



北京大学
PEKING UNIVERSITY
量化交易协会
Quantitative Trading Association

多因子选股I——策略基本逻辑

主讲人：邓珂雅 何金泽 沈廷威 胡逸凡



多因子选股I——策略基本逻辑

1 传统因子模型和他们之间的关系

2 因子类型及分布规律

3 因子评价指标

4 单因子有效性测试

资本资产定价模型CAPM

假设

- 通过预期收益与收益的方差来选择投资组合
- 投资者是风险厌恶类型，且永不满足
- 投资期是单期的
- 证券市场存在均衡状态
- 资产可以无限细分，投资规模无论是多少都是可以的
- 存在无风险资产，投资者可以按无风险利率借入借出无风险资产
- 没有税收和交易成本
- 投资者对每种证券的收益和风险有相同的预期

资本资产定价模型CAPM

模型

$$E(r_i) = r_f + \beta_i [E(r_M) - r_f]$$

期望收益率

无风险收益率

承受风险带来的补偿

其中， r_i 表示证券*i*的收益率， r_f 表示无风险收益率， r_M 表示市场基准收益率。

$$\beta_i = \frac{\text{Cov}(r_i, r_M)}{\sigma_M^2}$$

CAPM系统风险 β

系统风险 β

CAPM也可以表达成如下形式：

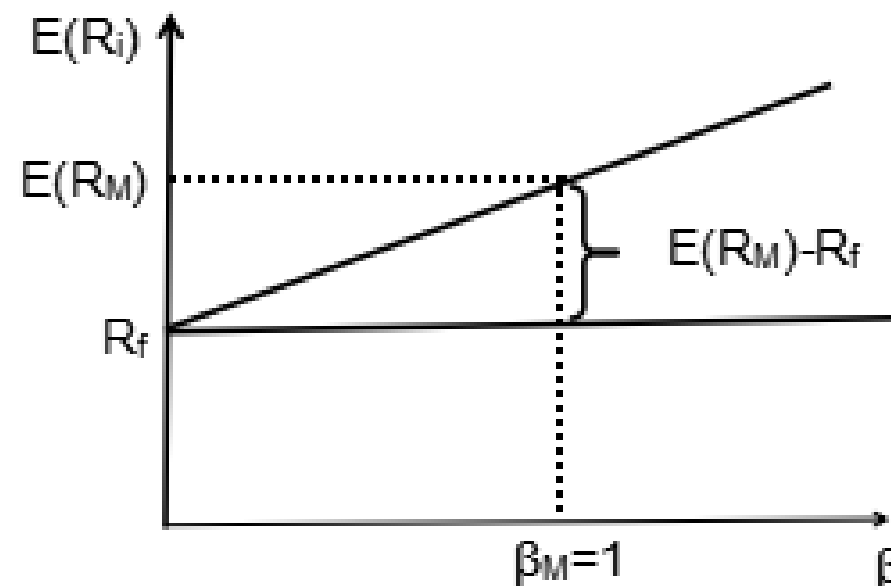
$$E(r_i) - r_f = \beta_i [E(r_M) - r_f]$$

$$\beta_i = \frac{\text{Cov}(r_i, r_M)}{\sigma_M^2}$$

模型当中，资产的超额收益率与 β_i 有关，其为证券相对于市场组合的风险度量， β_i 就是该证券的系统性风险。

- (1) $\beta = 0$ ：无风险资产与证券的协方差为0，任何无风险的资产 β 为0，任何 β 为0的资产的超额收益也是0；
- (2) $\beta = 1$ ：若某种证券与市场组合的协方差等于市场组合的方差，那么 β 为1，即该种风险资产可以获得市场组合的超额收益率；
- (3) β 值高时，获得的收益高， β 值低时，获得的收益低。

证券市场线(SML)



SML说明了，投资者面临的风险和投资者的收益成正比，即，世界上没有免费的午餐。

CAPM超额收益率 α

超额收益率

均衡市场中，所有证券都在证券市场线（SML）上，即

$$[E(r_i) - r_f] = \beta_i [E(r_M) - r_f]$$

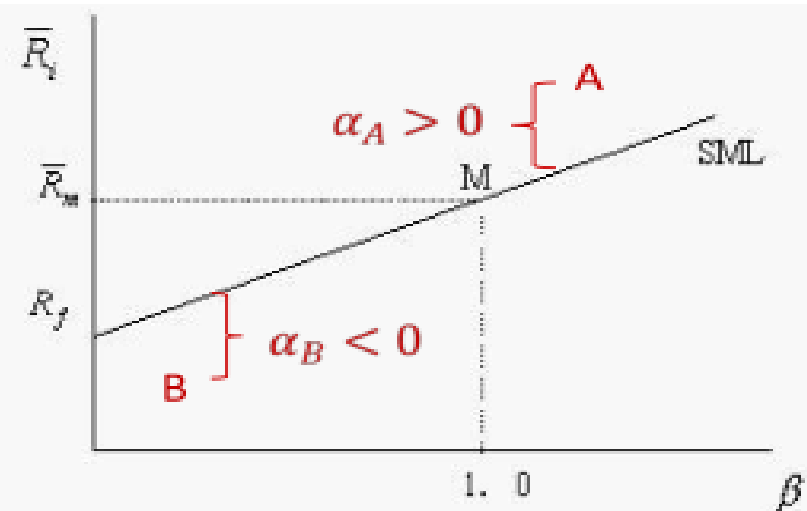
如果某只股票被高估或低估，就会偏离SML：

$$\alpha_i = [E(r_i) - r_f] - \beta_i [E(r_M) - r_f]$$

股票的 α 指的就是股票的实际期望收益与正常期望收益之间的差。

$\alpha > 0$ 时，证券组合实际收益位于 SML 之上，被低估推荐买入；

$\alpha < 0$ 时，证券组合实际收益位于 SML 之下，被高估推荐卖出；



Alpha策略：寻找未被正确定价的证券

当在市场中发掘存在正 α 的证券，我们就可以通过买入它的同时卖出等额的等 β 但 α 为0的证券获取超额收益，这一部分收益是没有对应风险的

Beta策略：改变组合beta

让市场上涨它涨得更多，市场下跌它跌得更少

套利定价理论APT

模型形式

套利定价模型把造成资产收益不确定性的两个源头分为公共因素（系统性）和特殊因素（可分散的非系统因素）。假定证券市场上，影响收益的公共因素有K个：

$$r_i = E(r_i) + \beta_1 F_1 + \beta_2 F_2 + \beta_3 F_3 + \cdots + \beta_k F_k + e_i$$

