



中央财经大学 金融学院

School of Finance, Central University of Finance and Economics

# Fama MacBeth回归模型和组合分析

朱一峰

中央财经大学金融学院

2024年1月11日

# 前言

- 横截面资产定价研究中主要使用Fama MacBeth回归和组合分析两种研究方法来做实证研究。
- 为什么需要用两种方法？
- Fama MacBeth回归可以加入更多的控制变量，是一种线性回归。
- 组合分析法最大的优点可能在于它是一种非参数技术分析的方法。这意味着我们不需要对需要检验的变量在横截面上的关系提出假设。许多其他方法依赖于对被检变量之间关系的函数形式的假设。例如，线性回归分析假设因变量和自变量之间的关系是线性的。组合分析不需要此假设。



# 提纲

- 第一节：Fama MacBeth回归分析
- 第二节：单变量组合分析
- 第三节：双变量独立排序组合分析
- 第四节：双变量序贯排序组合分析



# 第一节：Fama MacBeth回归分析



# 一、 Fama MacBeth回归分析

- 展示的回归系数是横截面回归的时间序列均值，t值是得到的回归系数时间序列的t检验值。
- 回归的自变量中包含了因子载荷（Beta）和资产特征。



**Table 4**

5% ES and expected returns for different sentiment regimes. The table reports the time-series averages of the slope coefficients and their  $t$ -values from the Fama–MacBeth regressions of the excess stock returns on various pricing variables (see the first column) using monthly data from July 1965 to December 2016 in high- and low-sentiment periods.

$$ExR_{i,t+1} = b_{0,t} + b_{1,t}ES_{i,t} + b_{2,t}X_{i,t} + \varepsilon_{i,t+1},$$

where  $ExR_{i,t+1}$  is the excess return, which is the difference between the monthly stock return on stock  $i$  and the one-month T-bill rate at time  $t+1$ .  $ES_{i,t}$  is ES at 5% for stock  $i$  at time  $t$ .  $X_{i,t}$  is a set of control variables.

Panel A: 5% ES	High-sentiment periods (the BW sentiment index is larger than its mean)				
	(1)	(2)	(3)	(4)	(5)
5% ES	-0.2878*** (-5.30)	-0.1479*** (-2.93)	-0.2835*** (-9.57)	-0.2419*** (-8.54)	-0.2489*** (-8.84)
SIZE			-0.2591*** (-5.23)	-0.2712*** (-5.56)	-0.2726*** (-5.60)
BM			0.3396*** (5.12)	0.3355*** (5.16)	0.3329*** (5.10)
MOM			0.0098*** (5.23)	0.0100*** (5.38)	0.0099*** (5.35)
TURN			0.0378 (0.88)	0.0607 (1.42)	0.0530 (1.25)
ILLIQ			0.0412* (1.78)	0.0541** (2.31)	0.0500** (2.17)
BETA			0.9121*** (3.58)	0.8352*** (3.13)	0.7994*** (2.99)
COSKEW				-0.6647* (-1.71)	-0.6539* (-1.69)
MAX			-0.0983*** (-12.30)	-0.0490*** (-3.83)	-0.0602*** (-4.97)
VOL		-0.2864*** (-8.74)		-0.2228*** (-4.61)	
IVOL					-0.1720*** (-4.08)
Constant	1.9580*** (7.60)	1.9510*** (7.64)	3.2993*** (8.71)	3.2637*** (8.73)	3.2723*** (8.76)
R-squared	0.034	0.037	0.086	0.094	0.093



## 第二节：单变量组合分析



## 二、单变量组合分析

- 我们从最基本的组合分析类型开始：单变量组合分析。单变量组合分析中只有一个排序变量 $X$ 。分析的目的是评估排序变量 $X$ 和结果变量 $Y$ （通常是资产收益率）之间的横截面关系。单变量组合分析不允许我们在检验二者关系时控制任何其他影响因素。
- 单变量组合分析共有四个步骤：第一步，计算将样本进行分组的断点值；第二步，使用这些断点对样本进行分组；第三步，计算每个时间段 $t$ 内，每个组合结果变量 $Y$ 的平均值；第四步，检验不同组合中 $Y$ 的平均值（不同的加权方式，等权重、流通市值、总市值）的变化情况。





## 二、单变量组合分析

- 量化投资使用的就是单变量组合分析。
- 通过做多收益率高的极端组，做空收益率低的极端组，每一时期的量化投资的收益就是收益率高的极端组平均收益减去收益率低的极端组平均收益。
- 当然我们也需要检验这样的收益是否会被常见的因子模型解释，如果被解释，说明发现的新策略信息已经包含在已有的公共因子中，如果没有，说明确实发现了新的量化投资策略。



