

基于中国 A 股数据的 资本资产定价模型实证分析

摘 要

本文使用中国 A 股市场 2000 年到 2019 年的交易数据，对资本资产定价模型进行实证分析。首先，采用 Fama-MacBetch 回归对标准 CAPM 进行检验，尽管结果成功验证了三个核心假设，但是得到的 β 风险对收益率影响并不显著。其次，又分析了 Fama-French 三因子模型，该模型综合使用市场因子、规模因子和价值因子进行资本资产定价，结果显示其能够捕捉到 A 股市场收益率的绝大部分变动，对所构建 25 个投资组合回归得到调整后 R^2 均大于 0.88。最后，本文考虑了中国股市特有的 IPO 监管造成的“壳价值污染”问题，通过剔除市值最低 30% 部分的股票来尽量削减其影响，这使得三因子模型中的部分异常结果得到有效修正。虽然本文并未提出较为创新的思路和方法，但是对核心模型的数据处理与回归分析过程做了细致呈现，这 1) 让笔者对资本资产定价模型有了更加深刻的理解；2) 为资本资产定价模型的基础学习提供了一套可以复现的开源资料¹。

关键词：中国 A 股市场数据 资本资产定价模型 Fama-French 三因子模型

¹ 本文所用配套代码已随论文提交，包含数据的整体工程文件也上传至 https://github.com/KarryRen/FinalPaper_2022_FinancialEconometrics_CAPM 完成公开，本文中的所有结果均可复现。

一、 资本资产定价模型概述

1.1 标准 CAPM

二十世纪六、七十年代, Sharpe(1964), Lintner(1965)和 Black(1972)将 Markowitz (1959)的理论延伸成为资本资产定价模型(Capital and Asset Pricing Model, CAPM), 标准 CAPM 用方程表示为:

$$E(\tilde{R}_i) = R_f + [E(\tilde{R}_m) - R_f]\beta_i \quad (1)$$

其中 $E(\tilde{R}_i)$ 是资产 i 的期望收益率; R_f 指无风险利率; $E(\tilde{R}_m)$ 为市场组合的期望收益率, 它在理论上是指由所有的风险资产共同构成的投资组合; β_i 表示资产 i 的系统风险, 它被定义为资产收益与市场组合收益之间的协方差同市场组合收益方差之间的比:

$$\beta_i = \frac{cov(\tilde{R}_i, \tilde{R}_m)}{\sigma^2(\tilde{R}_m)}$$

CAPM 模型是在严格的假设条件下进行的理论分析模型, 假设包括:

1. **均值方差假设:** 投资者通过观察一段时间内证券组合的收益率和标准差来评判证券组合, 若标准差相同, 投资者将选择具有较高收益率的证券组合, 若预期收益率等同, 投资者将选择具有较低标准差的证券组合。

2. **投资者一致假设:** 市场中所有投资者计划的投资时点和投资期限相同, 组成各个投资组合的证券数目相同, 投资者对证券的收益率、标准差、协方差看法相同, 投资者选择不同投资组合的原因只是风险偏好不同。

3. **完全市场假设**：市场不存在交易成本和税收，所有资产完全可分割、可交易，市场是完全竞争的，信息成本为零，所有市场参与者同时接受信息，且都是理性的。

标准 CAPM 的主要结论是：任何资产的期望收益率与其 β 值均呈线性正相关，如果市场达到均衡，市场上的所有资产的风险收益定价关系都应在证券市场线上。[Black、Jensen 和 Scholes \(1972\)](#) 以及 [Fama 和 MacBetch \(1973\)](#) 均证实了股票的平均收益和其 β 间存在着显著的正向线性关系。

1.2 Fama-French 三因子模型

二十世纪后期，一些实证分析结果向标准 CAPM 发起挑战，其中最引人瞩目的就是 [Banz \(1981\)](#) 提出的“规模效应”，他发现股票市值 ME（股票的价格乘以流通股数目）能够提高 β 值对平均收益率的横截面解释能力，小股票（低 ME）的平均回报率在用 β 估计时偏高，而大股票则偏低。

[Fama 和 French \(1992a\)](#) 也像 [Ringanum \(1981\)](#) 以及 [Lakonishok 和 Shapiro \(1986\)](#) 一样，发现了 1963 至 1990 年期间，美国股票市场风险 β 和平均收益之间的线性正相关关系消失了。他们注意到：股票平均收益率与规模、杠杆、E/P（盈利收益率）以及账面市值比之间的单变量关系很强。在多变量测试中，规模和平均收益率之间的负相关关系是稳健的，账面市值比与平均收益率之间的正向关系也持续存在，进而认为如果资产定价是理性的，那么股票风险应该是多维度的，风险的一个维度是由规模来代表，另一个维度由账面市值比来代表。

实证结果表明规模（ME）和账面市值比（BE/ME）为 1963 至 1990 年期间的股票收益率提供了简单而有效的解释，并在此基础上 [Fama 和 French \(1993\)](#) 提出了“Fama-French 三因子模型”：

$$R(t) - RF(t) = a + b[RM(t) - RF(t)] + sSMB(t) + hHML(t) + e(t) \quad (2)$$

其中 t 是表示第 t 期， $R(t)$ 、 $RF(t)$ 和 $RM(t)$ 分别表示 t 期投资组合的收益率、市场的无风险利率以及市场组合的收益率。 $SMB(t)$ 、 $HML(t)$ 为在标准 CAPM (1) 的基础上增加的两个因子，其含义分别是“小市值组合收益率减大市值组合收益率”以及“高账面市值比组合收益率减低账面市值比组合收益率”。

该模型实证结果表明：对时间序列回归而言，规模和账面市值比这两个因子可以解释不同投资组合平均收益的差异。

本文将基于中国 A 股市场数据对标准 CAPM 和 Fama-French 三因子模型进行实证分析，后续章节按照如下步骤展开：[第二节](#)将说明本文的数据来源；[第三节](#)将介绍 Fama-MacBatch 回归方法以及使用该方法对标准 CAPM 进行检验的结果；[第四节](#)则紧接着构建了 Fama-French 三因子模型并展开分析；[第五节](#)参考“中国版三因子模型”通过剔除市值最低 30% 部分的股票尽量削减“壳价值污染”的影响，进而对三因子模型的部分异常做出解释；[第六节](#)总结了相关结论。

二、数据来源

本文使用的股票的交易数据和指数的交易数据均来自 RESSET 数据库，无风险利率采用一年期定期存款利率，数据频率为月度，时间是从 2000 年 1 月至 2019 年 12 月。

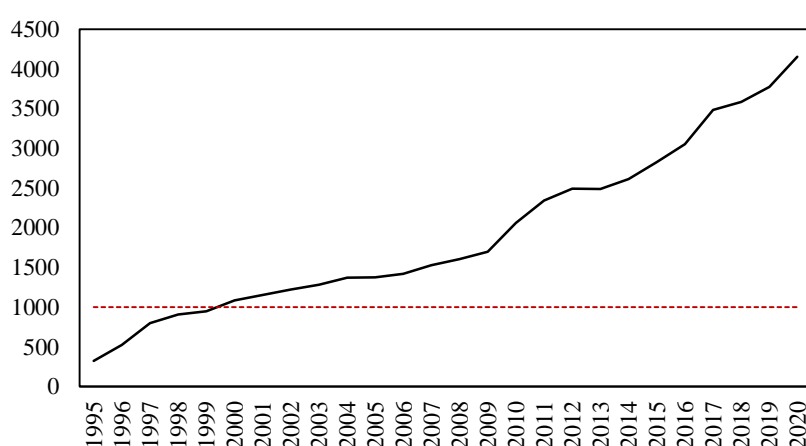


图 1. 1995 年至 2020 年中国上市公司数量

中国 A 股市场，始于 1990 年上交所和深交所成立。本文把分析重点放在 2000 年与 2019 年之间，主要有三个原因。首先，是为了**保证会计数据的统一性**：在中国，有关财务报告各方面的规则和条例实施直到 1999 年才基本成型。虽然 1993 年开始有了公平交易和财务披露原则，但企业在遵守这些原则方面得到的指导很少，各公司自由发挥，强加自己的标准，这都限制了各个公司会计数据的可比性。直到 1998 年和 1999 年关于交易和财务报告的法律和法规才得到更加彻底的设计和实现，1999 年会计准则的统一才得以广泛实现，因此 2000 年以后的投资组合收益样本所依赖的会计数据比前几年更具有可比性。其次，是为了**确保有足够多的观察数据**，如图 1. 所示直到 2000 年中国上市公司数量才突破 1000 家。最后，考虑到 2020 年以来新冠疫情对市场产生了巨大冲击，将 2020 年至今的数据进行剔除。

