目录

[SMB组合初探 2](#_Toc43480611)

[Barrar模型 4](#_Toc43480612)

[FF ts+cs（FAMA MACBETH） 7](#_Toc43480613)

[BARRA框架与FF框架的比较 11](#_Toc43480614)

[FF框架下的两个定价模型比较 11](#_Toc43480615)

[再品SMB组合 12](#_Toc43480616)

# SMB组合初探

1. 构建A股SMB组合： 逻辑：
   1. 对现存于市场的股票进行清洗，筛除停牌ST、待退市股票（借助JQ回测框架，可在每个rebalance时点得到哪些股票有这些标记的信息）
   2. 每个月第一个交易日获取各股上个月末收盘市值，根据百分位数划分成五组，市值最大组标记为1，最小组标记为5
   3. 得到下一个月经过筛选后的股票列表内的各股return
   4. 根据value weighted计算1、2、3、4、5rank组的return
   5. 最后得到1、2、3、4、5组时间序列上各月的return
   6. 组5收益率-组1收益率得到dollar neutral SMB monthly return
2. A股市场SMB因子组合表现
3. 时间段划分：（复利对趋势判断不友好，此处参考单利图像划分d）
   1. 2005年2月-2007年11月
   2. 2007年11月至2010年10月
   3. 2010年10月2013年1月
   4. 2013年1月至2016年12月
   5. 2016年12月至2018年2月
   6. 2018年2月至2020年5月
4. 将全局区间根据图像特征定义为三种阶段
   1. 有效阶段（净值阶段性上涨）
      1. 2007年11月至2010年10月
      2. 2013年1月至2016年12月
   2. 反转阶段（净值阶段性下跌）
      1. 2016年12月至2018年2月
   3. 无效阶段（净值在一定水平波动）
      1. 2005年2月-2007年11月
      2. 2010年10月2013年1月
      3. 2018年2月至2020年5月