# 单变量组合分析

**Table 3**

Decile portfolios for 5% ES. The table reports the equal-weighted averages of monthly stock returns in Column (1). The last seven columns show the alphas of the portfolios with respect to seven different factor models: (i) The FFCPS 5-factor alpha based on the following factors: market (MKT), size (SMB), book-to-market (HML), momentum (MOM), and liquidity risk (LIQ) factors of Fama and French (1993), Carhart (1997), and Pastor and Stambaugh (2003); (ii) The Fama–French 5-factor (FF5) alpha based on the five factors proposed in Fama and French (2015): market (MKT), size (SMB), book-to-market (HML), investment (CMA), and profitability (RMW); (iii) The Q-4 factor alpha based on Hou et al. (2015) using market (MKT), size (SMB), investment ( $R_{I/A}$ ), and profitability ( $R_{ROE}$ ); (iv) The FMAX-6 factor alpha based on the five factors inside the FFCPS-5 factor model together with the lottery demand factor (FMAX) proposed by Bali et al. (2017); (v) The M-4 factor alpha based on the mispricing factors MGMT and PEPF besides MKT and SMB following Stambaugh and Yuan (2017); (vi) The BF-3 factor alpha based on the long- and short-run behavioral factors FIN and PEAD besides MKT following Daniel et al. (2019); and (vii) The T-4 factor alpha based on the tail-risk factor proposed by Kelly and Jiang (2014) besides the three Fama–French factors: MKT, SMB, and HML, as well as their t-values, for decile portfolios sorted by the 5% ES in the previous month based on data  $t$  ( $t + 1$ ) from June (July) 1972 to November (December) 2016 in high- and low-sentiment periods.

Panel A: 5% ES		High-sentiment periods (the BW sentiment index is higher than its mean)						
Portfolio	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Excess return (%)	FFCPS alpha (%)	FF5 alpha (%)	Q-4 alpha (%)	FMAX-6 alpha (%)	M-4 alpha (%)	BF-3 alpha (%)	T-4 alpha (%)
1(lowest)	0.941*** (4.70)	0.240** (2.36)	0.098 (1.07)	0.187* (1.68)	0.003 (0.04)	0.071 (0.61)	0.250** (2.23)	0.285*** (2.90)
2	0.997*** (4.05)	0.259** (2.48)	−0.005 (−0.06)	0.125 (1.15)	−0.020 (−0.28)	0.011 (0.10)	0.282** (2.42)	0.260*** (2.63)
3	0.854*** (3.17)	0.113 (1.08)	−0.202** (−2.58)	−0.051 (−0.45)	−0.148* (−1.87)	−0.123 (−1.05)	0.242* (1.83)	0.069 (0.70)
4	0.968*** (3.31)	0.258** (2.34)	−0.096 (−1.12)	0.092 (0.77)	−0.000 (−0.00)	0.027 (0.22)	0.465*** (3.01)	0.184* (1.73)
5	0.806** (2.57)	0.111 (1.02)	−0.222** (−2.55)	−0.030 (−0.25)	−0.109 (−1.17)	−0.081 (−0.68)	0.439** (2.58)	0.029 (0.28)
6	0.752** (2.24)	0.156 (1.46)	−0.130 (−1.32)	0.055 (0.45)	−0.018 (−0.18)	0.042 (0.35)	0.482*** (2.68)	0.035 (0.33)
7	0.512 (1.37)	0.048 (0.49)	−0.201* (−1.94)	−0.002 (−0.02)	−0.050 (−0.51)	−0.002 (−0.02)	0.524*** (2.66)	−0.117 (−1.13)
8	0.151 (0.35)	−0.185* (−1.79)	−0.324*** (−2.84)	−0.177 (−1.54)	−0.143 (−1.35)	−0.067 (−0.55)	0.425** (2.03)	−0.368*** (−3.31)
9	−0.321 (−0.68)	−0.456*** (−3.58)	−0.481*** (−3.50)	−0.245* (−1.90)	−0.288** (−2.35)	−0.149 (−1.07)	0.310 (1.36)	−0.725*** (−4.99)
10(highest)	−1.259** (−2.38)	−1.198*** (−6.77)	−1.143*** (−6.36)	−0.921*** (−5.32)	−0.886*** (−5.54)	−0.659*** (−3.64)	−0.287 (−1.11)	−1.557*** (−7.85)
10-1 spread	−2.199*** (−4.76)	−1.438*** (−6.05)	−1.242*** (−5.48)	−1.108*** (−4.84)	−0.888*** (−4.69)	−0.729*** (−2.98)	−0.537* (−1.96)	−1.842*** (−7.17)