三、采用 Fama-MacBatch 回归验证标准 CAPM

3.1 假设

方程 (1) 所表示的标准 CAPM 有着 3 个需要被验证的基础条件。**C1**: 在任意投资组合 m 中, 预期收益与其风险之间的关系应是线性的; **C2**: 方程 (1) 中没有出现证券 i 的其他风险度量, 那么 β_i 应该是投资组合 m 中证券风险的完整度量; **C3**: 在一个由风险规避型投资者组成的市场中, 较高的风险应与较高的预期收益相关, 即 $E(\tilde{R}_m) - E(\tilde{R}_0) > 0$ 。

3.1.1 收益率的随机模型

方程 (1) 是通过预期收益的形式给出的, 必须用逐期的证券收益数据才能进行检验, 但是这些数据实际是不可得的。因此, 必须构造一个基于逐期收益形式的模型, 这样才能使用实际观察到的平均收益率来验证 **C1~C3**, 本文采用与 Fama 和 MacBetch (1973) 相同的方法对方程 (1) 做如下形式的随机概括:

$$\tilde{R}_{it} = \tilde{\gamma}_{0t} + \tilde{\gamma}_{1t}\beta_i + \tilde{\gamma}_{2t}\beta_i^2 + \tilde{\gamma}_{3t}s_i + \tilde{\eta}_{it} \quad (3)$$

其中, t 指的是第 t 期; \tilde{R}_{it} 是证券 i 从 $t-1$ 期到 t 期的单期百分比收益率。方程 (3) 允许 $\tilde{\gamma}_{0t}$ 和 $\tilde{\gamma}_{1t}$ 在不同期内随机变化。条件 **C3** 成立对应于方程 (1) 中斜率 $[E(\tilde{R}_{mt}) - E(\tilde{R}_{0t})]$ 是正的, 对应于 (3) 中风险溢价 $\tilde{\gamma}_{1t}$ 的期望是正的, 也即 $E(\tilde{\gamma}_{1t}) > 0$ 。

引入 β_i^2 是为了检验线性, 进而条件 **C1** 成立所对应的等价假设可以确定为 $E(\tilde{\gamma}_{2t}) = 0$ 。类似的, (3) 中引入的 s_i 是对证券 i 某种风险的衡量, 这种风险是未被

β_i 度量的, 那么条件 C2 所对应的假设应是 $E(\tilde{\gamma}_{3t}) = 0$ 。当然对比 (1) 和 (3) 后, 还可以确定一个显而易见需要被检验的假设 $E(\tilde{\gamma}_{0t}) = R_{ft}$ 。

3.1.2 模型的核心假设

总的来说, 方程 (3) 是对标准 CAPM (1) 的随机概括形式, 本文针对其给出标准 CAPM 如下四个可检验的核心假设:

$$H1 \text{ (线性): } E(\tilde{\gamma}_{2t}) = 0;$$

$$H2 \text{ (}\beta \text{ 是风险的完整度量): } E(\tilde{\gamma}_{3t}) = 0$$

$$H3 \text{ (正预期的收益-风险交易): } E(\tilde{\gamma}_{1t}) > 0$$

$$H4 \text{ (Sharpe-Lintner 假设): } E(\tilde{\gamma}_{0t}) = R_{ft}$$

当然, 本文在设计可检验假设时的确注意到 Fama 和 MacBetch (1973) 还提及了市场有效性假设, 但综合考虑到这一假设主要侧重于市场有效性并且实现起来较为复杂, 故本文并未涉及该假设。

3.2 检验方法

3.2.1 方法概述

检验标准 CAPM 不得不面对“变量误差”问题, 即方程 (1) 需要采用相对风险度量的真实值 β_i , 但在实际分析中只能使用估计值 $\hat{\beta}_i$ 。本文采用 $\hat{\beta}_i$ 的“定义式”对其进行估计:

$$\hat{\beta}_i = \frac{\widehat{cov}(\tilde{R}_i, \tilde{R}_m)}{\hat{\sigma}^2(\tilde{R}_m)}$$

其中 $\widehat{cov}(\tilde{R}_i, \tilde{R}_m)$ 和 $\widehat{\sigma}^2(\tilde{R}_m)$ 分别是使用月度收益率数据对 $cov(\tilde{R}_i, \tilde{R}_m)$ 和 $\sigma^2(\tilde{R}_m)$ 的估计。本文选用“上证综合指数”作为 \tilde{R}_{mt} 的近似替代，其样本股是在上海证券交易所全部上市股票，由派许公式计算的以报告期发行股数为权数的加权综合股价指数，使用“上证综合指数”一方面是考虑到其是中国的“大股指”能够较好地反映市场组合变动；另一方面是由于数据区间的匹配性²。

Blume (1970) 证明对于任何投资组合 p ，如果其是多支证券基于权重 x_{ip} ， $i = 1, 2, \dots, N$ 来定义的，那么投资组合 p 的风险就可以表示为：

$$\hat{\beta}_p = \frac{\widehat{cov}(\tilde{R}_p, \tilde{R}_m)}{\widehat{\sigma}^2(\tilde{R}_m)} = \sum_{i=1}^N x_{ip} \frac{\widehat{cov}(\tilde{R}_i, \tilde{R}_m)}{\widehat{\sigma}^2(\tilde{R}_m)} = \sum_{i=1}^N x_{ip} \hat{\beta}_i$$

经过严谨数学推导后，不难发现投资组合的 $\hat{\beta}_p$ 可以比单个证券的 $\hat{\beta}_i$ 更精确地估计出真实的 β 。相较于本小组在平时作业中采用估计单只股票 $\hat{\beta}_i$ 的方式来检验标准 CAPM，本文在此处最大的改进就是将采用构建投资组合的方式来估计 β 。

同时，为了减少因使用投资组合造成的信息损失，一般是在对单个证券 $\hat{\beta}_i$ 值排序的基础上构建投资组合，进而通过市值加权或平均的方法获得投资组合的 $\hat{\beta}_p$ 值。但需要注意的是，如果只是简单地执行上述操作很可能会导致严重的“回归现象”

(Regress Phenomenon)。在基于单个证券回归出的 $\hat{\beta}_i$ 截面上，较大的 $\hat{\beta}_i$ 往往高于相应的真实 β_i ，较小的 $\hat{\beta}_i$ 往往低于真实 β_i ，进而在对 $\hat{\beta}_i$ 排序的基础上构建投资组合会造成投资组合内正负抽样误差的堆积，其结果是一个较大的投资组合风险 $\hat{\beta}_p$ 会倾向于高估真实的 β_p ，而较小的 $\hat{\beta}_p$ 则倾向于低估真实值。

² 上证综合指数于 1990 发布，完整地覆盖了 2000 年到 2019 年整个数据区间

为了尽可能地避免“回归现象”，本文采用 Fama-MacBatch 回归(Fama 和 MacBetch (1973)) 的方法来完成投资组合的构建以及后续模型的计算：

步骤 1. 把基于一个时间段内数据计算出的单支股票的 $\hat{\beta}_i$ 进行排序，作为构建投资组合的依据，把这段时间称为“组合构建期”；

步骤 2. 用随后时间段中的数据，估计出这些投资组合中每一支股票的初始 $\hat{\beta}_i$ ，把这段估计初始 $\hat{\beta}_i$ 的时间称为“初始估计期”；

步骤 3. 将初始估计期按年向前滚动，估计后续时间段内投资组合每一年的 $\hat{\beta}_p$ ，同时获取月度的平均收益率数据。基于该数据完成模型的回归估计，将这段时间称为“模型检测期”。

这样使用后面时间段的“新”数据，能够保证在一个投资组合中单支股票 $\hat{\beta}_i$ 的估计误差在很大程度上是随机的，进而使得估计投资组合 $\hat{\beta}_p$ 时产生的“回归现象”能够被尽量减弱。

3.2.2 方法细节

图 2.以 2000 到 2012 年数据区间为例，展示了本文进行参数估计的细节。设 N 为证券总数，设 $\text{int}(N/20)$ 为等于或小于 $N/20$ 的最大整数。首先使用前 4 年(2000–2003)“组合构建期”的月度收益数据估计出每一股票的 $\hat{\beta}_i$ ，将其排序作为构建 20 个投资组合的参考，1 号投资组合到 20 号投资组合中股票的 $\hat{\beta}_i$ 递减。这样的话，如果 N 是偶数，第一个（ $\hat{\beta}_i$ 最大）和最后一个（ $\hat{\beta}_i$ 最小）组合各有 $\text{int}(N/20) + \frac{1}{2}[N - 20\text{int}(N/20)]$ 个证券，如果 N 是奇数则最后一个投资组合有相对多出的证券。

