算可得的实际值进行相关系数计算及其显著性检验。我们考察当预测周期  $t=1$  的情况下，模型预测系数（*IC of Alpha Model*）的计算结果，具体方法如下：

**Step1:** 对所有因子进行正交标准化处理，取目标因子残差截面  $\varepsilon_k^t$ ，即

$$X_k^t = \beta_{industry} X_{industry} + \beta_{style} X_{style} + \varepsilon_k^t$$

**Step2:** 针对给定预测周期  $t=1$ ，通过多次多元回归方程计算得到过去时间段 T 内过所有因子收益率向量簇  $\{f_k^t\}_K^T$ ，即

$$R = f_{industry} X_{industry} + f_{style} X_{style} + f_{k1} \varepsilon_{k1} + f_{k2} \varepsilon_{k2} + \dots + f_K \varepsilon_{K2} + \varepsilon$$

**Step3:** 利用过去 T 时间段因子收益率均值作为下期预测值，乘以当期每个目标因子残差截面，得到阿尔法收益截面预测向量，即

$$E(\varepsilon_{t+1}) = \sum_{k=1}^K ((\sum_{t=1}^T f_k^t / T) * \varepsilon_k^t)$$

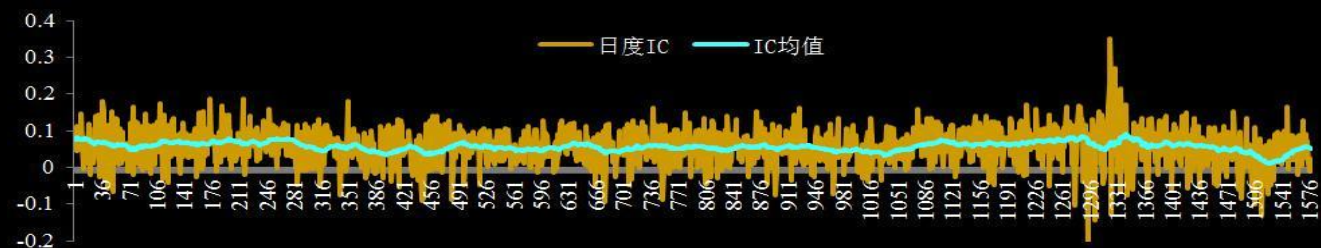
**Step4:** 计算各期模型预测系数，进而计算模型系数 T 检验结果，即

$$IC_{AlphaModel}^t = corr(E(\varepsilon_{t+1}), \varepsilon_{t+1})$$

$$TSta_{AlphaModel} = \overline{IC_{AlphaModel}} / (\sigma_{IC_{AlphaModel}} / \sqrt{T})$$

按照上述步骤，我们对因子体系的预测显著性进行了统计计算，其中时间自 2010 年 1 月至 2017 年 4 月，预测周期为 T+1 日频率，股票截面则为全 A 非 St，我们分别统计了每日模型 IC 及 30 日 IC 均值，具体计算结果如下：

图 21 模型预测系数



数据来源：国泰君安证券研究

表 9: 模型预测系数特征统计

日度 IC 均值	<i>TStas Alpha Model</i>	日度 IC 胜率
0.057	18.23	88.7%

数据来源：国泰君安证券研究

结果表明，因子体系在日度预测阿尔法收益截面的情况下，IC 均值为 0.057，模型预测系数显著性高达 18.23，整体预测胜率接近 90%，仅在 2015 年股灾期间有所波动，其余时间段稳定性较强。

因此从统计结果来看，我们构建的交易型阿尔法因子体系对短周期内阿尔法超额收益具有很强的预测能力，这既符合我们对交易型阿尔法收益特征的判断，也是随后构建策略获取较高超额收益的基础。

### 3.4.3. 因子个数与模型预测能力的关系

本节讨论因子数量的问题，可能有投资者会疑问，是否真的需要运用如此多将近 200 个阿尔法因子。我们将从因子数量与模型预测能力的角度出发，来考察这个问题。

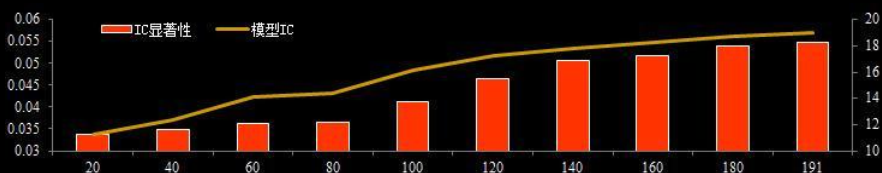
我们首先根据所有单因子检验的 *Factor Return IR* 排序，然后分别选取 20 个、40 个、60 个因子对阿尔法收益进行预测并计算模型预测系数及其显著性，以此类推直到用尽全部因子。其中检验时间仍为 2010 年 1 月至 2017 年 4 月，股票截面则为全 A 非 St，具体结果如下所示：

