

# PE-ROE 回归关系 - 行业选股测试

根据申万宏源研报《PE-ROE 股指原理与预期差选股》编写

## 选择股票池条件：

- 股票市值在历史上曾经大于过 100 亿

选择 1665 只股票作为初始股票池。

## 行业分类：

研报中采用中信二级分类指标，并算出回归的 MSE，认为用二级指标分类的效果比一级分类好。

中信一级分类指标 29 个类别，中信二级分类指标约 90 个类别。

由于数据库中没有完全一样的指标，我才用可以获取到的新浪财经分类指标，供 49 个类别，介于两者之间。

## 获取财务数据：

由于每次获取到财务报表的时间不一，故将季报的时间统一为报告有效期后的 20 天，将年报的时间统一为报告有效期后的 80 天。具体时间如下：

	有效日期	获取日期（回测时使用）
一季度报	3 月 31 日	4 月 20 日
半年报	6 月 30 日	7 月 20 日
三季度报	9 月 30 日	10 月 20 日
年报	12 月 31 日	次年 3 月 21 日

## 仓位分配：

按照公司流通市值作为权重。

## 调仓时间：

依据研报，调仓时间采用每月的最后一天。但财务数据的信号为季度频率。

这中间不免有一些空白期，此时根据市值变化来调仓。

## 数据剔除：

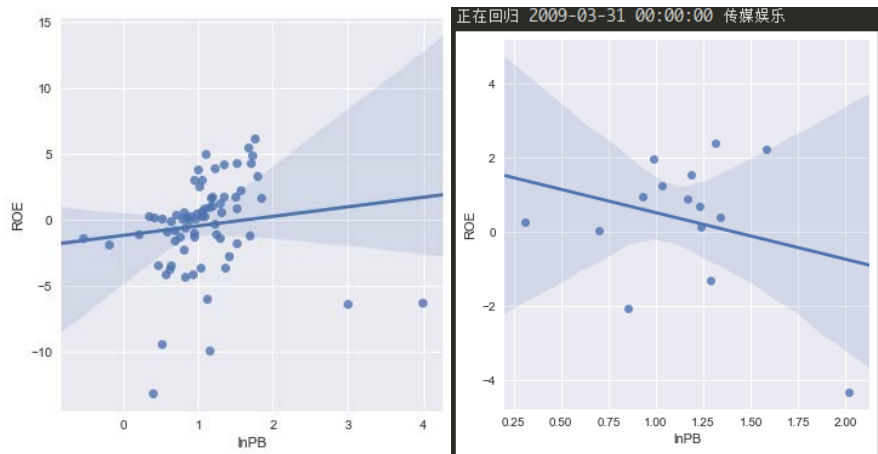
采用  $u \pm 2\sigma$  的剔除方式。剔除前 58189 条记录，剔除后 54603 条记录。

## 拟合方式：

将 ROE 对  $\ln(P/B)$  做回归，线性拟合。

若 ROE 实际值高于拟合结果线，则认为有超额 ROE，可以买入。

在查看结果是发现，即使经过异常值剔除，还是会有一些不正常的情况。



例如左图：正确的回归线显然应该更加陡峭一些；右图，正确的回归线应该斜率大于零。

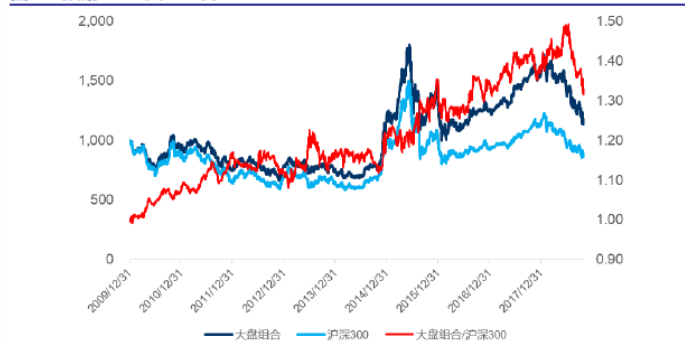
## 结果比较和分析

超额收益的计算方法是：选股策略净值/标的净值。是一个倍数的概念。

1. 首先，由于选择的是市值>100 亿的股票，所以可以用 HS300 作为标的。



图 5：预期 ROE 大盘组合



资料来源：申万宏源研究，注：将基准指数调整为 2009 年末 1000 点，之后均作此处理

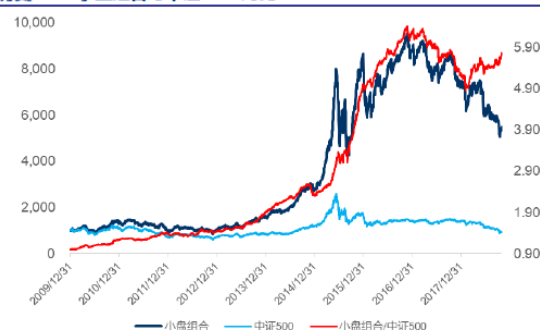
选股策略从 2009 至 2018，最终净值 2.3075，超额收益 1.9076 倍。除在 17 年有过下降外，超额收益保持比较稳定的增长。

研报大盘组合的超额收益，对于 HS300 而言约为 1.30 左右。

2. 作为尝试，用中证 500 作为标的，发现超额收益变得不稳定。



图 9：历史 ROE 小盘组合与中证 500 对比



资料来源：申万宏源研究

最终超额收益 1.6271 倍，主要的超额收益来自于 2015 年及以后。  
而研报中的小盘组合，超额收益能够达到约 5 倍。

经过上述对比，发现超额收益的来源可以分为两段：

15 年以前主要来自中小盘股，15 年以后主要来自大盘价值股。因此对比不同的标的，会表现出很不同结果。

究其原因，是我采用的股票池筛选条件是市值大于 100 亿，其中既包含了大盘股，也包含了很多中盘、小盘股。如果要严格筛选大盘股，则会导致回归数据点太少，结果不太正确。

3. 鉴于上述结论，我的投资组合实际上是介于大盘和中小盘的股票。因此用上证综指作为标的也是可行的。

对比发现，超额收益为 2.2088，而且增长相当稳定。



结语：市值因子在这个超额收益中扮演了很重要的角色。15 年及以前的超额收益来源，主要是中小盘股；15 年以后主要是大盘股。

新增部分：

- 修改选股条件为“股票市值在 2015 年以前大于 100 亿”，以避免选一些小盘股（只在牛市市值超过 100 亿），真正选出大盘股

通过上述选股，选出的数量为 700 个，可以认为已经完全提取出了大盘股，只不混杂小盘股。

此时，对于 HS300 的超额收益增长稳定，而且达到了 2 倍；尤其是 15 年以后，由于只包含了大盘股，超额收益增长稳定，较之前的结果有了提升。



进一步分析对比各年超额收益。