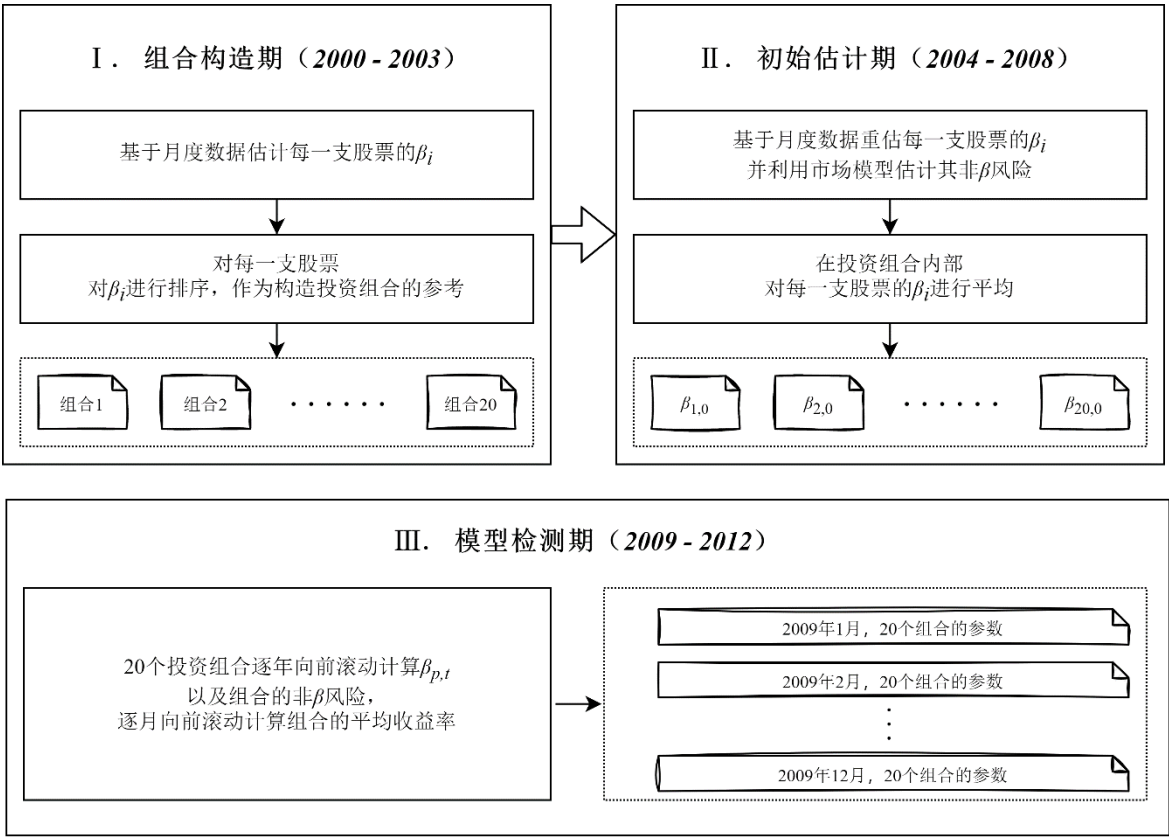


图 2. 采用 Fama-MacBatch 回归验证标准 CAPM 方法框架图
(以 2000 到 2012 年数据区间为例)

接下来的 5 年 (2004-2008 年) “初始估计期” 的数据被用来重新计算 $\hat{\beta}_i$, 这些数据在投资组合内的不同股票之间进行平均, 进而获得 20 个投资组合的初始 $\hat{\beta}_{pt}$ 。递增下标 t 表示在接下来的 4 年 (2008-2011 年) 的每年中这些 $\hat{\beta}_{pt}$ 被滚动重新计算为单个证券 $\hat{\beta}_i$ 的简单平均值, 从而逐年调整组合 $\hat{\beta}_{pt}$, 以考虑个别股票的退市。注意到, 本文和 Fama 和 MacBetch (1973) 所用的方法并不完全一致, 其对 $\hat{\beta}_{pt}$ 进行的是逐月滚动调整, 但本文考虑到数值的稳定性 $\hat{\beta}_{pt}$ 变动不会过于频繁, 故采用逐年滚动的方法进行调整。

另外, 使用 $s(\hat{\epsilon}_i)$ 作为对股票 i 非 β 风险的衡量, 其来源于估计市场模型 (4) 时所产生的最小二乘残差 $\hat{\epsilon}_{it}$ 的标准差:

$$\tilde{R}_{it} = \alpha_i + \beta_i \tilde{R}_{mt} + \tilde{\epsilon}_{it} \quad (4)$$

之所以可以这么表示, 是因为股票的风险在某种程度上可以由其收益率分布的离散度来衡量, 那么在由风险规避者主导的市场中, 一支股票的预期收益与它的总收益分散性有关, 而不仅仅是与该股票对有效投资组合的收益分散性的贡献有关。如果 $\beta_i = \text{cov}(\tilde{R}_i, \tilde{R}_m) / \sigma^2(\tilde{R}_m)$ 那么在 (4) 中 $\text{cov}(\tilde{\epsilon}_i, \tilde{R}_m) = 0$, 进而有:

$$\sigma^2(\tilde{R}_i) = \beta_i^2 \sigma^2(\tilde{R}_m) + \sigma^2(\tilde{\epsilon}_i) + 2\beta_i \text{cov}(\tilde{R}_m, \tilde{\epsilon}_i)$$

不难看出 $s(\hat{\epsilon}_i)$ 是对股票 i 收益分布中与 β_i 不直接相关的分散性的估计, 也即股票 i 的非 β 风险。

在 2009–2012 年的 4 年间, 20 个投资组合的逐月平均收益率也被计算出来, 注意本文在此处直接对投资组合中每支股票的收益率取平均。对于这一时期所有月份的数据, 进行如下截面回归:

$$R_{pt} = \hat{\gamma}_{0t} + \hat{\gamma}_{1t} \hat{\beta}_{p,t-1} + \hat{\gamma}_{2t} \hat{\beta}_{p,t-1}^2 + \hat{\gamma}_{3t} \bar{s}_{p,t-1} + \hat{\eta}_{pt} \quad (5)$$

$\hat{\beta}_{p,t-1}$ 是投资组合 p 中所有股票 $\hat{\beta}_i$ 的平均值; $\hat{\beta}_{p,t-1}^2$ 是这些 $\hat{\beta}_i$ 的平方的平均值; 而 $\bar{s}_{p,t-1}(\hat{\epsilon}_i)$ 是投资组合 p 中每一支股票 $s(\hat{\epsilon}_i)$ 的平均值。

方程 (5) 是 (2) 对一个投资组合中所有股票平均化后的形式, 把逐年的 $\hat{\beta}_{p,t-1}$, $\hat{\beta}_{p,t-1}^2$, 以及 $\bar{s}_{p,t-1}(\hat{\epsilon}_i)$ 作为解释变量, 逐月的平均收益率作为被解释变量, 对随机系数 $\hat{\gamma}_{0t}$, $\hat{\gamma}_{1t}$, $\hat{\gamma}_{2t}$ 以及 $\hat{\gamma}_{3t}$ 进行最小二乘估计。使用 2009–2012 这 4 年间的数据对

方程 (5) 进行回归得到的系数估计值 $\hat{\gamma}_{0t}$, $\hat{\gamma}_{1t}$, $\hat{\gamma}_{2t}$ 以及 $\hat{\gamma}_{3t}$ 是这一时期标准 CAPM 模型假设检验的依据, 要想获得其他时期的检验结果, 只需重复执行上述步骤: 本文用 7 年的数据来构建投资组合; 用接下来的 5 年的数据计算 (5) 中解释变量的初始值; 然后在紧接着的 4 年中逐月向前滚动得到平均收益率, 逐年向前滚动得到解释变量值, 并通过拟合方程 (5) 得到逐月的系数估计值。

表 1 显示本文基于数据选取的 3 个不同的投资组合构建期(除第一个都是 7 年的长度)、5 年的初始估计期和测试期 (除了最后一个都是 4 年)。选择 4 年的测试期是为了平衡计算成本。选择 7 年的投资组合形成期和 5 年的初始估计期来估计风险—收益回归中的独立变量 $\hat{\beta}_{p,t-1}$ 和 $\bar{s}_{p,t-1}(\hat{\epsilon}_i)$, 反映了平衡从静止过程中获得的大样本的统计能力与 p 的任何非恒定性所引起的潜在问题的愿望。因此要对不满足需要的数据进行剔除, 本文为了操作方便, 所以选取了 2000 年到 2019 年 20 年间持续进行交易的股票数据, 共有 724 支股票, 但是其中有部分股票个别月份停止交易, 这对整体计算影响并不大故不作考虑。

表 1. 显示本文基于数据选取的 3 组不同的“组合构建期”(除第一个都是 7 年的长度)、5 年的“初始估计期”和“模型测试期”(除了最后一个都是 4 年)。由于时间区间的限制, 需要剔除不满足条件的数据, 本文为了操作方便, 直接选取了 2000 到 2019 年这 20 年里均有交易记录的 724 支股票的数据, 尽管其中部分股票在个别月份曾有中断, 但对整体计算的影响并不大。

表 1. 三组“组合构建期”, “初始估计期”, “模型检测期”