表 10: 因子个数与模型预测能力比较

因子个数	20	40	60	80	100	120	140	160	180	191
模型 IC	0.0338	0.037	0.0424	0.0432	0.0483	0.0518	0.0534	0.0549	0.0562	0.057
IC 显著性	11.26	11.59	12.06	12.13	13.75	15.54	16.86	17.26	17.98	18.23

数据来源：国泰君安证券研究

图 22 因子个数与模型预测能力比较



数据来源：国泰君安证券研究

从上述分析结果中可以看到，随着有效因子数量的增加，因子体系的预测能力也呈现单调的上升趋势。由于，我们采用因子收益率 IR 排序，因此曲线呈现了边际递减的形态，但这并不妨碍绝对预测能力的提升。

所以说，目前的因子数量还远未达到可预测收益的边界区域，因子体系的空间还是比较大的。其实，在一些海外知名的对冲基金中，也有类似的交易型阿尔法策略体系的布局，其所使用的数据范围、因子数量和算法种类远超想象，甚至存在因子数量在百万、上亿级别的情况，可以说是一种崭新的交易模式。当然，一味追求因子个数的增加是不合理的，关键还是在于对任意单个因子，必须达到必要的统计显著结果才可纳入因子体系。

### 3.5. 组合换手率与超额收益

较高的组合换手率会使得投资者对策略实战效果有所担忧。对于大部分投资者而言，较高的策略换手率未必会带来更高的收益率，交易越频繁亏损越大也成为了较为普遍的现象。但是，对于量化模型获得较为稳定的阿尔法来源之后，提高投资换手率往往会带来收益率的显著提升。在众多的学术文献中，我们往往会看到一些选股策略或 CTA 策略在不考虑交易成本的情况下，实现了惊人的投资胜率和收益率。但是，在实际投资环境中，由于较高的换手率所导致的交易成本将大幅侵蚀策略的收益率，甚至导致最后的亏损。

因此，高换手率、低交易成本与高收益率可以说构成了投资的不可能三角，在策略设计的时候，必须寻找到三者之间的最优平衡状态。对于一些非常有价值的投资信号或阿尔法源，我们往往由于其所负担的每期 100% 的换手率而被迫将其放弃。但是，如果我们通过换手率与交易成本之间的平衡，将信号换手率控制在每月 60%、40% 甚至 20%，那么该策略所获取的收益仍然将是非常可观的。

具体而言，控制组合换手率的方式有很多，大体上分为因子层面、优化目标层面和风险约束层面。

1. 因子层面的换手率平衡，是指在因子构建过程中，就考虑组合过程中的高换手问题，因此做加长周期、平滑等处理，但是这种方式对短期阿尔法损失影响较大，适合于中低频的传统多因子模型，不适合短周期交易型阿尔法策略。
2. 优化目标层面的换手率平衡，是指在优化构建组合的过程中，以单次交易换手成本作为罚函数加入目标收益中，以实现考虑交易成本后的预期收益率最大化，具体的目标函数形式为：

$$\text{Maximize } \sum_{i=1}^N w_i E(r_i) - Tc \sum_{i=1}^N |w_i|$$

其中  $w_i$  为所需要求解的目标组合权重向量， $Tc$  为设定的客观交易成本。

这种目标函数的构建很大程度的降低了策略对交易成本的敏感程度，尤其是对短周期交易型阿尔法策略。一般情况下，如果只以最大化预期收益率为目标函数，很有可能出现任意单次换仓比率均接近 100% 的情况，虽然这样有利于阿尔法的收益空间，但是牺牲的交易成本也过大。因此，上述目标函数的构造实现了阿尔法与交易成本之间的平衡，是组合最优的拐点位置。

当然，这样的目标函数构建对超额预测  $E(\varepsilon)$  的预测精度要求较高，如果阿尔法模型的预测显著性不强，则也不适用这样的方法。

3. 风险约束层面的换手率平衡，是最为直观、简便的换手率控制方式，即在优化构建组合的过程中，在约束条件中严格控制单次换仓比例，即：

$$s.t. \quad \sum |w_t - w_{t-1}| / 2 \leq TurnOver$$

其中， $TurnOver$  为设定的单次换手率上限。

这种方法的优点在于简洁直观、效果明显，并且可以保证换手率的控制效果精确可控。但是其缺陷也非常明显，换手比例限制过于死板，并未综合考量超额收益与换手率的关系，在某些情况下，可能会以牺牲较多的阿尔法空间为代价。并且，在规划求解的过程中，容易出现由于和其他约束条件相冲突导致方程无解的情况发生。

在我们的策略设计中，采用了第 2 种优化目标层面的换手率平衡方式，因此随着客观交易成本  $Tc$  的提升，策略的整体换手率将逐步降低，超额收益也将有所降低。我们将在下一节中，对策略的交易成本敏感性做实证分析，考察模型可容忍的极限交易成本。