收益不确定性的公共因素解释部分

收益不确定性的特殊因素解释部分

F_i ：公共因素描述的风险，即系统性风险

β_i ：证券对因素 i 的敏感程度，称为因子载荷系数

$E(r_i)$ ：各因子对应的风险收益率为0时，证券的期望收益率

CAPM与APT的联系

充分分散非系统性风险

当市场上的证券数目 $n \rightarrow +\infty$ ，非系统性风险 e_i 对组合收益的影响变得很小，其方差趋于0，也就是说，被非系统性风险被分散掉了，那么充分分散后的投资组合的单因素模型就成了：

$$r_p = E(r_p) + \beta_1 F_1 + \beta_2 F_2 + \beta_3 F_3 + \cdots + \beta_k F_k$$

对共同因素敏感程度相同的证券或组合在均衡时（即对非系统风险进行分散化后）将具有相同的收益率。不然，“无风险套利”机会便会出现，投资者就会大量持有该资产的头寸而不管投资者的风险厌恶程度如何，直至机会消失。

演变成为CAPM模型

如果影响因素就一个，且代表的就是市场因素，用市场的风险溢价来衡量，那么在没有系统性风险（ $F_p = 0$ ）的时候，组合收益应该等于无风险收益率，那么我们就得到了CAPM模型：

$$E(r_i) = r_f + \beta_i [E(r_M) - r_f]$$

Fama-French 三因素模型

产生背景

在CAPM这个资产定价模型提出来以后，陆续有学者发现CAPM无法完全解释资产的预期收益。有实证研究表明，股票市值、账面市值比等指标可以很好的解释股票收益，而CAPM却无法解释这些异象。

模型

模型认为公司市值和账面市值比能近似地代替一些未知的基本变量，代理未被CAPM模型完全解释的风险，在实证中可以较好地预测平均收益。

$$r_i - r_f = \alpha + \beta_{iM}(r_M - r_f) + \beta_{iSMB}SMB + \beta_{iHML}HML + e_i$$

r_i 为证券组合在观察期内的平均收益， α 为超额收益率， β_i 为对应因子载荷系数， e_i 为残差项

SMB代表小减大，即小市值股票与大市值股票收益率之差

HML代表高减低，高账面市值比股票减低账面市值比股票的收益率之差

Fama-French 做了些什么？

解释变量的构造

以一个年份的数据为基础，把股票按市值、账面市值比的大小分成 $2 \times 3 = 6$ 组，同时对每组股票取交集处理后，可以得到6个不重复的组别。



分别计算解释变量：

$$SML = small - large$$

$$HML = high - low$$

Fama-French 做了些什么？

被解释变量的构造

与之前方法类似，把之前的股票按市值与账面市值大小划分为 $5 \times 5 = 25$ 组，并取交集得到25组被解释变量数据。也就是说Fama做了25次回归，每次回归的解释变量都是一样的，被解释变量不相同。

$$r_i - r_f = \alpha + \beta_{iM}(r_M - r_f) + \beta_{iSMB}SMB + \beta_{iHML}HML + e_i$$

$$i = 1, 2 \dots 25$$

	Low BE/ME	2	3	4	High BE/ME
Small size	$y_{11} \rightarrow x$	$y_{12} \rightarrow x$	$y_{13} \rightarrow x$	$y_{14} \rightarrow x$	$y_{15} \rightarrow x$
2	$y_{21} \rightarrow x$	$y_{22} \rightarrow x$	$y_{23} \rightarrow x$	$y_{24} \rightarrow x$	$y_{25} \rightarrow x$
3	$y_{31} \rightarrow x$	$y_{32} \rightarrow x$	$y_{33} \rightarrow x$	$y_{34} \rightarrow x$	$y_{35} \rightarrow x$
4	$y_{41} \rightarrow x$	$y_{42} \rightarrow x$	$y_{43} \rightarrow x$	$y_{44} \rightarrow x$	$y_{45} \rightarrow x$
Large size	$y_{51} \rightarrow x$	$y_{52} \rightarrow x$	$y_{53} \rightarrow x$	$y_{54} \rightarrow x$	$y_{55} \rightarrow x$

Fama-French 做了些什么？

模型有效性的说明

对于不同类别的股票组合，我们可以看到 R^2 基本上都在0.9以上，三因子模型的拟合程度非常好，是一个有效的模型。

	small_BE/ME	1	2	3	big_BE/ME
small_size	0.850504	0.932359	0.401501	0.952128	0.942172
1	0.928505	0.947654	0.928649	0.949936	0.943393
2	0.960149	0.924252	0.962277	0.926240	0.903547
3	0.916107	0.945470	0.944791	0.933945	0.924949
big_size	0.932492	0.929273	0.901979	0.933192	0.942885

模型改进及扩展

Carhart四因素模型

- FF3它并不能合理解释证券市场中普遍存在的动量效应（股票的收益率有延续原来的运动方向的趋势，即过去一段时间收益率较高的股票在未来获得的收益率仍会高于过去收益率较低的股票）
- 加入了收益动量因素UMD

Fama-French 五因素模型

- 增加盈利水平因子（高/低盈利股票投资组合的回报之差）
- 增加投资水平因子（低/高再投资比例公司股票投资组合的回报之差）

回顾多因子模型

多因子模型

多因子模型的基础理论认为：股票的收益是由一些共同的因子来驱动的，不能被这些因子解释的部分被称为股票的“特质收益率”。

$$r_i = \sum_j \beta_j F_j + e_i$$

其中 r_i 为股票 i 收益率， F_j 为因子收益率， β_j 为股票 i 在因子 F_j 上的暴露程度，即系数， e_i 为股票 i 的特质收益率。

Alpha因子：用以解释股票的特质收益率。

风险因子：又称系统因子、或者beta，用以解释系统风险。

- 两者最大的区别就是是否被市场所熟知。
- **风险因子的前身一定是alpha因子**，只是当市场上越来越多人知道以后，它变成一种系统性的可解释的行为，于是就变成了系统（风险）因子。这也是为什么alpha因子会失效的原因。

Barra结构化风险模型

模型介绍

Barra模型就是有赖于风险因子构造的模型，主要用于风险管理和对已知收益归因。

无论是 Barra USE3 还是 USE4 模型，在横截面上对股票收益率进行回归时，均需包含行业因子 f_i 及风格因子 f_s ，二者的区别仅在于是否加入截距项（国家因子 f_c ）

$$\text{USE3: } r_n = \sum_i \beta_{ni} f_i + \sum_s \beta_{ns} f_s + e_i$$

$$\text{USE4: } r_n = f_c + \sum_i \beta_{ni} f_i + \sum_s \beta_{ns} f_s + e_i$$

β_{ni} 为股票n 在行业因子 i 上的暴露度，即一个哑变量，股票所属的行业因子暴露度为 1，否则为 0。

β_{ns} 表示股票在风格因子上的暴露度。可以理解为是把证券的收益分解到各个因子上面。

Why Barra?

