

原创干货 | Fama-French 五因子模型在股票型基金业绩归因中的应用

前言：

早在 1993 年，Fama 和 French 两个人就已经发表了他们的三因子模型，认为股票的超额收益可以由市场风险、市值风险、价值风险来共同解释。关于这部分的实证应用，请点击链接

([原创干货 / Fama-French 三因子模型在股票型基金业绩归因中的应用](#)) 前往回顾。

Fama-French 三因子模型在解释股票型基金业绩归因中具有一定的学术地位，但在实际应用上，由于因子并未被完全提纯剥离，无法将所有投资绩效完全分析解释。后来，这两个人发现除了上述风险，还有盈利风险、投资风险（注意这里的投资风险不是二级市场的投资风险，而可以通俗地理解为企业扩大再生产的能力）也能带来个股的超额收益。

今天私募云通就带领大家简单了解这五个因子，并介绍纯因子组合的构造方法细节（细到不能再细，说明书级别），再展示一只股票型基金根据因子组合的业绩归因实证结果。



五因子火锅，既美味又营养

一. Fama-French 五因子模型

Ross(1976)提出套利定价理论（APT）模型，认为证券的收益不仅受市场组合变化的影响，也受市场中更多共同因素的影响。在讲解五因子模型之前，免不了回顾三因子模型：

$$r_{it} - r_{ft} = \alpha_i + \beta_{im}(r_{mt} - r_{ft}) + \beta_{is}SMB_t + \beta_{ih}HML_t + \varepsilon_{it}$$

变量	名称	算法	实务意义
r_{it}	投资组合 <i>i</i> 的收益率	由复权累计净值计算	投资组合 <i>i</i> 在统计区间的收益率
r_{ft}	t时刻的无风险收益率	可以用一年期国债利率等	
α_i	投资组合 <i>i</i> 的超额收益	通过多元回归计算	投资经理的择股能力
β_{im}	市场风险敞口	通过多元回归计算	投资组合对市场风险的敏感性
r_{mt}	市场组合的收益率	可由沪深 300 指数收益率等表示	
SMB_t	规模因子	上市公司规模不同导致的收益率差异	
β_{is}	规模风险暴露	通过多元回归计算	投资组合对不同规模公司股票的敞口
HML_t	账面市值比因子	由于上市公司账面市值比不同导致的收益率差异	
β_{ih}	账面市值风险暴露	通过多元回归计算	投资组合对不同账面市值比公司股票的敞口
ε_{it}	残差项		非系统性风险

根据 FF（1993）论文，如果三因子模型中的三个因子可以完全解释各种风险带来的超额收益，那么任何一个股票或者投资组合的 alpha 真实值应该为 0。然而不论是过去 20 多年的实证还是我们上篇文章的分析结果，都可以看到有些股票或基金的 α 值显著不为零。

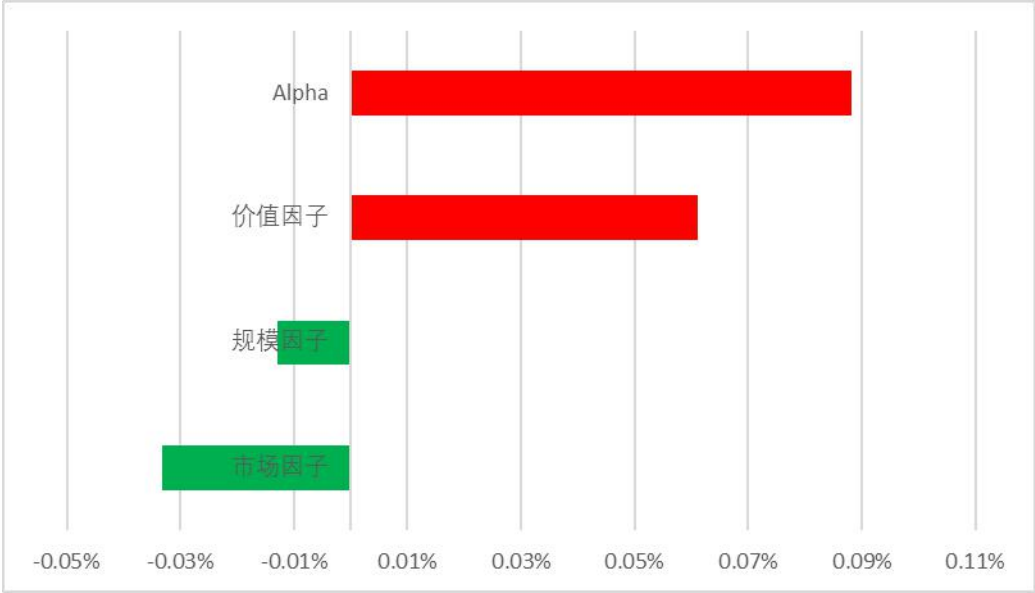


图 1. 三因子模型超额收益贡献分解示例

在示例图 1 中，超额收益贡献分解后，最大的收益贡献来源于 α ，虽然从拟合优度的角度三因子模型解释了 60%以上的超额收益，但仍不能解释所有的超额收益。于是，FF（2014）继续挖掘阿尔法，他们发现在上述风险之外，还有两个风险因子也能带来个股的超额收益，其表达式为：