### 3.6. 基于短周期价量特征的风格中性多因子选股策略

至此，我们已经完成了因子体系构建、阿尔法收益预测、因子预测周期评估、换手率与交易成本的平衡分析。本节，我们将构建基于短周期价量特征的风格中性多因子选股策略。

#### 3.6.1. 策略设计

在整个策略的设计过程中，我们保持着 2 个原则：首先，尽可能的提高策略收益率。由于  $Alpha$  正比于  $IC_{AlphaModel} \cdot \sqrt{N} \cdot \sigma_{Target} \cdot dis(mkt)$ ，因此在此在因子体系贡献了较高较稳定的  $IC_{AlphaModel}$  的基础上，我们尽可能的提高尝试次数  $\sqrt{N}$ 。而在主动风控端  $\sigma_{Target}$ ，我们并未强行设定目标跟踪误差，因此模型在设计上，整体收益的空间较大。



其次，收益端与风险端的分离。我们以交易型阿尔法为超额收益的来源设计组合优化的目标函数，而在风控约束端，保持组合在所有大类风格因子层面的中性处理。这种收益与风险完全分离的做法，使得模型在获得较高较稳定收益的同时，尽可能的降低了回撤风险。

策略具体构建步骤如下所示：

**Step1:** 对所有因子进行正交标准化处理，取目标因子残差截面  $\varepsilon_K^t$ ，即

$$X_K^t = \beta_{industry} X_{industry} + \beta_{style} X_{style} + \varepsilon_K^t$$

**Step2:** 针对给定预测周期  $t=2$ ，通过多次多元回归方程计算得到过去时间段  $T$  内所有因子收益率向量簇  $\{f_k^t\}_K^T$ ，即

$$R = f_{industry} X_{industry} + f_{style} X_{style} + f_{k1} \varepsilon_{k1} + f_{k2} \varepsilon_{k2} + \dots + f_K \varepsilon_{K2} + \varepsilon$$

**Step3:** 利用过去  $T=250$  日因子收益率均值作为下期预测值，乘以当期每个目标因子残差截面，得到阿尔法收益截面预测向量，即

$$E(\varepsilon_{t+1}) = \sum_{k=1}^K ((\sum_{t=1}^T f_k^t / T) * \varepsilon_k^t)$$

**Step4:** 以经交易成本调整后预测阿尔法收益截面为目标函数，以大类行业、风格中性控制为约束条件，求解组合权重向量  $w_t$ ，即

$$\begin{aligned} \text{Max } & w_t' \cdot E(\varepsilon) - Tc \cdot \sum |w_t - w_{t-1}| / 2 \\ \text{s.t. } & (w_t' - w_{bench\_mark}') X_{style} = 0 \\ & w_t' \cdot X_{industry} = w_{bench\_mark\_industry}' \\ & w \geq 0 \\ & \sum w = 1 \end{aligned}$$

其中， $w_{bench\_mark}$  和  $w_{bench\_mark\_industry}$  均为比较基准的个股权重和行业权重向量。

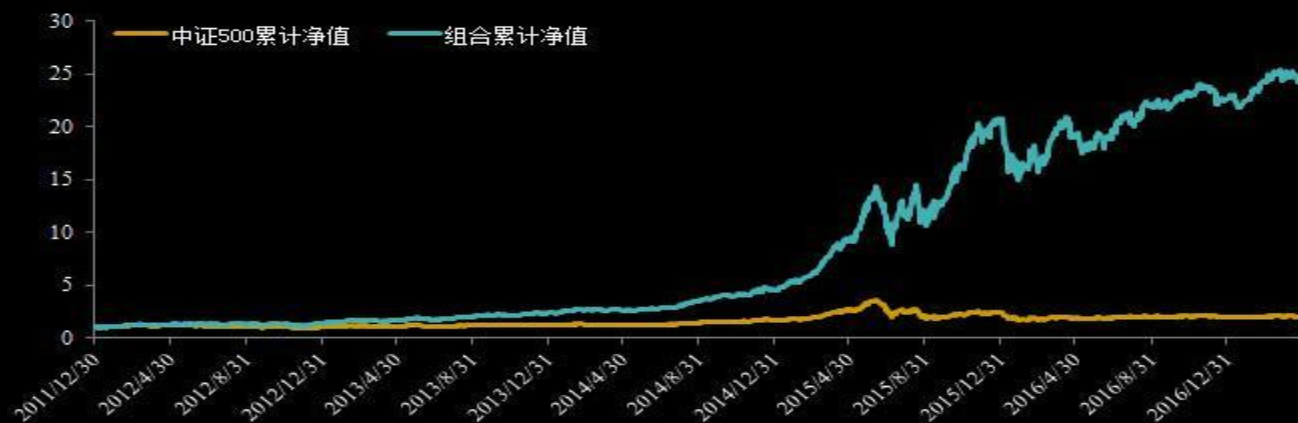
### 3.6.2. 实证分析

本节，我们将对策略进行实证回溯分析，其中相关参数设定如下：

- 1) 回测时间从 2012 年 1 月至 2017 年 4 月；
- 2) 比较基准为中证 500 指数；
- 3) 股票池选取全 A 非 ST 股票；
- 4) 交易成本为单边 0.1%，印花税 0.1%，即  $Tc=0.3\%$ ；
- 5) 行业中性约束中，因子敞口设定为  $\pm 5\%$ （相对比例）；
- 6) 风格中性约束中，因子敞口设定为  $\pm 0.1$ ；
- 7) 组合以次日开盘价模拟成交，并剔除停牌、涨跌停等异常情况；