均值-方差模型 (Markowitz Mean-Variance Model)

马科维茨认为，投资者可以用预期收益率 $E(R_p)$ ，以及收益率的标准差 σ_p 来完全构造和衡量一个投资组合。

依据他的观点，一个资产数目为 N 的投资组合（假设各资产头寸相同），如果已知每一个资产收益率的方差 $Var(R_i)$ ，和资产两两之间的协方差 $Cov(R_i, R_j)$ ，则我们可以计算这个投资组合的方差：

$$Var(R_p) = Var(R_1 + R_2 + \cdots + R_N) = \sum_{i=1}^N Var(R_i) + \sum_{i=j=1}^N Cov(R_i R_j)$$

对于一个包含 N 个资产的投资组合，我们需要计算 $N * (N-1) / 2$ 个协方差，用协方差矩阵来看就是：

$$\begin{pmatrix} Var(R_1) & Cov(R_1 R_2) & \cdots & Cov(R_1 R_N) \\ Cov(R_2 R_1) & Var(R_2) & \cdots & Cov(R_2 R_N) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(R_N R_1) & Cov(R_N R_2) & \cdots & Var(R_N) \end{pmatrix}$$

Why Barra?

$$\begin{pmatrix} Var(R_1) & Cov(R_1R_2) & \cdots & Cov(R_1R_N) \\ Cov(R_2R_1) & Var(R_2) & \cdots & Cov(R_2R_N) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(R_NR_1) & Cov(R_NR_2) & \cdots & Var(R_N) \end{pmatrix}$$

- (1) **数据量要求大**。要对协方差矩阵实现较为准确的估计，需要保证观测值数目大于矩阵的维数。考虑以沪深 300 的 300 个成分股作为投资组合，以月度数据计算经验协方差矩阵，则需要至少 $300/12 = 25$ 年的数据，因此缺乏现实可行性；
- (2) 依据历史数据进行协方差估计**无法反映投资组合中资产的结构性变化**（例如并购）；
- (3) 大量资产两两之间的协方差计算，容易出现**多重比较谬误**的问题，因而引起资产之间相关性的错误判断。
- (4) 历史数据中包含大量的噪音，因此简单使用资产的协方差矩阵进行预测会造成较大的偏差。

Why Barra?

核心思想是，我们可以选取一系列公共因子和特异因子来描述一个投资组合的风险。常用的公共因子有所属行业，成长性，市盈率等，特异因子则是和公共因子相对的概念，用于解释每个资产的收益率中不能用公共因子解释的部分。基于这个思路，投资组合的收益率 R_p 可以用资产头寸 W ，因子暴露矩阵 B ，因子收益率 f ，资产收益率向量 R ，特异因子收益率 e 表示：

$$R = Bf + e$$

$$R_p = W^T R = W^T (Bf + e)$$

假设特质因子收益率与公共因子收益率不相关，并且每只股票的特质因子收益率也不相关。那么在上述表达式的基础上，可以得到组合的风险结构为：

$$\sigma_p = \sqrt{W^T [BCov(f, f)B^T + \Delta]W}$$

其中 Δ 为特异因子收益率协方差矩阵。

投资组合的风险可以用因子收益率的协方差矩阵，而非投资组合中资产的收益率的协方差矩阵来描述。

How Barra?

Barra模型主要提供3个重要值：因子暴露、因子收益的协方差矩阵、个股的特别收益。

- 因子暴露：直接通过给定信息计算得到。
- 因子收益的协方差矩阵：通过横截面回归得到多组因子收益率数据，再计算因子收益率的协方差矩阵。
- 个股的特别收益：即为横截面回归得到的残差项。

通过上述三项，我们就可以很好地估计出组合的系统风险。

$$\sigma_p = \sqrt{W^T [BCov(R, R)B^T + \Delta]W}$$

How Barra?

- 因子收益的协方差矩阵：通过横截面回归得到多组因子收益率数据，再计算因子收益率的协方差矩阵。

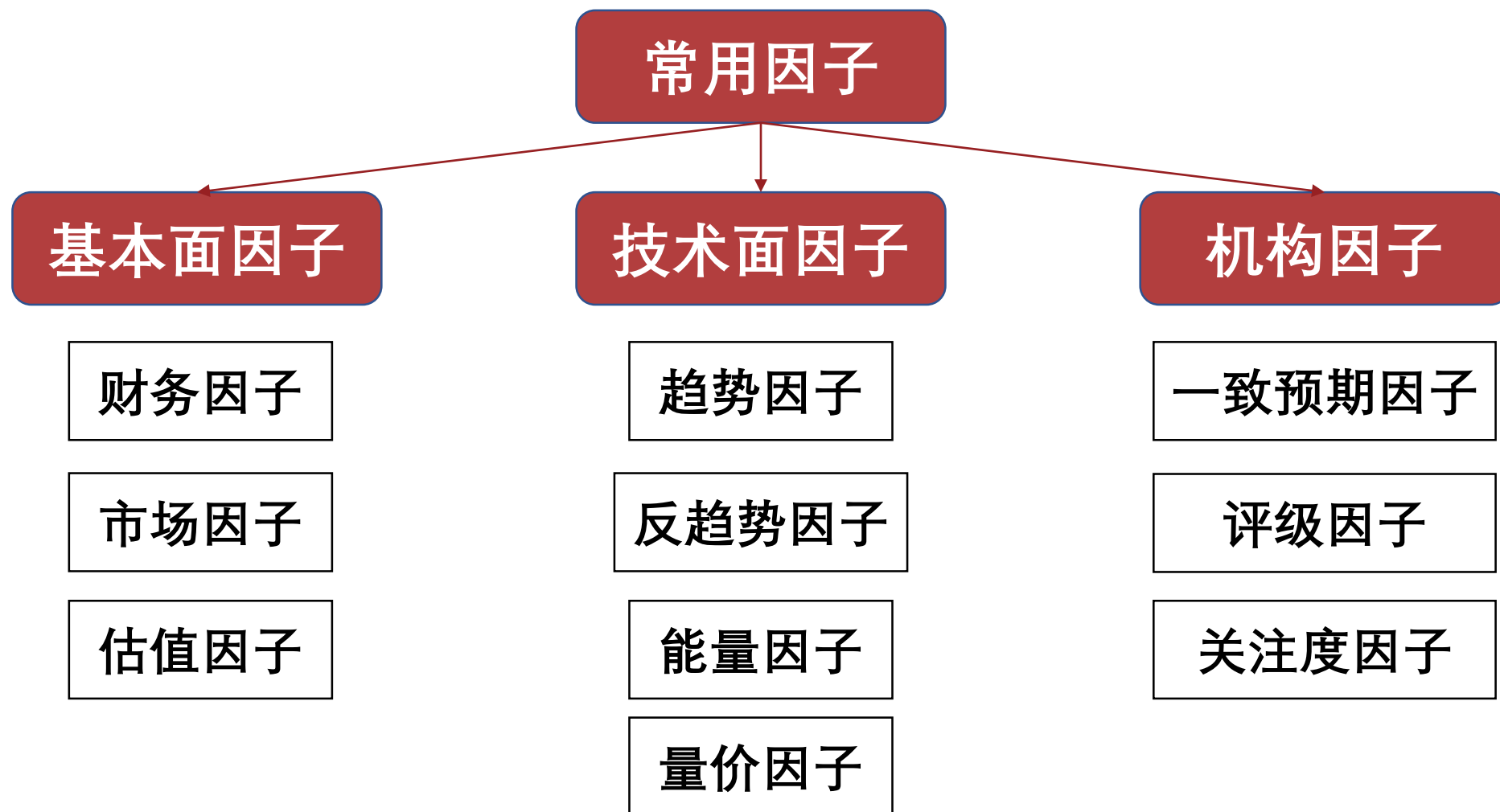
$$\text{USE4: } r_n = f_c + \sum_i \beta_{ni} f_i + \sum_s \beta_{ns} f_s + e_i$$