# 第三节：双变量独立排序组合分析



### 三、双变量独立排序组合分析

- 在双变量组合分析中，排序程序通常有两种不同的类型：独立排序和序贯排序。本节，我们将讨论独立排序组合分析。
- 双变量独立排序组合分析旨在评估两个排序变量（我们称为 $X_1$  和 $X_2$ ）与结果变量 $Y$ （通常是资产收益率）在横截面上的相关关系。
- 在双变量独立排序的组合分析中，组合是通过独立地对两个变量进行排序而形成的。



表 5.14

双变量独立排序组合的超额收益和异常收益

	系数	$\beta 1$	$\beta 2$	$\beta 3$	$\beta \text{ Diff}$	$\beta \text{ Avg}$
<i>MktCap</i> 1	超额收益	20.96 (6.12)	23.68 (6.81)	21.20 (3.62)	0.24 (0.06)	21.95 (5.66)
	FFC $\alpha$	11.96 (5.05)	18.63 (5.46)	9.90 (1.31)	-2.06 (-0.38)	13.50 (3.44)
<i>MktCap</i> 2	超额收益	11.07 (4.13)	13.45 (4.03)	11.49 (4.98)	0.41 (0.13)	12.00 (5.20)
	FFC $\alpha$	0.18 (0.08)	1.42 (0.45)	0.92 (0.43)	0.74 (0.16)	0.84 (0.65)
<i>MktCap</i> 3	超额收益	8.86 (3.01)	10.48 (5.50)	8.51 (5.18)	-0.36 (-0.11)	9.28 (5.52)
	FFC $\alpha$	-0.27 (-0.09)	1.33 (1.21)	-1.14 (-0.88)	-0.87 (-0.20)	-0.02 (-0.03)
<i>MktCap</i> 4	超额收益	6.45 (3.85)	8.99 (6.74)	8.64 (4.00)	2.19 (1.17)	8.03 (5.37)
	FFC $\alpha$	-0.91 (-0.81)	1.78 (3.23)	-1.41 (-0.94)	-0.50 (-0.24)	-0.18 (-0.25)
<i>MktCap</i> Diff	超额收益	-14.51 (-4.81)	-14.69 (-4.37)	-12.55 (-2.16)	1.95 (0.60)	-13.92 (-3.67)
	FFC $\alpha$	-12.87 (-3.83)	-16.85 (-5.23)	-11.31 (-1.57)	1.56 (0.43)	-13.68 (-3.29)
<i>MktCap</i> Avg	超额收益	11.84 (4.90)	14.15 (6.59)	12.46 (5.14)	0.62 (0.25)	12.82 (6.29)
	FFC $\alpha$	2.74 (1.72)	5.79 (5.12)	2.07 (0.84)	-0.67 (-0.19)	3.53 (3.00)

该表展示了根据  $\beta$  分类成 3 个投资组和根据 *MktCap* 分类成 4 个投资组所形成的组合的平均超额收益（标记为超额收益的行）和 FFC 阿尔法（标记为 FFC  $\alpha$  的行）。括号中的数字是  $T$  统计量，并根据 Newey 和 West（1987）六期滞后进行了调整，以检验组合时间序列平均超额收益或 FFC  $\alpha$  为 0 的零假设。



## 第四节：双变量序贯排序组合分析



## 四、双变量序贯排序组合分析

- 双变量序贯排序组合分析与其对应的独立排序类似，都是通过基于两个排序变量 $X_1$ 和 $X_2$ 的值对样本中的个体进行排序而形成的。
- 序贯排序和独立排序分析的唯一区别是，序贯排序在分析中，第二个排序变量的断点是根据第一个排序变量排序形成的每个组中形成。因此，当目标是分析在给定 $X_1$ 之后， $X_2$ 和 $Y$ 之间的关系时，可以使用序贯排序分析。在双变量序贯排序分析中，我们不再检验 $X_1$ 和 $Y$ 之间的关系， $X_1$ 仅用作控制变量。





## 四、双变量序贯排序组合分析

Table 5

Portfolios sorted by *IVOL* and *VaR1* during high consumer confidence period.