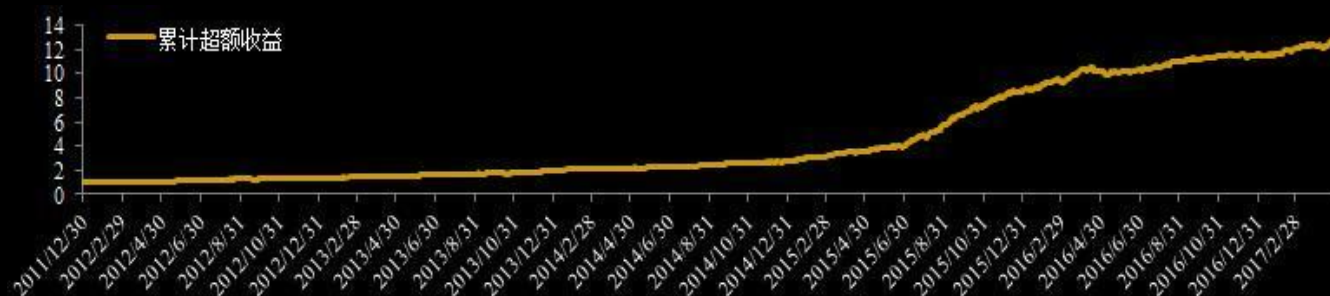
策略具体结果及相关绩效如下所示：

图 23 组合累计净值



数据来源：国泰君安证券研究

图 24 组合累计超额收益



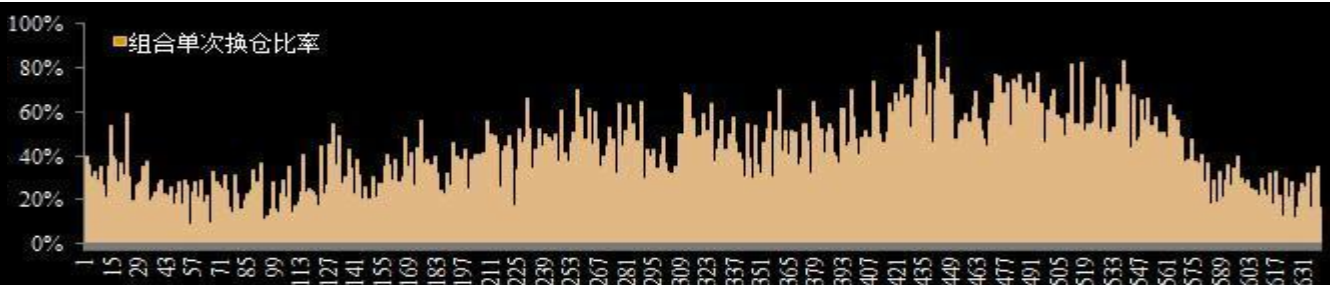
数据来源：国泰君安证券研究

表 11: 组合绩效统计

年化超额收益	50.20%	年化换手率（双边）	48.2 倍
年化跟踪误差	10.73%	平均单次换仓比率（双边）	38%
信息比率	4.67	平均每期持仓股票数	52 只
最大回撤	5.90%	最大回撤时间段	2014 年 3 月-4 月

数据来源：国泰君安证券研究

图 25 组合单次换仓比率



数据来源：国泰君安证券研究

表 12: 组合分年度绩效统计

年份	年化超额收益	信息比率	最大回撤
2012 年	36.86%	4.18	2.75%
2013 年	43.96%	3.99	4.54%
2014 年	40.00%	3.59	5.90%
2015 年	117.70%	7.81	4.29%
2016 年	35.92%	3.18	5.87%
2017 年（4 月末年化）	35.26%	4.92	2.95%

数据来源：国泰君安证券研究

上述结果表明，策略自 2012 年 1 月至 2017 年 4 月，相较中证 500 指数，实现年化超额收益 50.2%，信息比率 4.67，最大回撤 5.90%，发生在 2014 年 3 月至 4 月之间，组合整体收益表现优异。分年度表现来看，2015 年市场波动率大幅增加给予了交易型阿尔法策略更大的空间，当年超额收益超过 100%，其余年份超额收益表现均在 30%-40% 左右。值得注意的是，在市场风格发生较大变化的 2017 年，截止 4 月末，策略仍然实现年化 35% 左右的超额收益，表现出策略收益来源与组合构造均与市场风格较为独立的特性。

从换手率的角度来看，策略年化双边换手率为 48.2 倍，单次平均换仓比率约为 40%，而从分年化的换手情况来看，2015 年牛市阶段的换手率相对较高，这也是导致该年度策略收益表现较好的直接原因。

整体而言，从我们的测算结果来看，短周期交易型阿尔法策略体系在收益空间方面较之传统价值型阿尔法策略有巨大的提升，并且在风格中性的配置下组合的稳定性仍然得到了较好的保证。

