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：**

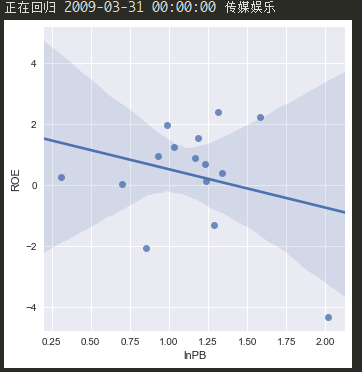
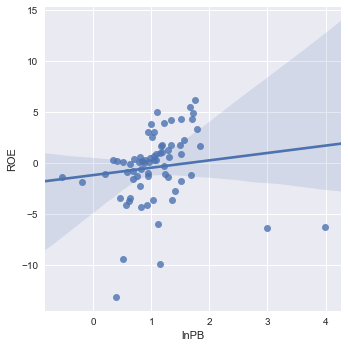
采用u±2\*σ的剔除方式。剔除前58189条记录，剔除后54603条记录。

**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。

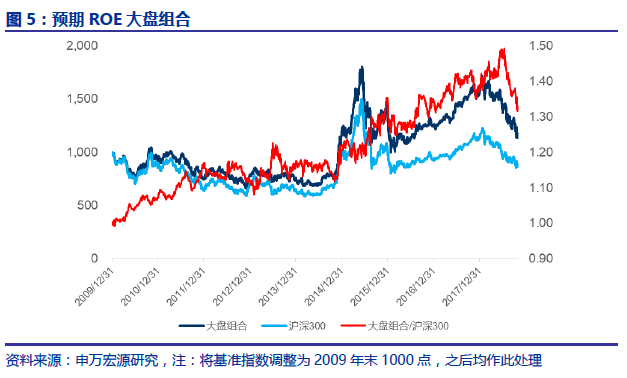
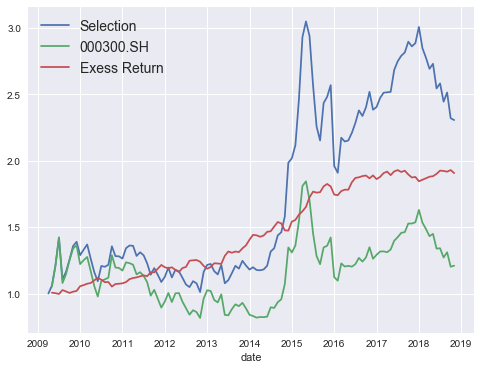


例如左图：正确的回归线显然应该更加陡峭一些；右图，正确的回归线应该斜率大于零。

**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。

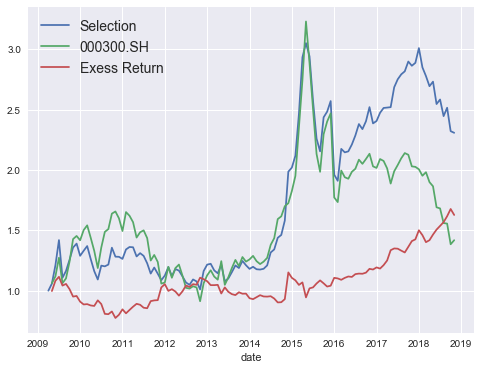
1. 首先，由于选择的是市值>100亿的股票，所以可以用HS300作为标的。



**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。

1. 作为尝试，用中证500作为标的，发现超额收益变得不稳定。

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。

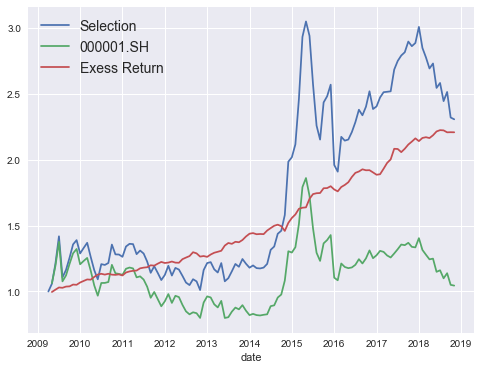
**经过上述对比，发现超额收益的来源可以分为两段：**

**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。**

1. 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。



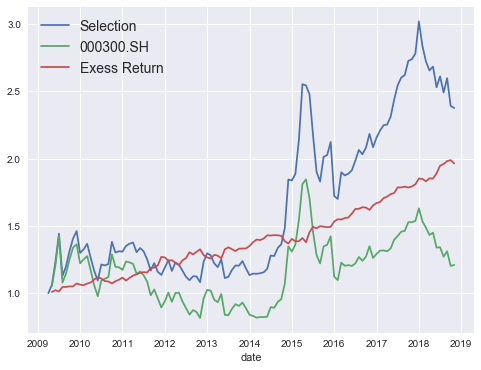
结语：市值因子在这个超额收益中扮演了很重要的角色。15年及以前的超额收益来源，主要是中小盘股；15年以后主要是大盘股。