采用市值的平方根加权最小二乘WLS方法进行回归

- 解决特质收益率的异方差性
- 很多研究观察到大市值股票的特质收益率方差通常较小，因此在实际回归中我们将以市值的平方根占比作为每只股票的回归权重

2. 因子类型及其行业分布规律

因子介绍



2. 因子类型及其行业分布规律

因子介绍

基本面因子

财务因子

1. 短期偿债因子
2. 长期偿债（资本结构）因子
3. 经营因子
4. 成长因子
5. 盈利因子

市场因子

1. 规模因子
2. 持股情况因子
3. 流动性因子

估值因子

1. 股价比值类：PE PB PCF PS
2. 企业价值类：EV/FCF
EV/EBITDA
3. 股利政策类：股息率、送股比

2. 因子类型及其行业分布规律

因子介绍

财务因子

短期偿债因子

1. 流动比率
剔除金融类

$$\text{流动比率} = \frac{\text{流动资产}}{\text{流动负债}}$$

2. 速动比率
剔除金融类

$$\begin{aligned} \text{速动比率} \\ &= \frac{\text{流动资产} - \text{存货}}{\text{流动负债}} \end{aligned}$$

3. 超速动比率
剔除金融类

$$\text{超速动比率} = \frac{\text{货币资金} + \text{短期投资} + \text{应收票据} + \text{应收账款}}{\text{流动负债}}$$

2. 因子类型及其行业分布规律

因子介绍

财务因子

长期偿债（资本结构）因子

1. 资产负债率

剔除金融类，作用不大

$$\text{资产负债率} = \frac{\text{总负债}}{\text{总资产}}$$

2. 股东权益比

剔除金融类，作用不大

$$\text{股东权益比} = \frac{\text{股东权益}}{\text{总资产}}$$

3. 现金占比

剔除金融类

$$\text{现金占比} = \frac{\text{货币资金}}{\text{总资产}}$$

2. 因子类型及其行业分布规律

因子介绍

财务因子

经营因子

1. 存货周转率

作用不大，针对特殊行业

$$\text{存货周转率} = \frac{\text{营业成本}}{\text{平均存货}}$$

2. 应收账款周转率

$$\text{应收账款周转率} = \frac{\text{营业收入}}{\text{平均应收账款}}$$

3. 固定资产周转率

针对重（固定）资产行业

$$= \frac{\text{营业收入}}{\text{平均固定资产}}$$

4. 总资产周转率

$$\text{总资产周转率} = \frac{\text{营业收入}}{\text{平均总资产}}$$

5. 股东权益周转率

$$\text{股东权益周转率} = \frac{\text{营业收入}}{\text{平均股东权益}}$$

2. 因子类型及其行业分布规律

因子介绍

财务因子

成长因子

1. 总资产周转率

$$\text{总资产增长率} = \frac{\Delta \text{总资产}}{\text{总资产}_0}$$

2. 固定资产增长率

针对重（固定）资产行业

$$\text{固定资产增长率} = \frac{\Delta \text{固定资产}}{\text{固定资产}_0}$$

3. 净利润增长率

$$\text{净利润增长率} = \frac{\Delta \text{净利润}}{\text{净利润}_0}$$

4. 主营利润增长率

$$\text{主营利润增长率} = \frac{\Delta \text{主营利润}}{\text{主营利润}_0}$$

5. 主营收入增长率

$$\text{主营收入增长率} = \frac{\Delta \text{主营收入}}{\text{主营收入}_0}$$

2. 因子类型及其行业分布规律

因子介绍

财务因子

盈利因子

1. 销售毛利率 (GPM)

$$GPM = \frac{\text{主营收入} - \text{主营成本}}{\text{主营收入}}$$

2. 销售净利率 (NPM)

$$NPM = \frac{\text{净利润}}{\text{主营收入}}$$

3. 营业利润率 (OPM)

$$\text{营业利润率} = \frac{\text{营业利润}}{\text{主营收入}}$$

4. 资产收益率 (ROA)

$$ROA = \frac{\text{净利润}}{\text{总资产}}$$

5. 权益收益率 (ROE)

$$ROE = \frac{\text{净利润}}{\text{股东权益}}$$

2. 因子类型及其行业分布规律

因子介绍

市场因子

规模因子



1. 总市值

总市值 = 当日股价 × 总股本数

2. 流通市值

流通市值 = 当日股价 × 流通股本数

2. 因子类型及其行业分布规律

因子介绍

市场因子

持股因子

1. 十大流通股东占比 (效果不强) 2. 十大流通股东持股变动 (效果不强)