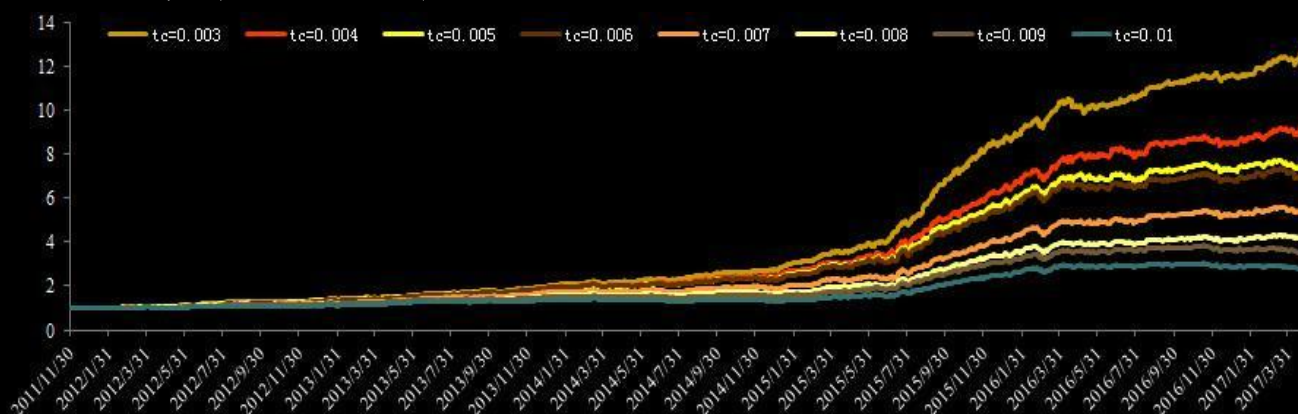
### 3.6.3. 交易成本敏感性分析

本节，我们将考察策略对交易成本的敏感性问题，以检验策略可承受的交易成本（包括冲击成本）的理论极限空间。

在组合换手率与超额收益一节中，我们提出加入单次换仓成本罚函数作为目标函数的做法，即  $Max \quad w'_t \cdot E(\varepsilon) - Tc \cdot \sum |w_t - w_{t-1}| / 2$ ，其中  $Tc$  即为设定的客观交易成本。

从目标函数的构造来看，为了实现经交易成本调整后的超额收益最大化，策略随着  $Tc$  的不断提升，组合各期单次换仓比率将逐步降低，而阿尔法也将有所降低。接下来，我们将对  $Tc$  进行敏感性分析，设定  $Tc$  自 0.3% 不断提升至 1%，观测策略整体的收益、风险及换手率等相关情况，具体统计结果如下图所示：

图 26 不同交易成本设定下组合累计超额收益



数据来源：国泰君安证券研究

表 13: 不同交易成本设定下组合绩效统计

	年化超额收益	年化换手率(双边)	平均单次换仓比率	最大回撤	信息比率
TC=0.3%	50.20%	48.2 倍	38.08%	5.90%	4.67
TC=0.4%	43.96%	41.7 倍	33.11%	6.53%	3.62
TC=0.5%	40.08%	38.8 倍	30.84%	6.68%	3.28
TC=0.6%	38.77%	36.1 倍	28.73%	7.86%	3.12
TC=0.7%	33.54%	33.2 倍	26.42%	7.50%	2.80
TC=0.8%	28.90%	30.2 倍	24.03%	7.69%	2.42
TC=0.9%	25.40%	28.2 倍	22.42%	8.99%	2.15
TC=1%	20.70%	26.1 倍	20.75%	10.50%	1.75

数据来源：国泰君安证券研究

从结果来看，随着设定的客观交易成本不断增加，策略的年化换手率也随之降低。尽管策略的超额收益也逐渐下降，但是整体而言，由于做了相对平衡的目标函数控制，因此策略对交易成本的敏感程度并不会导致灾难性的结果发生。

可以看到，当  $Tc=0.8\%$  时，组合年化超额收益率仍然接近 30%，信息比率接近 2.5，收益风险表现依然较为平稳。而当  $Tc=1\%$  的情况下，策略的整体收益表现接近极限位置，若交易成本再提升，模型将会趋于失效。

当然，对交易成本  $Tc$  的估计有许多种方式，不同域的股票所产生的冲击成本也不尽相同，因此在实际操作中可以根据不同股票的模拟冲击成本将  $Tc$  做向量化处理，使得整体结果更为精确。