**新增部分：**

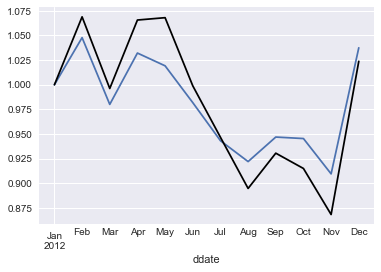
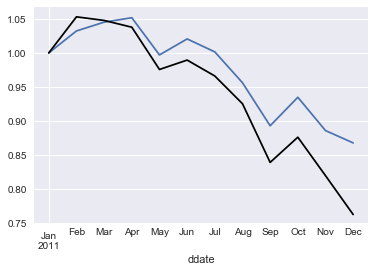
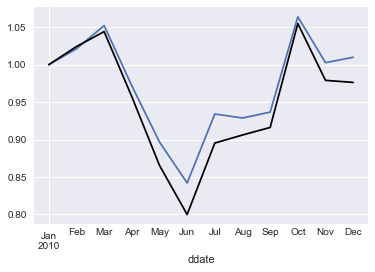
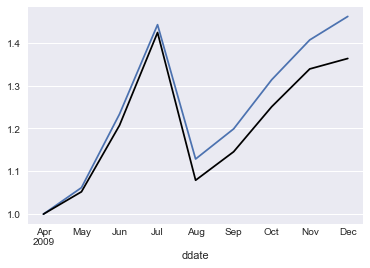
* **修改选股条件为“**股票市值在2015年以前大于100亿**”，以避免选一些小盘股（只在牛市市值超过100亿），真正选出大盘股**

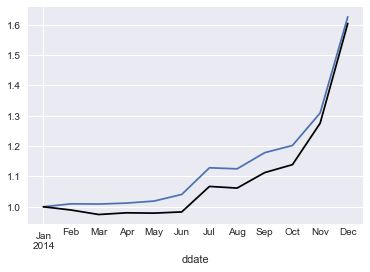
**通过上述选股，选出的数量为700个，可以认为已经完全提取出了大盘股，只不混杂小盘股。**

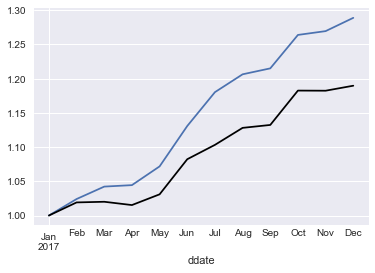
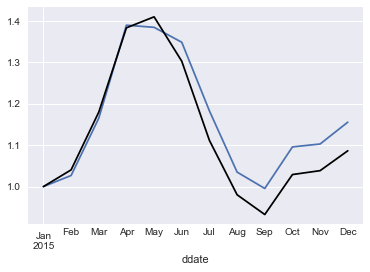
**此时，对于HS300的超额收益增长稳定，而且达到了2倍；尤其是15年以后，由于只包含了大盘股，超额收益增长稳定，较之前的结果有了提升。**