$$= \frac{\text{十大流通股东持股}}{\text{流通股本}}$$

$$= \frac{\text{十大流通股东持股变动}}{\text{流通股本}}$$

3. 机构持股比

$$= \frac{\text{机构持股数}}{\text{总股本}}$$

4. 机构持股变动

$$= \frac{\text{机构持股变动}}{\text{总股本}}$$

5. 机构家数
查看基金持股明细数据

2. 因子类型及其行业分布规律

因子介绍

市场因子

流动性因子

1. 换手率（20日、60日，**负向相关**）

$$= \frac{\text{成交量}}{\text{流通股本}}$$

2. 换手率变动（**负向相关**）

$$= \frac{\text{最近20日换手率}}{\text{最近60日换手率}}$$

3. 股东数（**效果不强**）

统计股东数目

4. 股东变动（**效果不强**）

$$= \frac{\Delta \text{股东数}}{\text{股东数}}$$

2. 因子类型及其行业分布规律

因子介绍

估值因子

股价比值类

1. 市盈率 (有盈利的企业)

$$PE_{TTM} = \frac{\text{总市值}}{\text{净利润}}$$

2. 市净率 (重资产企业)

$$PB_{LF} = \frac{\text{总市值}}{\text{股权账面价值}}$$

3. 市现率

$$PCF = \frac{\text{总市值}}{\text{经营性现金流量净额 (TTM)}}$$

4. 市销率 (成长期轻资产企业)

$$PS = \frac{\text{总市值}}{\text{主营业务收入 (TTM)}}$$

2. 因子类型及其行业分布规律

因子介绍

三种市盈率

1. PE(TTM, Trailing Twelve Months)——滚动市盈率

当前总市值/前面四个季度的总净利润

2. PE(LYR, Last Year Ratio) ——静态市盈率

当前总市值/去年净利润

3. PE(动) ——动态市盈率

当前总市值/预估今年净利润

2. 因子类型及其行业分布规律

因子介绍

两种市净率

1. PB——年报市净率

股价/净资产（该股上一年年末净资产）

2. PB(LF, Last File)——最新市净率

股价/净资产（最新财报净资产，一般为季报或者半年报）

2. 因子类型及其行业分布规律

因子介绍

估值因子

企业价值类

1. EV/EBITDA (重资产企业发展早期)

企业价值 (EV)

=MV of Debt and Equity - 现金及其等价物

EBITDA = 营业利润 + 折旧 + 摊销

$$PCF = \frac{\text{总市值}}{\text{经营性现金流量净额 (TTM)}}$$

2. EV/FCF

FCF = EBIT(1-T) + Interest(1-T) + D&A - Δ WC - Capex

4. 市销率

$$PS = \frac{\text{总市值}}{\text{主营业务收入 (TTM)}}$$

2. 因子类型及其行业分布规律

因子介绍

估值因子

股利政策类

1. 股息率（股利分配情况，**负相关**）

$$= \frac{Dividend}{Price}$$

2. 送股比例

e.g. 10送5

股份公司对原有股东采取无偿派发股票的行为。送股时，将上市公司的留存收益转入股本账户，留存收益包括盈余公积和未分配利润。

2. 因子类型及其行业分布规律

因子介绍

技术面因子

趋势指标

MA、MACD

反趋势指标

RSI、KDJ

能量指标

AR、BR

成交量指标

VMACD、
VROC

量价指标

OBV、MFI

2. 因子类型及其行业分布规律

因子介绍

技术面因子

趋势指标

1. MA(Moving Average)

黄金金叉：MA(5)上穿MA(10)

死亡金叉：MA(5)下穿MA(10)

2. MACD (Moving Average Convergence and Divergence)

买入信号：MACD由负转正

卖出信号：MACD由正转负

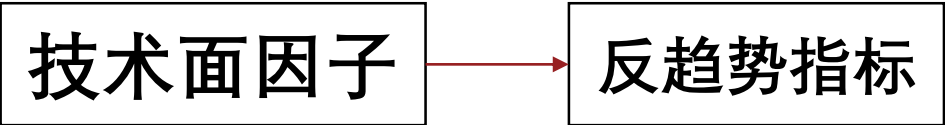
$DIFF = MA(12) - MA(26)$

$DEA = MA(DIFF, 9)$

$MACD = (DIFF - DEA) \times 2$

2. 因子类型及其行业分布规律

因子介绍



1. RSI (Relative Strength Index)

RS (相对强度) = $\sum[\Delta P_i]_+ / (\sum[\Delta P_i]_+ + \sum -[\Delta P_i]_-)$
通常取14天

RSI值 (%)	市场特征	投资操作
80-100	极强	卖出
50-80	强	买入
20-50	弱	卖出
0-20	极弱	买入

2. KDJ

$$K_t = 2/3 \times K_{t-1} + 1/3 \times RSV_t$$
$$RSV = \frac{CLOSE - LOW_N}{HIGH_N - LOW_N}$$
$$D_t = 2/3 \times D_{t-1} + 1/3 \times K_t$$
$$J_t = 3K_t - 2 D_t$$

情况	解读
D>80	超买现象
D<20	超卖现象
K上穿D	买入信号
K下穿D	卖出信号



2. 因子类型及其行业分布规律

因子介绍

技术面因子

能量指标

1. AR（人气指标） 反映当前市场多空双方力量发展对比的结果

$$AR = \frac{\sum(HIGH_t - OPEN_t)}{\sum(OPEN_t - LOW_t)}$$

2. BR（意愿指标） 反映当前市场多空双方力量争斗的结果

$$BR = \frac{\sum(HIGH_t - CLOSE_{t-1})}{\sum(CLOSE_{t-1} - LOW_t)}$$

策略：

- ①BR<AR, 且BR<100, 可考虑逢低买进
- ②BR<AR, 而AR<50时, 是买进信号
- ③BR急速上升, 而AR处在盘整或小跌时, 应逢高卖出

2. 因子类型及其行业分布规律

因子介绍

技术面因子

成交量指标

1. VMACD

$DIFF = MA(Vol, Sell, 12) - MA(Vol, Purchase, 12)$

$DEA = MA(DIFF, 9)$

$VMACD = DIFF - DEA$

2. VROC

$$VROC_t = (Vol_t - Vol_{t-n}) / Vol_{t-n}$$

买入信号：MACD由负转正
卖出信号：MACD由正转负

策略

- ① VROC向下跌破零，预示有主流资金撤退迹象。
- ② VROC向上突破零，预示有主流资金介入迹象。
- ③ 股价创新高，VROC未配合上升，显示上涨动力减弱。
- ④ 股价创新低，VROC未配合下降，显示下跌动力减弱。