	时期		
	1	2	3
组合构建期	2000~2003	2001~2007	2005~2011
初始估计期	2004~2008	2008~2012	2012~2016
模型检测期	2009~2012	2013~2016	2017~2019

3.2.3 方法实施示例

随本文提交代码中的 *FamaMacBatch.do* 文件对上述方法做了完整的实现，为了便于读者理解，此处本文以表 1. 中的时期 1（2000 至 2012 年）为示例，展示该方法的具体实施步骤。

步骤 1. 2000 年至 2003 年，基于 48 个月“组合构建期”中的数据，通过市场模型 (4) 估计出每一支股票的 $\hat{\beta}_i$ ：

$$(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_{724})'$$

对所有股票的 $\hat{\beta}_i$ 进行排序，由小至大等分为 20 份，相应地构建起 20 个投资组合：

$\hat{\beta}_i$ 最小 5% 对应的股票为投资组合 1（38 支股票），其次为投资组合 2（36 支股票）

以此类推，最大的 5% 对应的证券为投资组合 20（38 支股票）。

步骤 2. 2004 年至 2008 年，基于 60 个月“初始估计期”中的数据，用同样的方法估计出该区间中每一支股票的 $\hat{\beta}_i$ 以及 $s(\hat{\epsilon}_i)$ ，按照步骤 1. 得到的投资组合划分，在 20 个投资组合内部对每一支股票进行平均进而获得每个投资组合在当期的 $\hat{\beta}_{p,0}$ 以及 $\bar{s}_{p,0}$ ：

$$\begin{aligned}\hat{\beta}_{p,0} &= (\hat{\beta}_{p1,0}, \hat{\beta}_{p2,0}, \dots, \hat{\beta}_{p20,0})' \\ \bar{s}_{p,0} &= (\bar{s}_{p1,0}, \bar{s}_{p2,0}, \dots, \bar{s}_{p20,0})'\end{aligned}$$

然后计算“模型检测期”第一年 2009 年 12 个月 20 个投资组合的平均收益率

$$R_{p,1} = \begin{bmatrix} R_{p1,1} & \cdots & R_{p20,1} \\ \vdots & \ddots & \vdots \\ R_{p1,12} & \cdots & R_{p20,12} \end{bmatrix}$$

向前滚动 1 年，基于 2004 年至 2009 年 72 个月的数据，重复上述操作，得到当期 $\hat{\beta}_{p,1}$

以及 $\bar{s}_{p,1}$ ：

$$\begin{aligned}\hat{\beta}_{p,1} &= (\hat{\beta}_{p1,1}, \hat{\beta}_{p2,1}, \dots, \hat{\beta}_{p20,1})' \\ \bar{s}_{p,1} &= (\bar{s}_{p1,1}, \bar{s}_{p2,1}, \dots, \bar{s}_{p20,1})'\end{aligned}$$

紧接着计算“模型检测期”第二年 2010 年 12 个月 20 个投资组合的平均收益率

$$R_{p,2} = \begin{bmatrix} R_{p1,13} & \cdots & R_{p20,13} \\ \vdots & \ddots & \vdots \\ R_{p1,24} & \cdots & R_{p20,24} \end{bmatrix}$$

继续向前滚动 1 年，基于 2004 年至 2010 年 84 个月的数据，重复上述操作，得到当期 $\hat{\beta}_{p,2}$ 以及 $\bar{s}_{p,2}$ ：

$$\begin{aligned}\hat{\beta}_{p,2} &= (\hat{\beta}_{p1,2}, \hat{\beta}_{p2,2}, \dots, \hat{\beta}_{p20,2})' \\ \bar{s}_{p,2} &= (\bar{s}_{p1,2}, \bar{s}_{p2,2}, \dots, \bar{s}_{p20,2})'\end{aligned}$$

然后计算“模型检测期”第三年 2011 年 12 个月 20 个投资组合的平均收益率

$$R_{p,3} = \begin{bmatrix} R_{p1,25} & \cdots & R_{p20,25} \\ \vdots & \ddots & \vdots \\ R_{p1,36} & \cdots & R_{p20,36} \end{bmatrix}$$

继续向前滚动 1 年，基于 2004 年至 2011 年 96 个月的数据，重复上述操作，得到当期 $\hat{\beta}_{p,2}$ 以及 $\bar{s}_{p,2}$ ：

$$\begin{aligned}\hat{\beta}_{p,3} &= (\hat{\beta}_{p1,3}, \hat{\beta}_{p2,3}, \dots, \hat{\beta}_{p20,3})' \\ \bar{s}_{p,3} &= (\bar{s}_{p1,3}, \bar{s}_{p2,3}, \dots, \bar{s}_{p20,3})'\end{aligned}$$

然后计算“模型检测期”第四年 2012 年 12 个月 20 个投资组合的平均收益率

$$R_{p,4} = \begin{bmatrix} R_{p1,37} & \cdots & R_{p20,37} \\ \vdots & \ddots & \vdots \\ R_{p1,48} & \cdots & R_{p20,48} \end{bmatrix}$$

步骤 3. 汇总 2009 年至 2012 年 4 年的所有数据，并在时期 2 和时期 3 上重复上述操作，最终得到 20 个投资组合在 132 个月份上的数据，得到如表 2. 所示的数据结构，为后续回归估计做准备。

表 2. 数据预处理结果 (20 个投资组合在 132 个月份上的截面数据)

组合	时间	R_p	$\hat{\beta}_p$	$\hat{\beta}_p^2$	\bar{s}_p
组合 1	2009 年 1 月	$R_{p1,1}$	$\hat{\beta}_{p1,0}$	$\hat{\beta}_{p1,0}^2$	$\bar{s}_{p1,0}$
	\vdots	\vdots	\vdots	\vdots	\vdots
	2010 年 12 月	$R_{p1,12}$	$\hat{\beta}_{p1,0}$	$\hat{\beta}_{p1,0}^2$	$\bar{s}_{p1,0}$
	\vdots	\vdots	\vdots	\vdots	\vdots
	2012 年 1 月	$R_{p1,37}$	$\hat{\beta}_{p1,3}$	$\hat{\beta}_{p1,3}^2$	$\bar{s}_{p1,3}$
	\vdots	\vdots	\vdots	\vdots	\vdots
	2012 年 12 月	$R_{p1,48}$	$\hat{\beta}_{p1,3}$	$\hat{\beta}_{p1,3}^2$	$\bar{s}_{p1,3}$
	\vdots	\vdots	\vdots	\vdots	\vdots
	2009 年 1 月	$R_{p20,1}$	$\hat{\beta}_{p20,0}$	$\hat{\beta}_{p20,0}^2$	$\bar{s}_{p20,0}$
	\vdots	\vdots	\vdots	\vdots	\vdots
组合 20	2010 年 12 月	$R_{p20,12}$	$\hat{\beta}_{p20,0}$	$\hat{\beta}_{p20,0}^2$	$\bar{s}_{p20,0}$
	\vdots	\vdots	\vdots	\vdots	\vdots
	2012 年 1 月	$R_{p20,37}$	$\hat{\beta}_{p20,3}$	$\hat{\beta}_{p20,3}^2$	$\bar{s}_{p20,3}$
	\vdots	\vdots	\vdots	\vdots	\vdots
	2012 年 12 月	$R_{p20,48}$	$\hat{\beta}_{p20,3}$	$\hat{\beta}_{p20,3}^2$	$\bar{s}_{p20,3}$
	\vdots	\vdots	\vdots	\vdots	\vdots

3.2.4 初步统计结果

表 3.展示了 20 个投资组合在三个“初始估计期”第一个月的 $\hat{\beta}_{p,t-1}$ 的值，标准误 $s(\hat{\beta}_{p,t-1})$ 。还有由市场模型(4)计算得到的投资组合残差的标准差 $s(\hat{\epsilon}_p)$ 以及组合中所有单支股票残差的标准差的平均值 $\bar{s}_{p,t-1}(\hat{\epsilon}_i)$ 。

假设对于一个给定的股票，方程(4)中的干扰 $\tilde{\epsilon}_{it}$ 是连续独立的，与 \tilde{R}_{mt} 无关，并且通过时间完全分散，那么 $\hat{\beta}_i$ 的标准误差为：

$$\sigma(\hat{\beta}_i) = \frac{\sigma(\tilde{\epsilon}_i)}{\sqrt{n} \sigma(\tilde{R}_m)}$$

其中 n 是用于计算 $\hat{\beta}_i$ 的月数，类似的：

$$\sigma(\tilde{\beta}_{p,t-1}) = \frac{\sigma(\tilde{\epsilon}_p)}{\sqrt{n} \sigma(\tilde{R}_m)}$$

在表 3.中可以看出 $s(\hat{\epsilon}_p)$ 要比 $\bar{s}_{p,t-1}(\hat{\epsilon}_i)$ 小，这也说明通过组合来评估 β 要比使用单支股票进行评估更加精准。