（后续子区间研究将按此划分展开）

# Barrar模型

横截面回归，LHS 各股票return，RHS 各股票特征（作为factor loading）

回归出risk premium

若risk premium为正，return随char递增，反之递减

以size作为唯一char研究。

预期结果全局段 risk premium为负，即股票市值越大股票return越小；有效阶段负；无效阶段估计值不显著；反转段正

估计值数值大小参考意义不大，重点关注正负情况

（log size;lag0; non-normalized）

对全时段回归结果：

Avg para:0.00597

T:3.66

注：不符逻辑

区间时段回归结果：

1. 有效阶段（净值阶段性上涨）
2. 2007年11月至2010年10月

Para: -0.0004

T: -0.1257

1. 2013年1月至2016年12月

Para: 0.00077

T: 0.23

1. 反转阶段（净值阶段性下跌）
   * 1. 2016年12月至2018年2月

Para: 0.01506

T: 3.93

注：符合逻辑

1. 无效阶段（净值在一定水平波动）
   * 1. 2005年2月-2007年11月

Para: 0.01401

T: 2.6

* + 1. 2010年10月2013年1月

Para:0.00612

T: 2.06

* + 1. 2018年2月至2020年5月

Para:0.0088

T: 3.72

（log size;lag0;normalized）

对全时段回归结果：

Avg para: 0.005919

T :3.666221

注：不符逻辑

1. 有效阶段（净值阶段性上涨）
2. 2007年11月至2010年10月

Para: -0.001799326

T: -0.548428639

1. 2013年1月至2016年12月

Para: 0.000661295

T: 0.199363357

注：不能较好判断因子有效性

1. 反转阶段（净值阶段性下跌）
2. 2016年12月至2018年2月

Para: 0.011733735

T: 3.750313821

注：符合逻辑

1. 无效阶段（净值在一定水平波动）
2. 2005年2月-2007年11月

Para: 0.014743215

T: 2.821275994

1. 2010年10月2013年1月

Para: 0.006102102

T: 2.146339557

1. 2018年2月至2020年5月

Para: 0.010454062

T: 3.872674035

（log size;lag0;normalized）

对全时段回归结果：

Avg para: -0.007289144

T : -4.002411613

注：符合逻辑

1. 有效阶段（净值阶段性上涨）
2. 2007年11月至2010年10月

Para: -0.012072641

T: -3.81952119

1. 2013年1月至2016年12月

Para: -0.018866051

T: -4.199422095

注：lag1对有效阶段有较好判断

1. 反转阶段（净值阶段性下跌）
2. 2016年12月至2018年2月

Para: 0.002084408

T: 0.453528481

注：para符合逻辑，但不显著

1. 无效阶段（净值在一定水平波动）
2. 2005年2月-2007年11月

Para: 0.001606882

T: 0.31455253

1. 2010年10月2013年1月

Para: -0.001953602

T: -0.69955575

1. 2018年2月至2020年5月

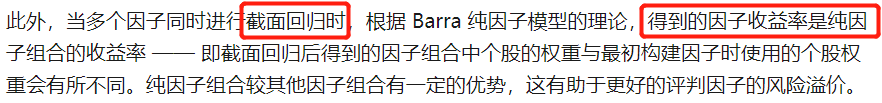
Para: -0.004026519

T: -0.968902935

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 总结 | 有效阶段 | 无效阶段 | 反转阶段 | 全局 |
| Log size; non-normalized; lag 0 | 不符逻辑 | 三段皆显著 | 符合逻辑 正号且显著 | 不符合逻辑 |
| Log size; normalized; lag 0 | 不符逻辑 | 三段皆显著 | 符合逻辑 正号且显著 | 不符合逻辑 |
| Log size; normalized; lag 1 | 符合逻辑，负号及T显著 | 三段皆不显著  无明显size和return的趋势 | 符合逻辑 正号且但不显著 | 不符合逻辑 |

思考：BARRA框架似乎能用于判断单个因子的有效性，而FF框架侧重于哪个定价模型最优（下一节探讨的内容）

# FF ts+cs（FAMA MACBETH）



（截图来源：<https://zhuanlan.zhihu.com/p/40984029> 搜索关键词：权重）

验证TS+CS回归出来的risk factor return 估计值和SMB return有关联。

TS\_window=6; CS\_window=6计算出各时间点回溯6个月的average return estimator 与对应真实SMB组合各时间点回溯6个月的avg return比较。

预期：SMB return与est return走势接近

注：符合预期，avg return 走势大部分时候相近，可以理解为是与SMB个股权重不同‘纯因子组合收益’

TS\_window=3; CS\_window=6

TS\_window=12; CS\_window=6

TS\_window=6; CS\_window=1

TS\_window=3; CS\_window=1

思考：FF框架主要考虑的是哪个定价模型α更小，即哪个定价模型最佳，但是针对子区间检验单个因子的有效性似乎帮助不大

# BARRA框架与FF框架的比较

前者factor loading 是 char（size数据）， FF factor loading 是beta，两个框架的风险暴露含义截然不同，前者是市值大小，后者是对SMB return的‘敏感度’