2. 因子类型及其行业分布规律

因子介绍

技术面因子

量价指标

1. OBV能量潮(On Balance Volume)

累积交易量的折线图，价格上升则加交易量，反之则减交易量，价格不变则OBV不变

2. MFI

$TYP_t = average(high_t, low_t, close_t)$

$$V1 = \frac{sum(if(TYP_t > TYP_{t-1}, TYP_t \times VOL_t, 0), 14)}{sum(if(TYP_t < TYP_{t-1}, TYP_t \times VOL_t, 0), 14)}$$

$$MFI_t = 100 - \frac{100}{1 + V1}$$

策略

- ①当股价上升而OBV线下降，表示买盘无力，股价可能会回跌。
- ②股价下降时而OBV线上升，表示买盘旺盛，逢低接手强股，股价可能会止跌回升。
- ③OBV线缓慢上升，表示买气逐渐加强，为买进信号。
- ④OBV线急速上升时，表示力量将用尽为卖出信号。。

策略

- ①MFI<20，资金短期冷却，MFI突破20，资金转向
- ②MFI>80，资金短期过热，MFI跌破20，资金转向

2. 因子类型及其行业分布规律

因子介绍

机构因子

一致预期因子

评级因子

债券评级

关注度因子

因子绝对值

因子变化量

历史相对位置

2. 因子类型及其行业分布规律

因子介绍

机构因子

一致预期因子

- ①针对股票，挑选符合条件的研报（若分析师重复，则以该分析师最新的研报为准）
- ②考虑三个维度：业绩预测准确度、推荐股票市场反应、报告时效性