$$r_{it} - r_{ft} = \alpha_i + \beta_{im}(r_{mt} - r_{ft}) + \beta_{is}SMB_t + \beta_{ih}HML_t + \beta_{ir}RMW_t + \beta_{ic}CMA_t + \varepsilon_{it}$$

变量	名称	算法	实务意义
RMW_t	盈利因子	上市公司盈利水平不同导致的收益率差异	
β_{ir}	盈利风险暴露	通过多元回归计算	投资组合对不同盈利水平公司股票的敞口
CMA_t	投资因子	由于上市公司投资再生产能力不同导致的收益率差异	
β_{ic}	投资风险暴露	通过多元回归计算	投资组合对不同投资水平公司股票的敞口
其他		调整算法与三因子时略有不同	同三因子模型

从表达式看，五因子模型比三因子模型多出来了两项： RMW_t 是（高-低）盈利股票投资组合的因子收益率，而 CMA_t 是（低-高）再投资比例公司股票投资组合的回报之差。这两项分别描述了盈利水平风险、投资水平风险。与三因子类似，参数估计的方法仍然是用多元线性回归的方法，这里的 α_i 则是五因子模型里尚未解释的超额收益。

另外值得注意的是，FF（2014）提到，模型的表现相对于因子的具体取值并不敏感，举个例子：不论用流通市值还是总资产用以描述市值（规模）因子，对模型的影响并不会太大，主要原因在于对该维度下的股票分组影响不大，因此如果在其他文献中看到的定义与本文有所不同，也无需大惊小怪。

二. 组合因子构造

以沪深三百成分股为例。为稳定样本，采用 2014.01-2017.01 期间周度数据。

因子对应指标（代理变量）：

- 市场因子指标：沪深 300 指数；
- 规模因子指标：总市值；
- 价值因子指标：账面市值比，即市净率的倒数；
- 盈利因子指标：扣除非经常性损益、摊薄的 ROE；
- 投资因子指标：当期总资产相对上期总资产变化率；
- 无风险利率、股票型基金收益率用私募云通 CHFDB 数据库直接调取；
- 以自然周为频率，计算因子值并重新分层并分类计算对应因子收益率。

这里的分层和分类是什么意思呢？在五因子模型中的因子收益率，实际上是沿着某一个或几个因子维度，将股票分层，算出不同股票组合的收益率，再通过某种运算规则计算得到。具体来说，在每个自然周的周末，我们按因子大小将所有股票分成 2~3 层，计算下一周每层内部股票的平均收益率，之后用不同层股票的平均收益率做差，层间的差异实质上衡量了该

种因子的收益效应。通过这个过程，我们完成了因子收益率的计算，且使得各个因子的量纲得到统一，不同因子收益率之间也具有可比性。

1. 对规模数据分类

Fama 和 French 按照规模的中位数将股票划分为小规模(S)和大规模(B)两类。

2. 对价值、盈利、投资数据分层分类

在价值数据层面，将 1 中样本按照账面市值比从高到低排序后分为三组，前 30%为高账面市值比(H)，后 30%为低账面市值比(L)，剩余 40%为中等账面市值比(M)。

在盈利数据层面，将 1 中样本按照摊薄 ROE 从高到低排序后分为三组，前 30%为高盈利 (强)水平(R)，后 30%为低盈利 (弱)水平(W)，剩余 40%为中盈利 (中)水平(M)。

在投资数据层面，将 1 中样本按照资产变化率从高到低排序后分为三组，前 30%为高投资 (强)水平(A)，后 30%为低投资 (保守)水平(C)，剩余 40%为中投资 (中)水平(M)。

3. 组合交叉

组合交叉得到 18 个组合，即第一层：(S, L), (S, M), (S, H), (B, L), (B, M), (B, H)；第二层：(S, W), (S, M), (S, R), (B, W), (B, M), (B, R)；第三层：(S, C), (S, M), (S, A), (B, C), (B, M), (B, A)。

4. 规模因子 SMB

三层分组均有规模数据，因此规模因子最为特殊，是三个层面的平均。在第一层上：

$$SMB_t = (r_{(S,L)} + r_{(S,M)} + r_{(S,H)})/3 - (r_{(B,L)} + r_{(B,M)} + r_{(B,H)})/3$$