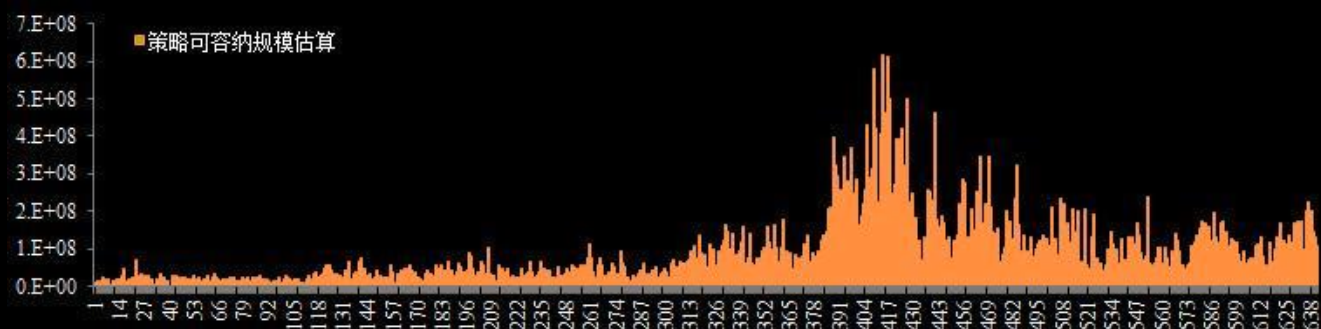
### 3.6.4. 策略可容纳规模估算

收益、风险和资金容纳规模构成投资的不可能三角，在策略实现相对较高收益、较低风险的情况下，可容纳的资金规模相对不会很大，因此有必要对这类策略的规模容量做一定的估算。

我们以持仓组合中，各只股票当日成交金额的 10% 作为个股可容纳资金规模上限。同时，为了避免极端小成交金额对组合整体估计的偏差，我们以作用组合权重后所有个股可容纳资金规模上限倒序排列的 95% 分位

数，作为当日策略可容纳规模上限估计值，测算时间与组合构造时间保持一致，具体结果如下所示：

图 27 策略可容纳规模估算



数据来源：国泰君安证券研究

从上述结果可以看到，策略可容纳资金规模与市场整体成交量相关，2012 年至 2014 年，由于市场整体成交萎缩，导致策略可容纳规模在 1 亿左右。而 2014 年至 2016 年牛市阶段，市场成交活跃，因此策略可容纳规模大幅提升，最高接近 7 亿水平。而在 2017 年当前市场成交水平，策略可容纳规模在 2 到 3 亿左右。

因此，客观而言，对于资金量较大的机构投资者而言，完全利用该种交易型策略进行阿尔法投资不太现实，策略比较适合运用于多策略体系中的一支，或大部分资金可以布局于容量较大但收益偏低的传统价值型阿尔法策略中，而小部分的资金量布局于该类高换手且收益较高的交易型阿尔法策略中，提高策略组合的整体收益水平。

当然，从策略的设计上而言，同样可以通过降低一定的收益率来获取更大的资金容纳规模，期间的因子运用、组合构造等步骤是相对灵活、开放的。

## 4. 总结与展望

本篇报告，我们开创性的构建了全新的多因子模型体系——短周期交易型多因子阿尔法选股体系。其中，我们将超额收益的来源锁定在由于交易行为所产生的机会中，试图获取与传统价值型多因子模型相独立的阿尔法来源。

由于 A 股市场交易行为随机性很强，并且机构投资者的交易效率亦不十分高效，加之 T+0 交易、融券卖空交易模型的缺乏，导致在短周期内由交易产生的定价非有效情况较为常见。通过研究我们发现，与传统价值型阿尔法策略相比，交易型阿尔法的收益空间更高，稳定性也更强。

在本篇报告所介绍的交易型阿尔法体系中，我们总共构建、运用了将近 200 个阿尔法因子。研究分析表明，因子体系呈现出了高显著、低相关的特征，对超额收益具有很强的预测能力，这为后面策略构建提供了坚实的保障。



在策略设计方面，我们通过超额收益的因子化预测转换、换手率与交易成本的平衡、收益端与风险端的分离等方式，实现了策略整体的设计，回测结果表明，在可接受的交易成本和资金容量的情况下，模型可实现较高且稳定的超额收益。

尽管如此，在整个研究过程中，我们感觉到只是看到了交易型阿尔法空间的冰山一角。在海外成熟市场，类似的统计套利型的阿尔法策略占到了整个量化投资的半壁江山，因此该类策略或许也是量化投资的重要发展方向之一。

并且，随着近些年来，人工智能、机器学习算法的飞速发展，未来交易型阿尔法模型体系将大量运用该类算法去进行因子挖掘，但是这并不意味着机器将替代人类，对于投资者而言，底层的投资逻辑和策略设计仍然掌握在其逻辑框架中，这是任何技术都无法改变的。

在后期的研究中，数据源的拓展也是一个核心的方向，当前策略体系我们仅仅是利用了日频率的价格信息，而日内的高频数据则包含了更为显著的预测信息，是值得深入研究的方向。此外，场外的数据信息同样包含大量的预测信息，通过整体的策略设计，模型可提升的空间仍然较大。