表 3. 20 个组合在三组“初始估计期”的部分统计量

统计量	1	2	3	4	5	6	7	8	9	10
2004~2008 年 “初始估计期”										
$\hat{\beta}_{p,t-1}$	1.019	.989	.981	.976	1.072	1.012	1.117	.969	1.062	.998
$s(\hat{\beta}_{p,t-1})$.052	.038	.035	.037	.044	.052	.243	.035	.035	.032
$s(\hat{\varepsilon}_p)$.023	.024	.030	.031	.027	.036	.026	.024	.030	.033
$\bar{s}_{p,t-1}(\hat{\varepsilon}_i)$.128	.127	.127	.126	.128	.127	.130	.127	.126	.127
2008~2012 年 “初始估计期”										
$\hat{\beta}_{p,t-1}$	1.536	1.264	1.201	1.121	1.159	1.109	1.108	1.071	1.053	1.030
$s(\hat{\beta}_{p,t-1})$.123	.027	.021	.021	.018	.021	.022	.020	.018	.019
$s(\hat{\varepsilon}_p)$.032	.027	.023	.024	.030	.036	.025	.024	.024	.031
$\bar{s}_{p,t-1}(\hat{\varepsilon}_i)$.133	.127	.127	.126	.126	.126	.126	.126	.126	.126
2012~2016 年 “初始估计期”										
$\hat{\beta}_{p,t-1}$	1.512	1.290	1.234	1.187	1.147	1.125	1.092	1.075	1.036	1.053
$s(\hat{\beta}_{p,t-1})$.062	.018	.018	.015	.014	.016	.017	.016	.014	.016
$s(\hat{\varepsilon}_p)$.031	.033	.032	.026	.024	.030	.029	.025	.023	.024
$\bar{s}_{p,t-1}(\hat{\varepsilon}_i)$.131	.127	.127	.126	.126	.126	.126	.126	.126	.126

表 3 (续)

统计量	11	12	13	14	15	16	17	18	19	20
2004~2008 年 “初始估计期”										
$\hat{\beta}_{p,t-1}$	1.001	.976	1.036	1.062	1.068	.938	.968	1.047	.941	.971
$s(\hat{\beta}_{p,t-1})$.043	.034	.138	.043	.047	.035	.032	.045	.033	.039
$s(\hat{\varepsilon}_p)$.028	.034	.034	.027	.036	.024	.027	.031	.033	.033
$\bar{s}_{p,t-1}(\hat{\varepsilon}_i)$.127	.127	.128	.127	.127	.127	.127	.128	.127	.127
2008~2012 年 “初始估计期”										
$\hat{\beta}_{p,t-1}$	1.015	.966	.987	.929	.952	.894	.901	.854	.806	.734
$s(\hat{\beta}_{p,t-1})$.018	.021	.019	.020	.020	.019	.018	.017	.016	.021
$s(\hat{\varepsilon}_p)$.031	.025	.034	.033	.033	.028	.029	.027	.025	.033
$\bar{s}_{p,t-1}(\hat{\varepsilon}_i)$.126	.126	.126	.126	.126	.127	.127	.127	.128	.128
2012~2016 年 “初始估计期”										
$\hat{\beta}_{p,t-1}$	1.006	1.001	.994	.975	.941	.925	.880	.855	.831	.666
$s(\hat{\beta}_{p,t-1})$.017	.018	.015	.015	.014	.015	.013	.014	.013	.015
$s(\hat{\varepsilon}_p)$.030	.031	.035	.032	.029	.026	.029	.028	.026	.035
$\bar{s}_{p,t-1}(\hat{\varepsilon}_i)$.126	.126	.126	.126	.126	.126	.126	.127	.127	.129

3.3 估计结果与假设检验

3.3.1 估计结果呈现

本文对全样本区间（2009~2019）和三个子样本区间（2009~2012、2013~2016、2017~2019）分别做了分析，表 4. 呈现了四个时间区间下相关参数的估计结果。本文分四组对方程（4）的四种形式进行回归，D 组是基于（4）本身。A 到 C 组里方程（4）中的一个或多个变量被移除。表格中显示的 $\bar{\gamma}_j$ 是逐月回归系数 $\hat{\gamma}_{jt}$ （132 个）的平均值； $s(\hat{\gamma}_j)$ 是每月相应估计值的标准差，本表还展示了用于检验 $\bar{\gamma}_j = 0$ 这一假设的 t 统计量：

$$t(\bar{\gamma}_j) = \frac{\bar{\gamma}_j}{s(\hat{\gamma}_j)/\sqrt{n}}$$

其中 $n = 132$ 是时期内的月份数，也是用于计算 $\bar{\gamma}_j$ 和 $s(\hat{\gamma}_j)$ 的 $\hat{\gamma}_{jt}$ 数量。

在解释这些 t 统计量时，本文参考了 Fama（1965a）和 Blume（1970）的结论。这些结论表明，相对于正态分布而言，普通股票的收益分布是“厚尾”的，而且可能更符合非正态对称稳定分布，而不是正态的。从 Fama 和 Babiak（1968）的观点来看，这一结论意味着，当人们在基础变量为正态的假设下解释大的 t 统计量时，得到的显著性水平可能是高估的。但需要注意的是，除了假设 $H3$ （正的预期收益-风险权衡）之外，向上偏的概率水平会导致产生拒绝标准 CAPM 假设的偏差，因此如果在正态的假设下解释 t 统计量时不能拒绝这些假设，那么考虑收益分布的厚尾时，这些假设就会更加可信。

表 4. 估计结果汇总

$$R_{pt} = \hat{\gamma}_{0t} + \hat{\gamma}_{1t} \hat{\beta}_{p,t-1} + \hat{\gamma}_{2t} \hat{\beta}_{p,t-1}^2 + \hat{\gamma}_{3t} \hat{s}_{p,t-1} + \hat{\eta}_{pt}$$

时间区间	参数								
	$\overline{\hat{\gamma}_0 - R_f}$	$t\left(\overline{\hat{\gamma}_0 - R_f}\right)$	$\overline{\hat{\gamma}_1}$	$t\left(\overline{\hat{\gamma}_1}\right)$	$\overline{\hat{\gamma}_2}$	$t\left(\overline{\hat{\gamma}_2}\right)$	$\overline{\hat{\gamma}_3}$	$t\left(\overline{\hat{\gamma}_3}\right)$	\bar{r}^2
A 组:									
2009—2019	.0107	1.06	.0013	.176509
2009—2012	.0066	.32	.0103	.7004
2013—2016	.0306	2.02	-.0070	-.7410
2017—2019	-.0071**	-.67	.0004	.0215
B 组:									
2009—2019	-.0393	-.21	.0993	.27	-.0477	-.2617
2009—2012	-.1608	-.31	.3268	.32	-.1488	-.3012
2013—2016	.0113	.12	.0402	.22	-.0270	-.3217
2017—2019	.0553	.81	-.1250	-.93	.0595	.9622
C 组:									
2009—2019	.0207	.17	.0008	.11	-.0738	-.08	.17
2009—2012	-.2253	-.13	.01	.742254	.17	.14
2013—2016	.0767	.31	-.0091	-1.02	-.3437	-.18	.16
2017—2019	.0039	.02	.0005	.03	-.1128	-.08	.21
D 组:									
2009—2019	-.0372	-.14	.1085	.27	-.0528	-.27	-.0236	-.02	.23
2009—2012	-.2823	-.50	.3874	.37	-.1807	-.35	.7314	.54	.22
2013—2016	.0233	.04	.0375	.13	-.0248	-.18	-.0912	-.03	.22
2017—2019	.1974	.82	-.1686	-1.14	.0804	1.16	-.9403	-.60	.28

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.3.2 核心假设检验

先来看 *H1*, 表 4. 中 B 组和 D 组的结果不能否定标准 CAPM 模型的条件 C1, 即预期收益率和 β 之间的关系的确是线性的。不难发现, B 组和 D 组之中无论是全样本

时间段还是子时间段内的 $t(\hat{\gamma}_2)$ 均十分接近于 0，尽管子时间段 2017–2019 内的 $t(\hat{\gamma}_2)$ 较大，但也都小于 1.8，仍然无法显著地拒绝原假设。

再考虑标准 CAPM 条件 C2 所对应的假设 **H2**，即除了 β 以外，没有任何其他风险会系统性地影响预期收益。表 4. 中 C 组和 D 组的结果也都没有拒绝这一假设，可以发现 $t(\hat{\gamma}_3)$ 都十分接近于 0，且 $t(\hat{\gamma}_3)$ 的符号随机分为正负。

如果只看前面两个假设的话，标准 CAPM 模型似乎能很好地解释数据。然而，如果关键条件 C3 对应的假设 **H3** 被拒绝，以上得到结论的价值都会大打折扣。也就是说，除非风险和收益之间平均存在正的相关性，否则本文对标准 CAPM 的验证就无法达到满意的地步。但令人失望的是，本文的结果和 Fama 和 MacBetch (1973) 的结果大相径庭，所有模型的任意时间区间内的 $t(\hat{\gamma}_1)$ 均十分小，也就是说表 4. 中的结果能够拒绝 **H3**。