同样我们可以计算第二层、第三层的 SMB₂、SMB₃，平均求得最终的规模因子。

5. 价值因子 HML、盈利 RMW、投资因子 CMA

HML 是高账面市值比公司的平均收益率与低账面市值比公司平均收益率的差额，表达形式如下：

$$HML_t = (r_{(S,H)} + r_{(B,H)})/2 - (r_{(S,L)} + r_{(B,L)})/2$$

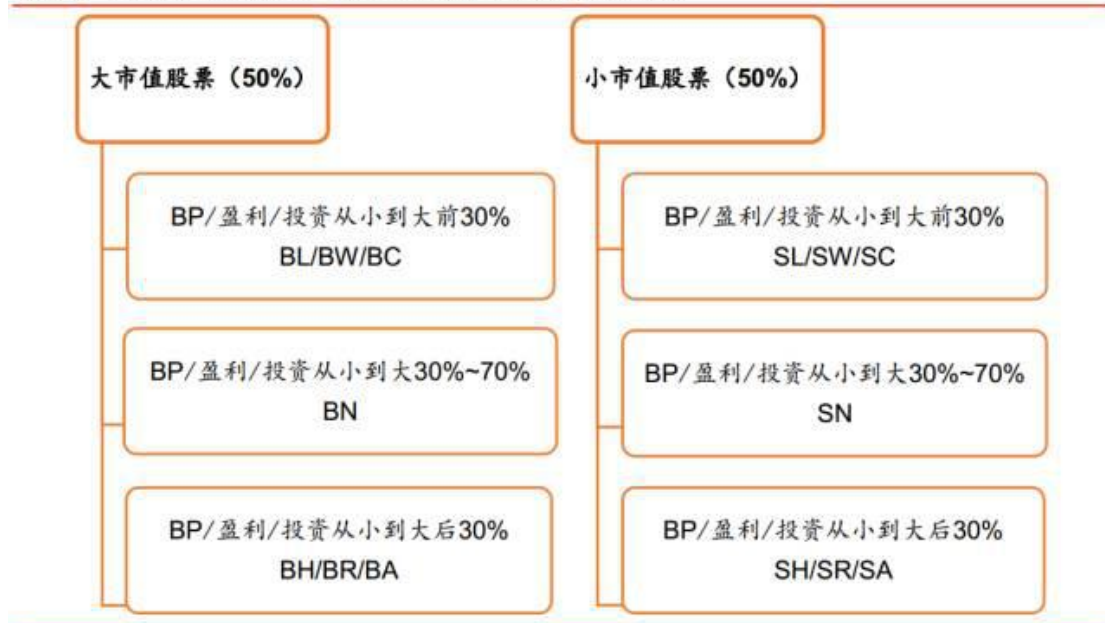
盈利、投资因子完全类似。借用华泰证券研究所资料来总结以上步骤 4、5 可得：

图表2： 因子收益率计算表

分组方式	分组计算方式	因子构造
规模：按中位数均分为两组，分别记为大(B)或小(S)：	规模与 BP 两维度分组	$SMB_{BP}=(SH+SN+SL)/3-(BH+BN+BL)/3$
BP、盈利、投资：按 30%, 70% 分位数分成三组，分别记为	规模与盈利两维度分组	$SMB_{op}=(SR+SN+SW)/3-(BR+BN+BW)/3$
BP——高(H)、中(N)、低(L)	规模与投资两维度分组	$SMB_{inv}=(SC+SN+SA)/3-(BC+BN+BA)/3$
盈利——强(R)、中(N)、弱(W)	以上三种分组方式的均值	$SMB=(SMB_{BP}+SMB_{op}+SMB_{inv})/3$
投资——保守(C)、中(N)、强(A)	规模与 BP 两维度分组	$HML=(BH+SH)/2-(BL+SL)/2$
	规模与盈利两维度分组	$RMW=(BR+SR)/2-(BW+SW)/2$
	规模与投资两维度分组	$CMA=(BC+SC)/2-(BA+SA)/2$

资料来源：Wind，华泰证券研究所

图表3： 因子构造 2x3 分层表（规模维度与 BP 或盈利或投资维度）



资料来源：Wind，华泰证券研究所

三. 小结

看到这里就该展示一只股票型基金根据因子组合的业绩归因实证结果了，正如我们在（[原创干货 / Fama-French 三因子模型在股票型基金业绩归因中的应用](#)）中所展示的图表一样，但是如果按部就班直接放图，对于各位看官来说就毫无参与感了，所以云通在这留一个**课后作业**，请大家按照以上思路自己做一遍因子收益率的构造以及基金业绩归因，要求：

1. 五因子时间序列统计特性（均值、标准差、t 值）
2. 五个纯因子组合累计净值曲线
3. 五因子之间的相关性+与基金超额收益之间的相关性
4. 基金超额收益贡献分解（平均结果+变化雷达图）