**进一步分析对比各年超额收益。**







**12/5与分析师通电话后，获知以下细节：**

**Q: u+-3\*σ的剔除函数，不知道是在一个时间横截面上剔除，还是在一个个股的横截面上剔除？**

**A: 只是剔除了PB>30**

**Q: 回归斜率线出现负数如何处理？**

**A：可以理解为中国股市出现不理智现象，应剔除这一时间、这一个行业**

**Q: 在行业分类上，研报中采用了申万二级，我这边采用了我所能找到的Sina一级分类。申万二级的分类达到90个类别，会不会造成每一个行业内数据点数量太少？**

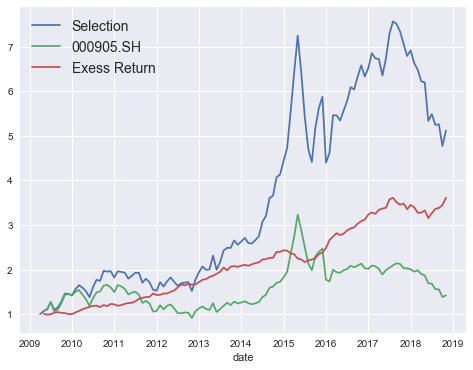
**A: 细分以后，确实会很少，只要斜率不为负，不影响结果。**

**Q：股票池需要进一步细分为中盘，大盘，小盘，但是细分标准需要统一？**

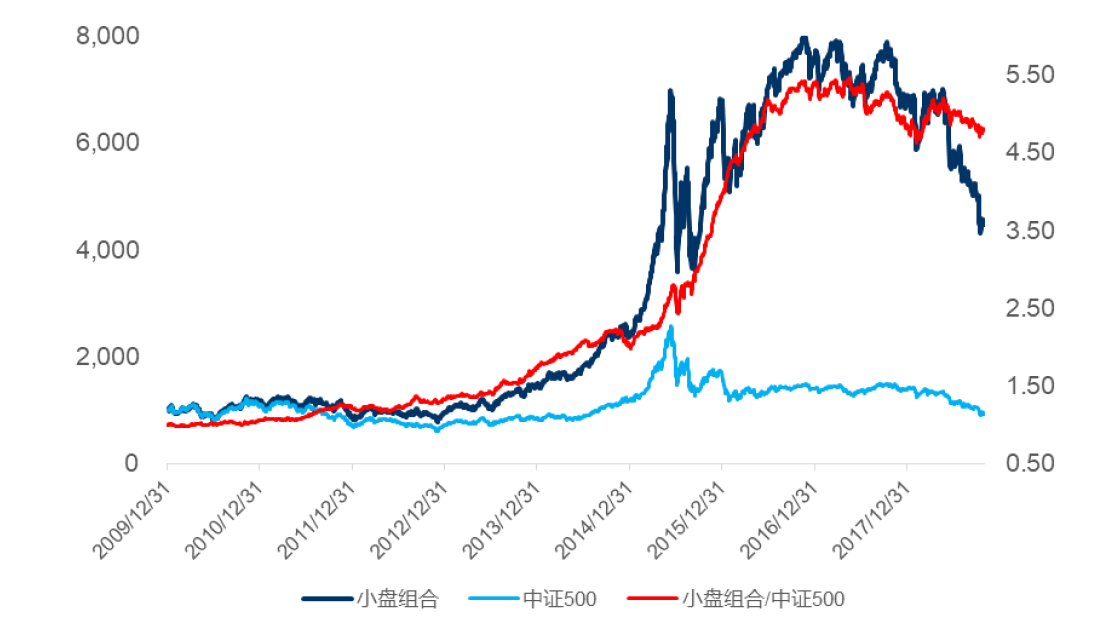
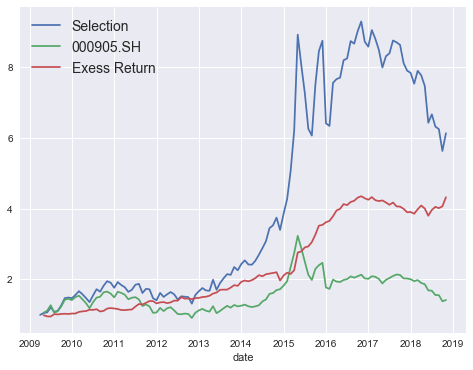
**A: 细分标准是：市值排名前20%大盘股 后30%小盘股**

**更新结果：**

**小盘股（以中证500为标的）**



**超额收益达到近4倍，但是对比研报的近6倍还是有所差距。**



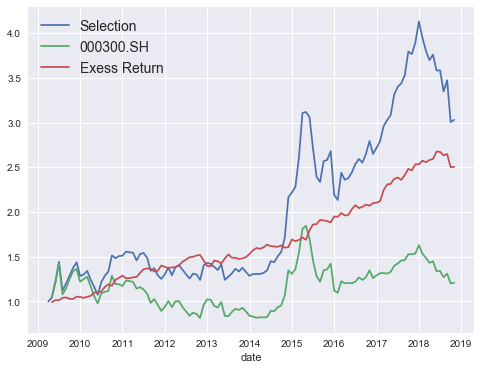
**修正小盘的权重，修正为等值加权后，效果和研报几乎完全相同。**

**只是开仓时间的区别，导致了最终超额收益的差异：研报巧妙地采用了09年末作为起点，因此整个回测时间段的标的涨幅几乎为0%。在本文回测中，用了09年初作为起点，因此整个回测时间段标的涨幅为50%；另外，观察到09年的超额收益几乎可以忽略不计。综上所述，已经成功复制了结果。**

**大盘股（以沪深300为标的）**



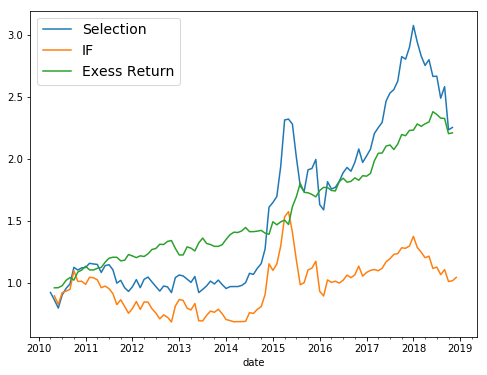
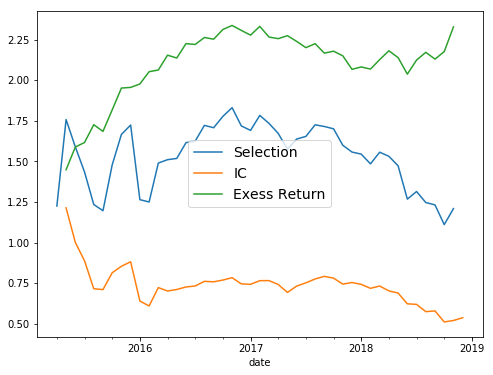
**超额收益约为1.75倍，略优于研报的1.4倍。**



**如果加上原先的选股条件：“2015年以前市值>100亿”，效果会更好**

**和期货对比**

**起始日期：小盘对比IC：2015-04-16，大盘对比IF：2010-04-30**



**IF超额收益统计**

策略收益和标的收益直接相减，单利相加

****

**HS300超额收益统计**

****