$$Score = (Accuracy + Impact) / 2 \times TimeDecay$$

2. 因子类型及其行业分布规律

因子分布规律

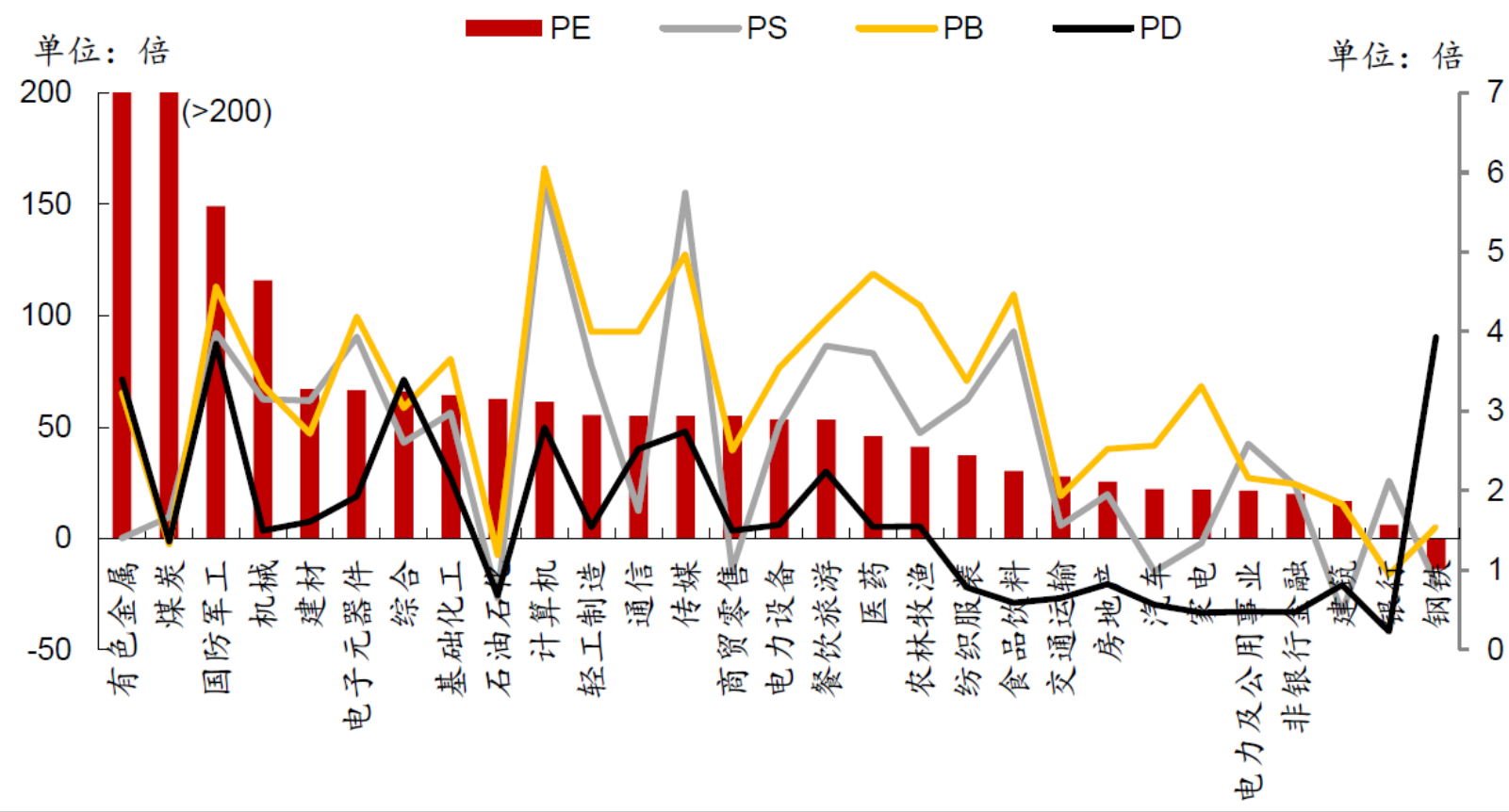
要点

- 估值因子在不同行业中，差异较大
- 估值因子在不同规模的公司中，差异较大

2. 因子类型及其行业分布规律

因子分布规律

图2： 各一级行业 PE、PB、PS、PD 值比较（2016/08/31）



在不同行业中：

- ①各种倍数相差最高达200倍
- ②钢铁、煤炭PB、PS低
- ③银行、非银金融PE低

2. 因子类型及其行业分布规律

因子分布规律

要点

- 行业中性分层法

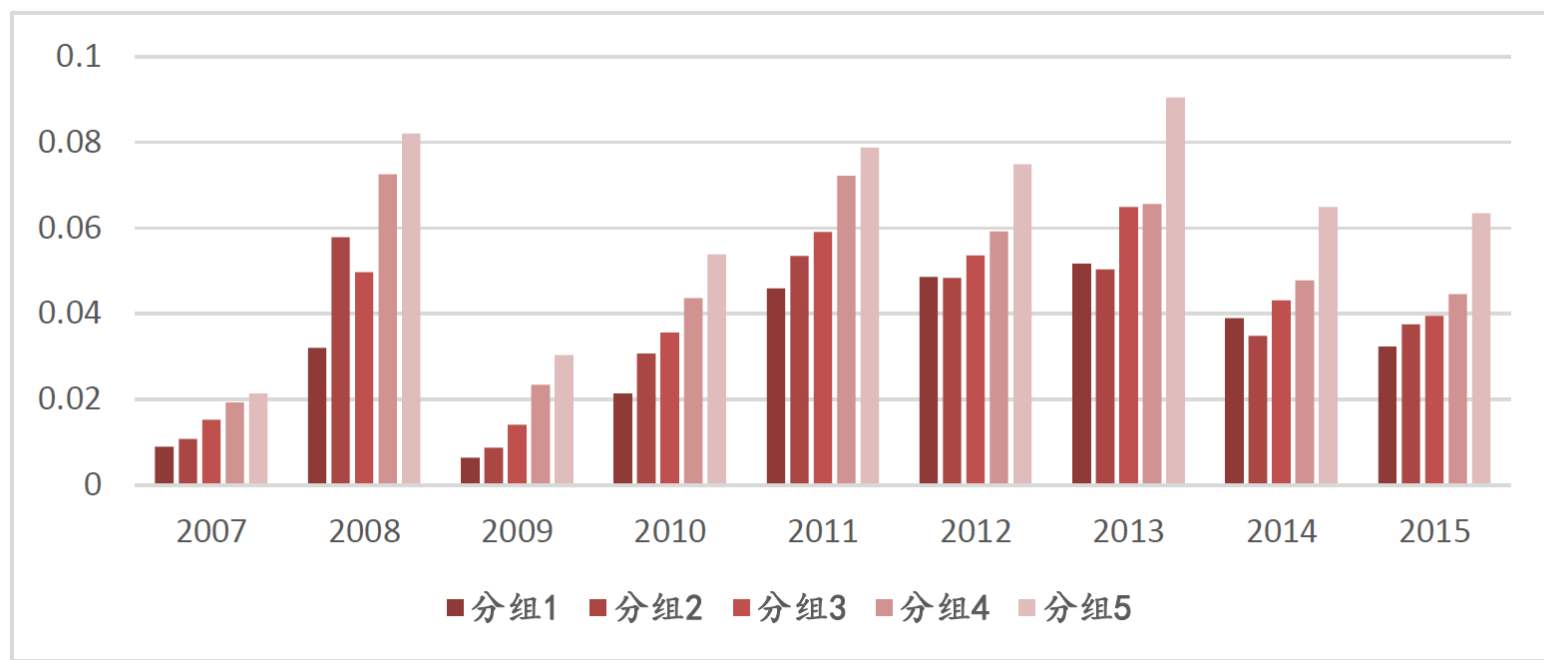
①在每个一级行业中分组，按照市值大小等分成5组（从小到大）

②合并：组号相同&跨行业

③进行EP的比较

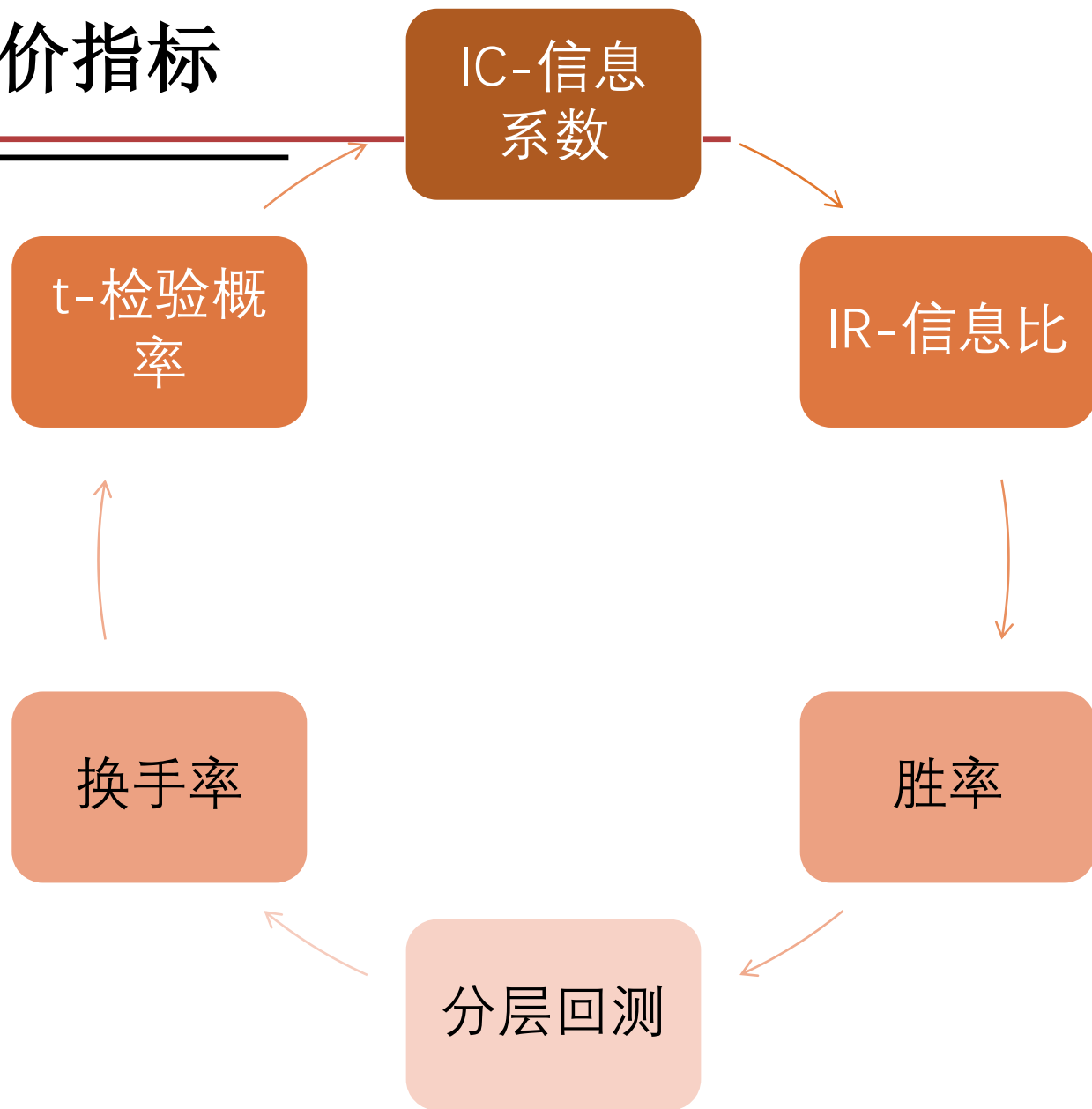
企业规模（市值）越大，
EP越大，PE越小

图3：按市值均分为五层的EP因子均值随时间变化图



资料来源：Wind，华泰证券研究所

3. 因子有效性评价指标



3. 因子有效性评价指标-因子IC-信息系数

概念

- 期初因子 与 期末股票收益率相关系数
- 相关系数绝对值越大，该因子选股能力越强
- 相关系数为负，因子值越小越好；相关系数为正，因子值越大越好

因子IR-信息比率

概念

- 投资组合相对于基准组合的超额收益稳定性
- IR越大，该因子所指导的投资组合的超额收益越稳定，该因子的预测能力越好