然而值得欣慰的是，在对 **H4** 检验上，结果似乎并不像 Fama 和 MacBetch (1973) 中提到的是“ambiguous”的，四组模型中仅有 A 组中两个阶段的数据支持本文拒绝 **H4**，其余部分的 $t(\hat{\gamma}_0 - R_f)$ 均十分接近于 0，表 4. 中的结果无法拒绝 **H4**。

3.4 对标准 CAPM 检验的结论

综上，本文的结果在一定程度上证实了标准 CAPM 的有效性。一方面，可以说尽管在不同时期存在“随机非线性”，但是不能拒绝这样的假设：即在做出投资组合决策时，投资者应该认定证券的投资组合风险和其预期收益之间的关系是线性的。同样结果也不支持拒绝“除了投资组合风险外，没有任何风险措施会系统地影响平均收益”这一假设。另外还说明了无风险利率在标准 CAPM 中扮演的重要角色。但在另一方面，不可否认的是，在所截取的数据上无法得到“平均而言，收益和风险之间

应该有一个正向的关系”这一重要的核心结论，这也促使本文进一步展开了对 Fama-French 三因子模型的实证研究，想要尝试探索一种更加恰当的模型来解释中国股票市场收益率。

四、 Fama-French 三因子模型

4.1 数据补充说明

由于本文后续将引入新的因子，故此处对上述数据进行一定的补充说明。首先，参考 Fama 和 French (1993) 的做法排除了金融公司，因为这些公司的部分指标可能与非金融公司的意义不同；其次，剔除了状态为 ST（特别处理）、*ST（退市风险警示）或 PT（特别转让）的股票，以及账面价值为负的股票，避免这些状态异常的股票对数据产生干扰。

本文把公司资产负债表中披露 $t-1$ 年末的账面所有者权益作为 t 年账面价值，把公司披露的 $t-1$ 年末的市场总值作为 t 年的市场价值。同时，不再直接采用上证综合指数作为市场回报率，而是采用流通市值加权平均计算得出。

在我国以往的信息披露规范制度下，上市公司本年度财务报表一般在次年 3、4 月份公布，4 月底为年报披露截止日期。这就造成了财务报表信息披露有一定的滞后性，进而导致数据库中公司公布的财务数据与实际市场数据时间错位，也就是说 t 年末的财务数据直到 $t+1$ 年 4 月份披露财务报表时才得到更新。所以本文选取 t 年 5 月至 $t+1$ 年 4 月作为组合构建周期，尽可能减少时间错位带来的不利影响。例如某只股票母公司在 2002 年 4 月底才披露 2001 年的年度财务报表，那么 2002 年 1 月至 4 月只能依靠 2001 年披露的 2000 年末的账面价值和市场价值数据进行分组，而 2002 年 5 月至 2003 年 4 月可以采用 2002 年 4 月底披露的数据进行分组，以此类推。

4.2 时间序列回归的输入

不同于上述参数估计时采用的截面回归，分析 Fama-French 三因子模型需要进行时间序列回归。正如 (2) 式所呈现的那样，解释变量包括 **市场因子**（市场组合的超额收益率）、**规模因子**（市值）和 **价值因子**（账面市值比）三个因素，被解释变量为

根据规模（市值）和账面市值比（ BE/ME ）构建的 25 个股票组合的平均超额收益率。本小节将详细介绍这些输入数据的获取方法，并对最终数据结构做展示，之后进行了一些描述性统计。

4.2.1 解释变量：股票三因子

添加因子的动机——规模和账面市值比是解释平均股票收益的特别变量，它们能代表收益中的基本风险因素。Fama 和 French（1992b）记录了公司规模和账面市值比与收益率之间的关系，他们发现高账面市值比的公司往往有较高的资产收益，同时控制账面市值比变量后，小公司往往比大公司的资产收益率高。另外，小公司可能会遭受长期的收益率低落，而大公司则可以避免这种情况。

以上发现表明规模与一个市场风险因子有关，这可能解释了规模和平均收益率之间的**负相关关系**。而账面市值比与收益率之间的关系则表明，相对盈利能力也是平均收益率中一个市场风险因子的来源，这可能解释了账面市值比和平均收益率的**正相关关系**。三因子模型的核心目的正是通过新增规模因子和价值因子来更好地刻画风险，进而更为准确地解释收益率变动。

计算因子的组合构建——为了研究月度的经济基本面，Fama 和 French（1992b）基于在规模和账面市值比上排序的所有股票构建了六个投资组合。本文采用同样的方法构建投资组合，旨在模仿收益中与规模因子、价值因子相关的基本风险因素。

表 5. 基于规模和账面市值比联合划分的六种公司类型

	中 BE/ME		
	低 BE/ME (L)	(M)	高 BE/ME (H)
小规模 (S)	S/L	S/M	S/H
大规模 (B)	B/L	B/M	B/H

Fama 和 French (1992a) 证明了账面市值比在解释股票收益率上比规模具有更大的作用, 因此根据账面市值比将公司分为三类, 根据规模将公司分为两类。表 5.展示了本文划分公司类型的具体方法: 一方面以 t 年 4 月 A 股所有股票的市值中位数将 t 年 5 月至 $t+1$ 年 4 月“组合构建周期”中的公司分为小规模和大规模 (S 和 B) 两组; 另一方面根据此时股票 BE/ME 底部 30% (Low)、中间 40% (Medium) 以及最高 30% (High) 三个分位点将公司分为低、中和高三组 (L、M 和 H)。联合基于规模的两个分组以及基于 BE/ME 的三个分组就可以构建出六个投资组合 (S/L , S/M , S/H , B/L , B/M , B/H)。其中 S/L 投资组合包含在小规模组中的股票, 它们也在低 BE/ME 组中; B/H 组合包含大的规模组中的股票, 它们也在高 BE/ME 组中。六个投资组合的月加权收益率从 t 年 5 月至 $t+1$ 年 4 月计算, 投资组合在 $t+1$ 年 4 月份重新分组。正如上面所提到的, 本文选择从 t 年 5 月开始计算收益率, 是为了确保 t 年的市值以及账面市值比这两大指标是已知的。

规模因子——用投资组合计算 SMB (*small - big*) 来模仿与规模相关的收益率风险因子, 具体计算方法为每个月 3 个小市值投资组合 (S/L , S/M , S/H) 以及 3 个大市值投资组合 (B/L , B/M , B/H) 在平均收益率之间的差额:

$$SMB(t) = \frac{R_{S/H}(t) + R_{S/M}(t) + R_{S/L}(t)}{3} - \frac{R_{B/H}(t) + R_{B/M}(t) + R_{B/L}(t)}{3} \quad (6)$$

将账面市值比作为控制变量, 这种差额应该在很大程度上不受 BE/ME 的影响, 而是集中在大规模与小规模公司的平均收益率差异上。

账面市值比因子——类似的, 采用 HML (*high - low*) 模仿与账面市值比相关收益率的风险因子。 HML 是每个月里两个高 BE/ME 投资组合 (S/H 和 B/H) 平均收益率和两个低 BE/ME 投资组合 (S/L 和 B/L) 平均收益率之间的差额:

$$HML(t) = \frac{R_{S/H}(t) + R_{B/H}(t)}{2} - \frac{R_{S/L}(t) + R_{B/L}(t)}{2} \quad (7)$$

这两个平均收益率之间的差额应该在很大程度上与市值因子无关,而是集中在高 BE/ME 和低 BE/ME 公司之间的不同收益率上。

注意,和标准 CAPM 检验不同的是,此处投资组合的平均收益率不再是由组合中个股收益率直接平均得到,而是由个股流通市值加权计算而来。这样做一方面是为了最小化方差,另一方面也可以更为准确地模拟真实的投资组合。

市场因子——股票收益率的市场因子反映的是超额市场收益率,即 $RM - RF$ 。其中, RM 不再直接用上证综合指数来代表,而是由个股流通市值加权平均法计算得到, RF 是使用一年期定期存款利率表示的无风险利率。

4.2.2 被解释变量：收益率

依据规模和账面市值比的排序构建起 25 个投资组合,并将各个组合的平均超额收益率作为时间序列回归的被解释变量。之所以使用按规模和账面市值比排序构建投资组合,是因为要确定 SMB 和 HML 这两个新增因子是否涵盖了与规模、账面市值比相关的股票收益率中的风险。