Panel A: Equal-weighted (%)								
	Excess Return				FF5 Alpha			
<i>VaR1</i>	<i>V1</i>	<i>V3</i>	<i>V5</i>	<i>V5-V1</i>	<i>V1</i>	<i>V3</i>	<i>V5</i>	<i>V5-V1</i>
<i>IVOL1</i>	1.378** (2.02)	1.362* (1.69)	1.178 (1.42)	-0.200 (-0.57)	0.367 (1.54)	0.186 (0.78)	0.051 (0.22)	-0.317 (-1.03)
<i>IVOL2</i>	1.363* (1.91)	1.492* (1.83)	1.183 (1.36)	-0.180 (-0.47)	0.212 (1.02)	0.352** (2.12)	0.105 (0.45)	-0.107 (-0.33)
<i>IVOL3</i>	1.542** (2.08)	1.024 (1.23)	0.938 (1.07)	-0.605 (-1.47)	0.469** (2.03)	0.108 (0.52)	-0.169 (-0.74)	-0.638* (-1.72)
<i>IVOL4</i>	1.458* (1.91)	1.233 (1.45)	0.504 (0.53)	-0.954** (-2.07)	0.451* (1.96)	0.167 (0.47)	-0.485* (-1.81)	-0.936** (-2.58)
<i>IVOL5</i>	0.687 (0.88)	0.037 (0.04)	-0.669 (-0.73)	-1.356*** (-3.16)	-0.330 (-1.23)	-0.705** (-2.60)	-1.436*** (-5.00)	-1.106*** (-2.98)
Avg ( <i>IV1-IV5</i> )	1.286* (1.79)	1.030 (1.27)	0.627 (0.72)	-0.659* (-1.90)	0.234 (1.52)	0.022 (0.14)	-0.387** (-2.19)	-0.621** (-2.40)
Panel B: Value-weighted (%)								
	Excess Return				FF5 Alpha			
<i>VaR1</i>	<i>V1</i>	<i>V3</i>	<i>V5</i>	<i>V5-V1</i>	<i>V1</i>	<i>V3</i>	<i>V5</i>	<i>V5-V1</i>
<i>IVOL1</i>	1.442** (2.04)	1.112 (1.39)	0.774 (0.96)	-0.668 (-1.45)	0.340 (0.95)	-0.026 (-0.09)	-0.340 (-1.22)	-0.681 (-1.60)
<i>IVOL2</i>	1.269* (1.73)	1.455* (1.75)	0.702 (0.82)	-0.567 (-1.07)	0.044 (0.13)	0.428 (1.56)	-0.241 (-0.90)	-0.285 (-0.62)
<i>IVOL3</i>	1.851** (2.49)	0.872 (1.11)	0.552 (0.65)	-1.299** (-2.00)	0.940*** (2.63)	0.154 (0.59)	-0.428 (-1.44)	-1.368** (-2.55)
<i>IVOL4</i>	1.364* (1.74)	1.238 (1.51)	0.203 (0.22)	-1.162* (-1.91)	0.313 (0.91)	0.263 (0.64)	-0.638** (-2.09)	-0.951** (-2.08)
<i>IVOL5</i>	0.829 (1.08)	-0.008 (-0.01)	-0.654 (-0.73)	-1.483** (-2.56)	-0.207 (-0.62)	-0.607* (-1.72)	-1.206*** (-3.51)	-0.999** (-2.17)
Avg ( <i>IV1-IV5</i> )	1.351* (1.95)	0.934 (1.19)	0.315 (0.37)	-1.036** (-2.29)	0.286 (1.62)	0.042 (0.25)	-0.571*** (-2.92)	-0.857*** (-2.99)

This table presents both equal- and value-weighted averages of monthly stock returns and their *t*-values (given in parentheses) for quintile portfolios sorted by *IVOL* and then by *VaR1* from January 1997 to December 2019 during high consumer confidence period. At the beginning of each month, we first sort stocks by *IVOL* into five quintiles. Within each *IVOL* portfolio, we then sort stocks into five sub-quintiles by *VaR1*. *IVOL1* and *IVOL5* denote the lowest and highest quintiles for *IVOL*, respectively. *V1* and *V5* denote the lowest and highest quintiles for *VaR1*. The *V5-V1* spread denotes the average equal-weighted (Panel A; value-weighted for Panel B) excess return difference between the highest and lowest *VaR1* within the same *IVOL* quintile or the average of five *IVOL* portfolios (Avg (*IV1-IV5*)). The table also reports its [Fama and French \(2015\)](#) 5-factor (FF5) alpha. Avg (*IV1-IV5*) reports the return average or return spread average of the five *IVOL* quintiles. Significance at the 1%, 5%, and 10% levels is indicated by \*\*\*, \*\*, and \*, respectively.





## 四、双变量序贯排序组合分析

- 双变量独立排序和序贯排序组合分析会产生类似的结果。虽然通常如此，但不一定总是这样。因此，使用这两种方法同时进行分析是一个好主意。如果独立排序和序贯排序分析的结果产生完全不同的结论，就有必要做进一步调查。
- 两种类型的双变量组合分析方法之间最显著的差异在于，序贯排序分析只能用于在控制 $X_1$ 之后检查第二个排序变量 $X_2$ 和结果变量 $Y$ 之间的关系，而独立排序分析则允许检查这种关系以及控制 $X_2$ 之后 $X_1$ 和 $Y$ 之间的关系。