$$IC = \frac{Cov(factor, return)}{\sqrt{Var(factor)Var(return)}}$$

$$IR = \frac{R_P - R_B}{\sqrt{Var(R_P - R_B)}}$$

3. 因子有效性评价指标-胜率

概念

- 在时间样本空间里，因子收益为正的次数占全样本的比值
- 比值越大，概率越大，该因子挑选的超配组合收益战胜低配组合收益的成功率越高
- 该因子的有效性越显著

年化收益率

作用

- 通过将收益率年化，可以更加直观的看到因子的收益有效性

$$P_{win} = \sum_{i=1}^n 1(R_P > 0) / n$$

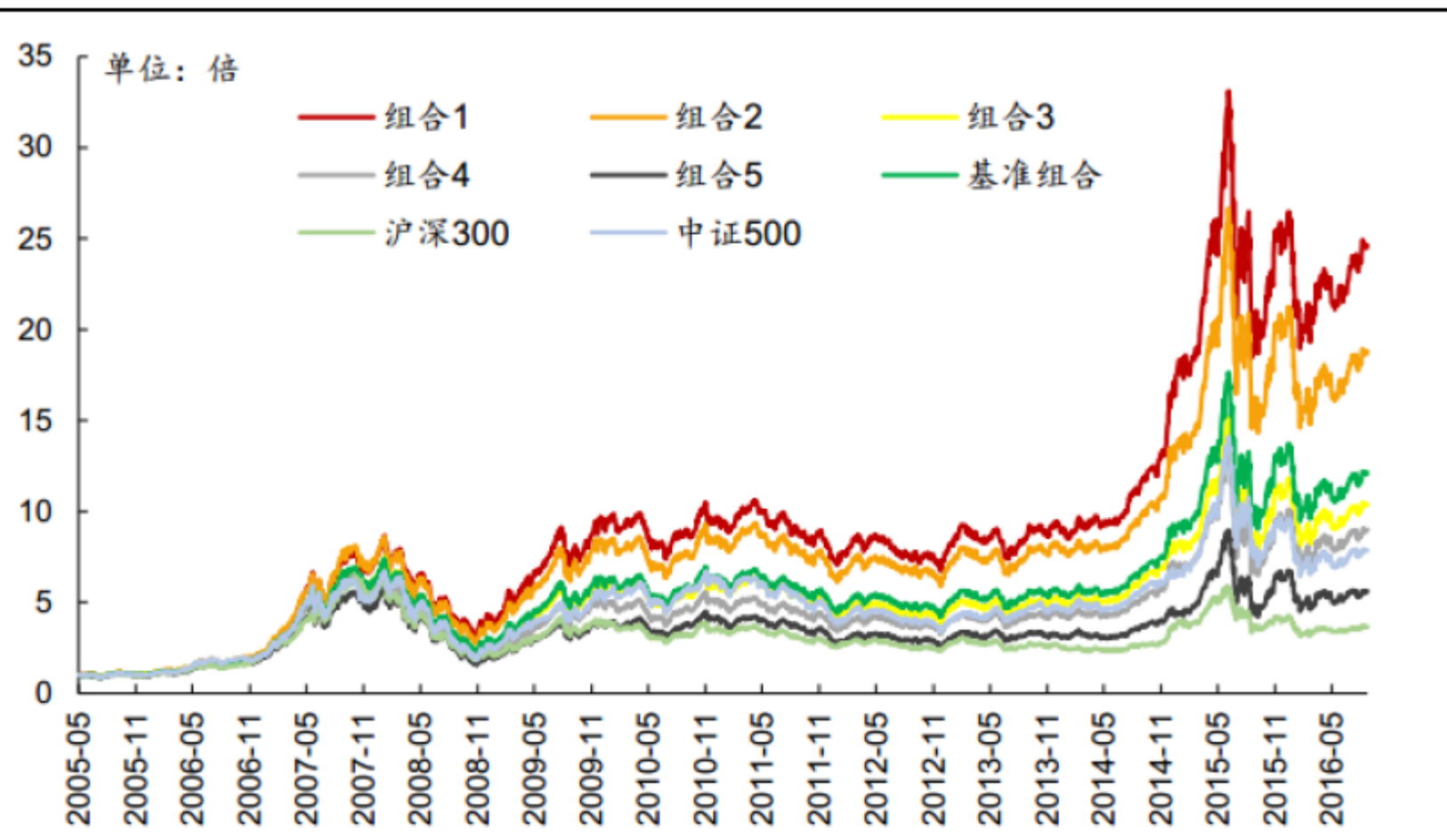
$$R_{year} = \left(\frac{P_{end}}{P_{start}} \right)^{\frac{250}{n}} - 1$$

3. 因子有效性评价方法-分层回测

概念

- 按照因子值
- 对这10个排
- 对10组投资

BP因子分层组合回测净值



投资组合

对比

3. 因子有效性评价指标-换手率

概念

- 每个月末投资组合中被更换的股票所占原股票的比例
- 由于更换投资组合会产生手续费和冲击成本，因子的换手率应在保证收益率的前提下尽量低

t-检验概率

$$t = \frac{\bar{R}_P}{s(R_P)/\sqrt{n}} \sim t(n-1)$$

- 实证检验表明，当P-value<0.2时，就可以证明出该因子具有较好的选股效果