此处构建 25 个 $SIZE - BE/ME$ 投资组合的方式类似于上文解释的 6 个 $SIZE - BE/ME$ 的投资组合。在 t 年 5 月,将 A 股所有股票按照规模和账面市值比进行分类,规模和 BE/ME 都是由 t 年 4 月底前披露的 $t-1$ 年相关数据测度的。依靠 A 股所有股票的 5 个规模分位点以及 5 个账面市值比分位点,构建起 25 个投资组合,并计算各投资组合从 t 年 5 月至 $t+1$ 年 4 月的月度加权平均收益率。这 25 个投资组合在 2000 年 1 月至 2019 年 12 月中的月度加权平均收益率是时间序列回归中的被解释变量。

4.2.3 数据处理结果展示

图 3.展示了构建时间序列回归输入数据的流程,对上面提到的股票三因子以及收益率的构造过程做了简洁总结。表 6.则呈现了处理完成的时间序列数据结构。

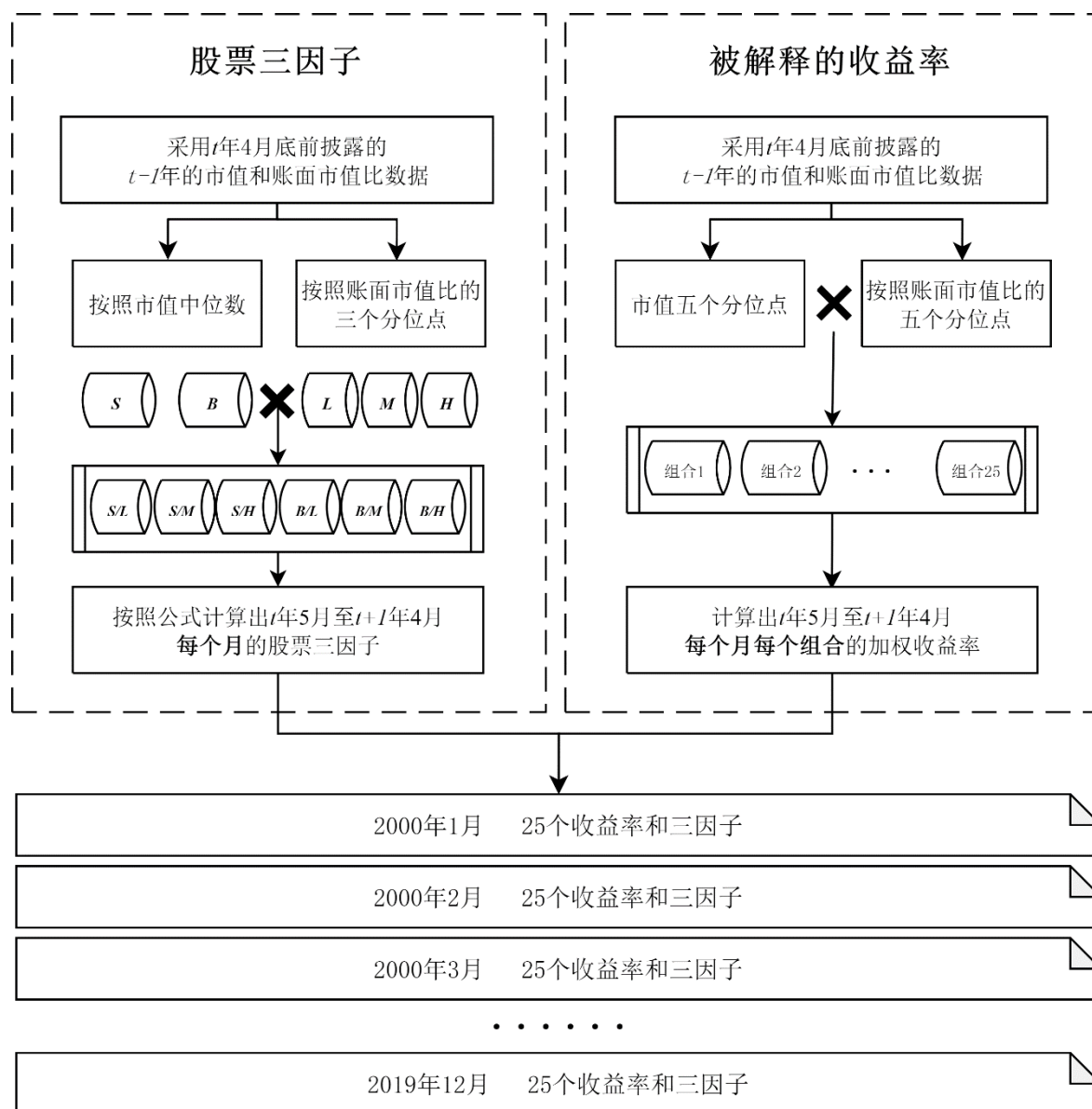


表 6. 时间序列数据结构

组 合	时间	被解释收益率	股票三因子		
			$RM - RF$	SMB	HML
组 合 1	2000 年 1 月	$R_1(1) - RF(1)$	$RM(1) - RF(1)$	$SMB(1)$	$HML(1)$
	⋮	⋮	⋮	⋮	⋮
	2000 年 12 月	$R_1(12) - RF(12)$	$RM(12) - RF(12)$	$SMB(12)$	$HML(12)$
	⋮	⋮	⋮	⋮	⋮
	2019 年 1 月	$R_1(229) - RF(229)$	$RM(229) - RF(229)$	$SMB(229)$	$HML(229)$
	⋮	⋮	⋮	⋮	⋮
	2019 年 12 月	$R_1(240) - RF(240)$	$RM(240) - RF(240)$	$SMB(240)$	$HML(240)$
	⋮	⋮	⋮	⋮	⋮
	2000 年 1 月	$R_{25}(1) - RF(1)$	$RM(1) - RF(1)$	$SMB(1)$	$HML(1)$
	⋮	⋮	⋮	⋮	⋮
组 合 25	2000 年 12 月	$R_{25}(12) - RF(12)$	$RM(12) - RF(12)$	$SMB(12)$	$HML(12)$
	⋮	⋮	⋮	⋮	⋮
	2019 年 1 月	$R_{25}(229) - RF(229)$	$RM(229) - RF(229)$	$SMB(229)$	$HML(229)$
	⋮	⋮	⋮	⋮	⋮
	2019 年 12 月	$R_{25}(240) - RF(240)$	$RM(240) - RF(240)$	$SMB(240)$	$HML(240)$

4.2.4 数据的初步描述性统计

表 7. 展示了本文使用 A 股所有股票交易数据规模和账面市值比分位点构建的 25 个投资组合基本统计信息。可以看出，表中的结果和 Fama 和 French（1993）基于纽约证券交易所的分位点形成投资组合的描述性统计有所不同，公司数量和规模大小以及账面市值比之间没有明显的关系，分布较为均匀。当然，正如所预想的一样，5 个最大规模的投资组合的市值约为市场总价值的 45.97%，并且最大市值和最低账面市值比的股票占据了 25 个组合总价值的 16% 以上。

表 7. 基于规模和账面市值比构建的 25 个投资组合的描述性统计；2000 至 2019 年，20 年

规模	账面市值比									
	Low	2	3	4	High	Low	2	3	4	High
	平均每个月的市场值					平均每个月的公司数				
Small	9.3	9.0	8.3	7.7	7.2	.34	.58	.79	1.08	2.24
2	14.8	14.8	14.7	14.8	14.6	.35	.57	.79	1.06	1.95
3	22.7	22.8	22.6	22.5	22.5	.35	.58	.78	1.06	1.99
4	37.6	37.2	37.4	36.9	37.3	.35	.57	.78	1.06	2.11
Big	138.4	115.6	157.2	190.1	204.5	.32	.57	.78	1.07	2.35
	平均每个月的市场值占比					平均每个月的公司数				
Small	.0059	.0082	.0087	.0100	.0100	30.1	46.3	60.3	90.3	111.1
2	.0123	.0153	.0167	.0155	.0125	60.9	70.3	75.1	72.4	60.9
3	.0197	.0228	.0232	.0226	.0184	67.7	74.9	76.1	69.6	53.0
4	.0373	.0382	.0350	.0311	.0275	77.7	82.9	72.1	61.0	49.7
Big	.1688	.0921	.1017	.0971	.1496	98.1	67.7	59.2	50.9	69.2

如表 8. 所示，基于规模和账面市值比构建的 25 个股票组合产生了范围比较广的月平均超额收益率，从每月 0.49% 到 2.06%。基于本表，可以得出与 Fama 和 French（1992a）中部分相同的结论：一方面，按照账面市值比划分的五组中，平均收益率均随着规模增大而减少，且规模最大组和最小组之间的平均收益率差异也是较为显著的；但在另一方面，按照市值划分的五组中，平均回报率不仅没有呈现“随着 BE/ME