短周期交易型阿尔法策略体系的构建，希望对投资者的研究思路有所拓展，在量化模型最擅长的领域发挥其真正的威力。

## 5. 附录

### 附录 1 大类风格因子定义明细

表 14: 大类风格因子定义明细（小类因子加权方式详见 BARRA CNE5 相关报告）

大类因子	小类因子	因子计算方式
Beta	BETA	$r_i = \alpha + \beta r_m + e_i$ ；利用个股收益率序列和沪深 300 指数收益率序列进行一元线性回归，速率序列长度取 250 交易日。股收益率序列和沪深 300 指数收益率序列均以半衰指数加权，半衰期为 60 日。
Momentum	RSTR	$RSTR = \sum_{t=L}^{T+L} w_t [\ln(1 + r_t)]$ ；其中 T=500, L=21，收益率序列以半衰指数加权，半衰期为 120 日。
Size	LNCAP	$LNCAP = LN(total\_market\_capitalization)$ ；个股总市值对数值。
	EPIBS	$EPIBS = est\_eps / P$ ；其中 est_eps 为个股一致预期基本每股收益。
Earnings Yield	ETOP	$ETOP = earnings\_ttm / mkt\_freeshares$ ；历史 EP 值，利用过去 12 个月个股净利润除以当前市值。
	CETOP	$CETOP = Cash\_earnings / P$ ；个股现金收益比股票价格。

	DASTD	$DASTD = \left( \sum_{t=1}^T w_t \cdot (r_t - \mu(r))^2 \right)^{1/2}$ ; 其中收益率序列长度取 250 个交易日, 半衰期设定为 40 日。
Volatility	CMRA	$CMRA = \ln(1 + \max\{Z(T)\}) - \ln(1 + \min\{Z(T)\})$ ; 其中 $Z(T) = \sum_{\tau=1}^T [\ln(1 + r_{\tau})]$ ; $r_{\tau}$ 表示个股月收益率, T 代表过去 12 个月。
	HSIGMA	$HSIGMA = std(e_i)$ ; 其中残差 $e_i$ 为 BETA 计算中所得。
Growth	SGRO	过去 5 年企业营业总收入复合增长率。
	EGRO	过去 5 年企业归属母公司净利润复合增长率。
	EGIB	未来 3 年企业一致预期净利润增长率。
	EGIB_S	未来 1 年企业一致预期净利润增长率。
Value	BTOP	$BTOP = common\_equity / current\_market\_capitalization$ ; 计算企业总权益值除以当前市值。
	MLEV	$MLEV = (ME + LD) / ME$ ; 其中 ME 表示企业当前总市值, LD 表示企业长期负债。
Leverage	DTOA	$DTOA = TD / TA$ ; 其中 TD 表示总负债 TA 表示总资产。
	BLEV	$BLEV = (BE + LD) / BE$ ; 其中 BE 表示企业账面权益, LD 表示企业长期负债。
Liquidity	STOM	$STOM = \ln(\sum_{i=1}^{21} (V_i / S_i))$ ; 其中 $V_i$ 表示当日成交量, $S_i$ 表示流通股本。
	STOQ	$STOQ = \ln(\frac{1}{T} \sum_{\tau=1}^T \exp(STOM_{\tau}))$ ; 其中 T=3。
	STOA	$STOA = \ln(\frac{1}{T} \sum_{\tau=1}^T \exp(STOM_{\tau}))$ ; 其中 T=12。

数据来源: 国泰君安证券研究

## 5.1. 附录 2 表 6 相关变量及函数说明

表 15: 表 6 相关变量及函数说明

变量及函数	定义
OPEN	开盘价
HIGH	最高价
LOW	最低价
CLOSE	收盘价
VWAP	均价
VOLUME	成交量
AMOUNT	成交额
BANCHMARKINDEXCLOSE	基准指数的开盘价
BANCHMARKINDEXOPEN	基准指数的收盘价
RET	每日收益率 (收盘/前收盘-1)
DTM	$(OPEN <= DELAY(OPEN, 1) ? 0 : MAX((HIGH - OPEN), (OPEN - DELAY(OPEN, 1))))$
DBM	$(OPEN >= DELAY(OPEN, 1) ? 0 : MAX((OPEN - LOW), (OPEN - DELAY(OPEN, 1))))$