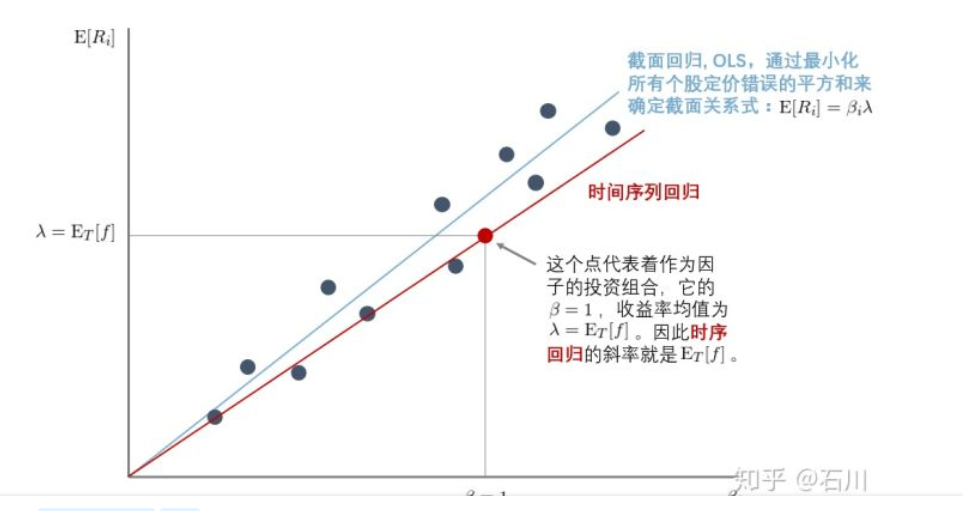
两者有必然联系吗？问： char高必然beta高吗？否定

# FF框架下的两个定价模型比较

1. TS
2. TS+CS

两个定价模型最后都是为了展示横截面上不同股票因子暴露不同带来的对应因子的return贡献，最终目的都是为了看α是否为0来判断因子组合的优劣

（来源<https://zhuanlan.zhihu.com/p/40984029>，搜索关键词时序回归 vs 截面回归）



总结来说：FF不适用于单个因子的有效性验证，barra可以一定程度用于检验单个因子是否有效

# 再品SMB组合

深挖SMB组合在三种不同阶段的细节

（先抛小结，下文为具体分段）

小结：

1. BIG与SMB在任何阶段都无线性关系（但美股里BIG 与SMB的拟合度较SMALL与SMB拟合度相当）
2. 有效阶段Rs Rb两极化（return一正一负）情况较少，而在反转及无效阶段两极化数量增加，更重要的是在反转时 两极化数量占比非常大！
3. 通常，SMALL组合和ZZ500的拟合度非常高，BIG组合和HS300的拟合度非常高，但在反转或无效阶段，SMALL与ZZ500的拟合度会降低，而BIG与HS300的拟合度全时段保持非常高拟合的水平（R^2 0.99），在有效阶段可以考虑用一定比例的多ZZ500空HS300在等效替代SMB组合，而其他阶段可以考虑多SMALL组合，空HS300，因为ZZ500不一定具对SMALL组合有强代表性
4. SMB组合对HS300指数有效市场中性（因为HS300与BIG组合的高拟合度，可替代，做空BIG即做空HS300），而对ZZ500指数不一定有效市场中性
5. SMALL与SMB的线性关系 有待进一步挖掘更多信息

将SMB分解成return\_small和return\_big

1. 有效阶段（净值阶段性上涨）

2007年11月至2010年10月

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | small return | big return | SMB return | HS300 | ZZ500 |
| std | 0.130853563 | 0.11977131 | 0.065806257 | 0.120967898 | 0.127216973 |
| mean | 0.027133487 | -0.00666296 | 0.033796447 | -0.00682777 | 0.011240974 |
| sharp\_ratio  (ignore Rf) | 0.207357648 | -0.05563069 | 0.513574986 | -0.05644283 | 0.088360644 |

观察：因子有效阶段，big与small return基本同向，且多数情况下黄线（small return）在灰线之上，即SMB正收益

推断：size因子有效阶段，SMALL与BIG走势相同（所谓走势即同向，同正同负），存在极少数两端现象（即点出现在二、四象限）。在同向时，Small return>Big return，带来SMB正收益，即有效阶段SMB组合净值向上

BIG与SMB无线性关系，SMALL似乎有一些

ZZ500相对HS300与SMALL更拟合，即ZZ500可以代表小市值组合

HS300相对ZZ500更拟合大市值组合，符合逻辑

思考：若大市值组合可以被HS300代替，小市值组合可以被ZZ500代替，那么构建SMB多空组合是不是可以直接用这两个指数代替，毕竟HS300可被做空，而做空整个big组合可操作性降低（答：看SMB所在阶段，有效阶段两个指数分别具有大小市值组合的代表性，但其他（反转）阶段未必！！！）

观察：SMB相对HS300作为市场benchmark，被有效market neutralized（R^2极低），这是因为大市值组合和HS300高拟合，SMB（Rs-Rb）相当于Rs-Rmarket HS300.

2013年1月至2016年12月

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | small return | big return | SMB return | HS300 | ZZ500 |
| std | 0.129718616 | 0.077123882 | 0.105992101 | 0.081581211 | 0.090459039 |
| mean | 0.052852633 | 0.009529767 | 0.043322866 | 0.008920661 | 0.01774931 |
| sharp\_ratio | 0.407440616 | 0.123564416 | 0.408736737 | 0.109346999 | 0.196213776 |

此有效阶段small 与SMB 的拟合度相较上一个有效阶段提高明显（思考能提炼出什么）

SMB与big 依然无线性关系

SMALL依然被ZZ500更好的代表

BIG依然被HS300更好的代表

1. 反转阶段（净值阶段性下跌）

2016年12月至2018年2月

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | small return | big return | SMB return | HS300 | ZZ500 |
| std | 0.060738344 | 0.027870695 | 0.059750151 | 0.02993703 | 0.034345603 |
| mean | -0.023050413 | 0.015035256 | -0.038085669 | 0.014041341 | -0.003871134 |
| sharp\_ratio | -0.379503484 | 0.539464691 | -0.637415449 | 0.469029191 | -0.112711207 |

IMPORTANT:

1. 推断：size因子反转阶段，SMALL与BIG走势不再高度一致（两端情况发生频繁），存在较多数两端现象（主要出现在第二象限）。
2. 出现较多异常情况，即rb>0>rs， 导致rs-rb更加负（跌的更厉害），涨的时候很少出现利好SMB的异常。考虑到BIG与HS300走势接近，SMB反转由small组合return异常带来（背离市场走向）。而不是同向rs<rb带来（看过美股反转时，主要时同向且Rs<Rb带来）
3. 推断：因子反转的原因是small return异于big（market）第二象限出现大量远离原点的显著异常（称第二四象限出现远离原点的点为有效异象，离远点较近的点不算异象）。而有效时并small return并没有异于market方向

反转时，依然big于SMB无线性关系，SMALL与SMB的线性拟合度相对有效阶段进一步提高（有挖掘点）

反转时，ZZ500和SMALL的拟合度下降明显，ZZ500不再能够代表SMALL，也即反转时，不适合用ZZ500替代SMALL做多

BIG与HS300的拟合度在反转阶段依然保持高拟合度

反转阶段，相对HS300市场指数 SMB市场中性化表现依然很好（BIG与HS300依然高拟合度）

1. 无效阶段（净值在一定水平波动）

2005年2月-2007年11月

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | small return | big return | SMB return | HS300 | ZZ500 |
| std | 0.128639571 | 0.09379391 | 0.098070511 | 0.09462669 | 0.15887732 |
| mean | 0.04988192 | 0.052437693 | -0.004138444 | 0.057949431 | 0.078410169 |
| sharp\_ratio | 0.387764977 | 0.55907354 | -0.042198656 | 0.61240048 | 0.493526511 |

同跌时，SMALL跌更多

BIG与 SMB依然无线性关系

SMALL与ZZ500拟合度优秀

BIG与HS300拟合度优秀

SMB相对HS300市场中性，但对ZZ500仍有较弱线性关系

2010年10月2013年1月

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | small return | big return | SMB return | HS300 | ZZ500 |
| std | 0.086594868 | 0.056538455 | 0.062517224 | 0.067930238 | 0.076432064 |
| mean | -0.002399344 | -0.00104253 | -0.001356811 | -0.001009848 | -0.007564154 |
| sharp\_ratio | -0.02770769 | -0.01843936 | -0.021702997 | -0.014865951 | -0.098965718 |

1. 推断：size因子无效 阶段，SMALL与BIG走势相同（所谓走势即同向，同正同负），存在较有效期07-10更多的两端现象（即点出现在二、四象限）。
2. 同正收益时，small return更多时候更高，同负收益时，big return更多时候少跌

BIG与SMB 在无效阶段依然无线性关系

SMALL与ZZ500保持‘良好拟合度’的线性关系

BIG与HS300保持优秀拟合度的线性关系

SMB较HS300市场中性化程度高，较ZZ500市场中心化程度低

2018年2月至2020年5月

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | small return | big return | SMB return | HS300 | ZZ500 |
| std | 0.070540773 | 0.048125318 | 0.053947435 | 0.053041794 | 0.063109789 |
| mean | 0.00495485 | -0.00176045 | 0.006715296 | -0.002249534 | -0.002966663 |
| sharp\_ratio | 0.070240932 | -0.03658046 | 0.124478505 | -0.042410586 | -0.047007972 |

因子无效阶段，big与small return基本同向，且同涨时，BIG SMALL正收益率相近，偶尔big更高；同跌时，small跌更多

同正收益时，small return更多时候更高，同负收益时，big return更多时候少跌

无效阶段似乎第二象限两极化的点增多，但同向的点并没有像反转阶段那么少（反转阶段第二象限点主导走势）

BIG与SMB无线性关系

SMALL与ZZ500的拟合度不算优秀

BIG与HS300依然高拟合

SMB与ZZ500也不拟合，较好的市场中性化，可能是SMALL与ZZ500拟合度下降的原因？