TR	$\text{MAX}(\text{MAX}(\text{HIGH}-\text{LOW}, \text{ABS}(\text{HIGH}-\text{DELAY}(\text{CLOSE}, 1))), \text{ABS}(\text{LOW}-\text{DELAY}(\text{CLOSE}, 1)))$
HD	$\text{HIGH}-\text{DELAY}(\text{HIGH}, 1)$
LD	$\text{DELAY}(\text{LOW}, 1)-\text{LOW}$
HML SMB MKE	Fama French 三因子
SELF	特殊变量，出现在 Alpha143，表示 t-1 日的 Alpha143 因子计算结果
RANK(A)	向量 A 升序排序
MAX(A, B)	在 A, B 中选择最大的数
MIN(A, B)	在 A, B 中选择最小的数
STD(A, n)	序列 A 过去 n 天标准差
CORR(A, B, n)	序列 A、B 过去 n 天相关系数
DELTA(A, n)	$A_i - A_{i-n}$
LOG(A)	自然对数函数
SUM(A, n)	序列 A 过去 n 天求和
ABS(A)	绝对值函数
MEAN(A, n)	序列 A 过去 n 天均值
TSRANK(A, n)	序列 A 的末位值在过去 n 天的顺序排位
SIGN(A)	符号函数 $\begin{cases} 1 & \text{if } A > 0 \\ 0 & \text{if } A = 0 \\ -1 & \text{if } A < 0 \end{cases}$
COVANCE(A, B, n)	序列 A、B 过去 n 天协方差
DELAY(A, n)	$A_{i-n}$
TSMIN(A, n)	序列 A 过去 n 天的最小值
TSMAX(A, n)	序列 A 过去 n 天的最大值
PROD(A, n)	序列 A 过去 n 天累乘
COUNT(condition, n)	计算前 n 期满足条件 condition 的样本个数
REGBETA(A, B, n)	前 n 期样本 A 对 B 做回归所得回归系数
REGRESI(A, B, n)	前 n 期样本 A 对 B 做回归所得的残差
SMA(A, n, m)	$\hat{Y}_{i+1} = (A_i m + \hat{Y}_i (n - m)) / n$ ，其中 $\hat{Y}$ 表示最终结果
SUMIF(A, n, condition)	对 A 前 n 项条件求和，其中 condition 表示选择条件
WMA(A, n)	计算 A 前 n 期样本加权平均值权重为 0.9i, (i 表示样本距离当前时点的间隔)
DECAYLINEAR(A, d)	对 A 序列计算移动平均加权，其中权重对应 d, d-1, ..., 1 (权重和为 1)
FILTER(A, condition)	对 A 筛选出符合选择条件 condition 的样本
HIGHDAY(A, n)	计算 A 前 n 期时间序列中最大值距离当前时点的间隔
LOWDAY(A, n)	计算 A 前 n 期时间序列中最大值距离当前时点的间隔
SEQUENCE(n)	生成 1~n 的等差序列
SUMAC(A, n)	计算 A 的前 n 项的累加
&	逻辑运算与
	逻辑运算或
A? B:C	若 A 成立，则为 B，否则为 C

数据来源：国泰君安证券研究

## 本公司具有中国证监会核准的证券投资咨询业务资格

### 分析师声明

作者具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，保证报告所采用的数据均来自合规渠道，分析逻辑基于作者的职业理解，本报告清晰准确地反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响，特此声明。

### 免责声明

本报告仅供国泰君安证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为本公司的当然客户。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。

本报告的信息来源于已公开的资料，本公司对该等信息的准确性、完整性或可靠性不作任何保证。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。

本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。在任何情况下，本公司、本公司员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。

本公司利用信息隔离墙控制内部一个或多个领域、部门或关联机构之间的信息流动。因此，投资者应注意，在法律许可的情况下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。在法律许可的情况下，本公司的员工可能担任本报告所提到的公司的董事。

市场有风险，投资需谨慎。投资者不应将本报告作为作出投资决策的唯一参考因素，亦不应认为本报告可以取代自己的判断。在决定投资前，如有需要，投资者务必向专业人士咨询并谨慎决策。

本报告版权仅为本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制、发表或引用。如征得本公司同意进行引用、刊发的，需在允许范围内使用，并注明出处为“国泰君安证券研究”，且不得对本报告进行任何有悖原意的引用、删节和修改。

若本公司以外的其他机构（以下简称“该机构”）发送本报告，则由该机构独自为此发送行为负责。通过此途径获得本报告的投资者应自行联系该机构以要求获悉更详细信息或进而交易本报告中提及的证券。本报告不构成本公司向该机构之客户提供的投资建议，本公司、本公司员工或者关联机构亦不为该机构之客户因使用本报告或报告所载内容引起的任何损失承担任何责任。

### 评级说明

#### 1. 投资建议的比较标准

投资评级分为股票评级和行业评级。以报告发布后的 12 个月内的市场表现为比较标准，报告发布日后的 12 个月内的公司股价（或行业指数）的涨跌幅相对同期的沪深 300 指数涨跌幅为基准。

#### 2. 投资建议的评级标准

报告发布日后的 12 个月内的公司股价（或行业指数）的涨跌幅相对同期的沪深 300 指数的涨跌幅。

	评级	说明
股票投资评级	增持	相对沪深 300 指数涨幅 15%以上
	谨慎增持	相对沪深 300 指数涨幅介于 5%~15%之间
	中性	相对沪深 300 指数涨幅介于-5%~5%
	减持	相对沪深 300 指数下跌 5%以上
行业投资评级	增持	明显强于沪深 300 指数
	中性	基本与沪深 300 指数持平
	减持	明显弱于沪深 300 指数

## 国泰君安证券研究所

	上海	深圳	北京
地址	上海市浦东新区银城中路 168 号上海银行大厦 29 层	深圳市福田区益田路 6009 号新世界商务中心 34 层	北京市西城区金融大街 28 号盈泰中心 2 号楼 10 层
邮编	200120	518026	100140
电话	(021) 38676666	(0755) 23976888	(010) 59312799
E-mail:	gtjaresearch@gtjas.com		