的增加而增加的趋势”，账面市值比最高组和最低组之间的差异还并不显著，甚至在 Small 和 1 组中为负数，这是匪夷所思的，本文将会在后面进行补充讨论。

表 8. 基于规模和账面市值比构建的 25 个投资组合的月平均超额收益率

规模	账面市值比					
	Low	2	3	4	High	High-Low
Small	1.96	2.06	1.91	1.84	1.50	-0.459
2	1.44	1.59	1.57	1.61	1.28	-0.158
3	0.85	1.07	1.38	1.22	1.14	0.286
4	0.76	0.84	1.00	1.03	1.01	0.252
Big	0.49	0.76	0.69	0.78	0.79	0.295
Big-Small	-1.467***	-1.305***	-1.226***	-1.056**	-0.713	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.3 时间序列回归结果

在时间序列回归中，斜率和 R^2 值是不同因子是否捕捉到股票收益变化的直接证据。本文和 Fama 和 French (1993) 一样，通过三种回归来验证三因子在股票收益中发挥的作用，分别是：

- (a) 只使用市场因子即超额市场收益 $RM - RF$ 进行回归；
- (b) 使用市值因子 SMB 和账面市值比因子 HML 作为解释变量的回归；
- (c) 使用三因子 $RM - RF$ 、 SMB 和 HML 进行回归。

本节将报告上述三种回归的结果，分析三因子对股票市场超额收益的解释能力。

4.3.1 市场因子

表 9. 显示, 超额市场收益 $RM - RF$ 的确能捕捉到股票超额收益率的变化。然而, 更值得注意的是, 市场因子在股票收益中留下了许多可以由其他因素解释的变化。25 个组合回归调整 R^2 值中, 只有 4 个较高市值的组合超过了 0.9, 对于小市值股票组合, 调整 R^2 值大都在 0.7 至 0.9 之间, 最小规模和最小账面市值比的组合调整 R^2 甚至低于了 0.7。而正如接下来将要展示的那样, 对于这些股票组合, 市值和账面市值因素 (SMB 和 HML) 将有更好的解释力。

表 9. 25 个组合的超额股票收益率对市场因子 $RM - RF$ 的回归结果

2000 年 1 月至 2019 年 12 月, 240 个月

$$R(t) - RF(t) = a + b[RM(t) - RF(t)] + e(t)$$

相关回归参数值										
账面市值比										
规模	Low	2	3	4	High	Low	2	3	4	High
b						$t(b)$				
Small	1.13***	1.11***	1.13***	1.14***	1.10***	20.39	18.91	21.97	19.37	20.98
2	1.10***	1.12***	1.11***	1.13***	1.12***	19.23	21.36	22.60	24.46	20.32
3	1.11***	1.09***	1.11***	1.13***	1.12***	25.37	29.44	26.03	25.68	30.26
4	1.06***	1.10***	1.12***	1.15***	1.09***	24.16	26.46	33.22	33.10	45.13
Big	0.95***	1.04***	1.04***	1.00***	0.94***	33.45	50.02	40.32	40.82	25.40
调整 R^2						$s(e)$				
Small	0.65	0.68	0.69	0.70	0.73	6.66	6.21	6.10	6.04	5.45
2	0.73	0.76	0.76	0.79	0.81	5.34	5.13	5.01	4.63	4.42
3	0.78	0.81	0.82	0.84	0.88	4.77	4.26	4.18	3.99	3.30
4	0.79	0.85	0.89	0.90	0.92	4.35	3.75	3.25	3.14	2.57
Big	0.85	0.93	0.94	0.90	0.83	3.28	2.24	2.20	2.69	3.43

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.3.2 规模因子和价值因子

表 10. 显示, 在没有市场因子的情况下, *SMB* 和 *HML* 的确也能捕捉到股票收益的少量时间序列变化: 25 个调整 R^2 值中有 20 个高于 0.1, 12 个高于 0.3。然而, 特别是对于规模较大的五分之一的投资组合, *SMB* 和 *HML* 对股票收益的解释能力并不强, 这刚好和表 9. 形成了“互补”。反过来, 对于表 9. 中调整 R^2 较低的几个组合表 10. 中的调整 R^2 明显较高。

表 10. 25 个组合的超额股票收益率对规模因子 (*SMB*) 和价值因子 (*HML*) 的回归结果

2000 年 1 月至 2019 年 12 月, 240 个月

$$R(t) - RF(t) = sSMB(t) + hHML(t) + e(t)$$

规模	相关回归参数值									
	账面市值比									
	Low	2	3	4	High	Low	2	3	4	High
	<i>s</i>					<i>t(s)</i>				
Small	1.63***	1.61***	1.76***	1.80***	1.71***	7.62	8.53	10.24	9.34	9.10
2	1.48***	1.50***	1.52***	1.51***	1.52***	8.17	9.18	8.50	8.78	8.03
3	1.25***	1.27***	1.30***	1.35***	1.19***	7.31	7.57	7.99	7.77	7.00
4	1.05***	1.13***	1.05***	1.07***	0.92***	6.87	6.40	5.50	5.40	5.20
Big	0.24	0.38**	0.48**	0.42**	0.13	1.47	2.27	2.60	2.38	0.89
	<i>h</i>					<i>t(h)</i>				
Small	0.08	-0.07	0.16	0.32*	0.50	0.29	-0.21	0.61	1.05	1.90
2	0.09	0.09	0.19	0.44**	0.76	0.33	0.34	0.67	1.59	2.49
3	-0.15	0.10	0.23*	0.54***	0.76	-0.51	0.37	0.93	1.84	2.71
4	-0.26	0.11	0.30*	0.45***	0.83	-1.00	0.38	1.13	1.68	3.04
Big	-0.68**	-0.08	0.49***	0.81***	0.92	-2.48	-0.30	1.59	2.74	3.67
	调整 R^2					<i>s(e)</i>				
Small	0.37	0.41	0.45	0.44	0.41	8.99	8.37	8.16	8.24	8.00
2	0.36	0.37	0.37	0.34	0.32	8.25	8.23	8.16	8.31	8.31
3	0.31	0.29	0.29	0.27	0.21	8.43	8.23	8.34	8.51	8.58
4	0.27	0.23	0.18	0.17	0.13	8.20	8.45	8.68	8.91	8.57
Big	0.10	0.03	0.03	0.05	0.08	7.94	8.55	8.55	8.29	7.97

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.3.3 三因子

表 10.显示，单独使用 *SMB* 和 *HML* 对股票超额收益的解释力并不强，但似乎与市场因子形成互补。表 10.展示了将三因子全部包含在模型后的回归结果，这个表可以说是本部分最为重要的表。令人欣慰的是，在本文的数据上，Fama-French 三因子模型得到了较好的验证，这三个股票因子的确捕捉到了股票超额收益的大部分变化。

首先，*SMB* 因子斜率的 t 统计量几乎都大于 2，大多数都大于 10。这是十分显著的，*SMB* 作为对规模风险的模拟，明显捕捉到了市场因子和 *HML* 所忽略的股票收益率变化。此外，正如预想的那样，*SMB* 的斜率与规模大小明显相关，在每一个账面市值比的五分位数分组中，*SMB* 的斜率从较小规模到较大规模均呈现单调递减的趋势。

其次，对于 *HML* 的斜率，即模拟账面市值风险的回报，与账面市值比也有系统的关系。在每个规模的五分之一分组中，*HML* 的斜率单调地增加，从最低账面市值比五分之一的显著负值到最高五分之一的显著正值，而且几乎所有的值都在 1% 的显著性水平下是显著的。*HML* 明显捕捉到了市场因子和 *SMB* 所忽略的股票收益的共同变化，这与账面权益有关。

最后，在上述 *SMB* 因子和 *HML* 因子的显著斜率下，三因子共同回归必然导致 R^2 的大幅增加。表 9.中单独使用市场因子回归，只有四个组合的调整 R^2 大于 0.9，而表 11.中调整 R^2 小于 0.9 则变成了“凤毛麟角”（25 个中只有 1 个小于 0.9）。对于规模最小的五分之一的五个投资组合，调整 R^2 从表 9.的 0.65 到 0.73 显著增加成为了表 11.中的 0.87 和 0.96 之间。在表 11.中即使是最小的 R^2 （最小规模和最小账面市值比组合的 0.88），也远远大于市场因子单独产生的 0.65。

表 11. 25 个组合的超额股票收益率对三因子的回归结果
 2000 年 1 月至 2019 年 12 月, 240 个月
 $R(t) - RF(t) = a + b[RM(t) - RF(t)] + sSMB(t) + hHML(t) + e(t)$

Size quintiles	相关回归参数值									
	账面市值比									
	Low	2	3	4	High	Low	2	3	4	High
	b					$t(b)$				
Small	1.02***	1.00***	1.00***	1.01***	0.98***	27.71	38.64	57.32	60.37	44.19
2	0.99***	1.01***	1.01***	1.02***	1.01***	56.27	59.67	59.57	52.64	41.08
3	1.03***	1.01***	1.02***	1.04***	1.05***	40.88	55.52	60.05	44.14	52.33
4	1.00***	1.03***	1.06***	1.09***	1.05***	44.09	53.40	49.09	54.58	60.45
Big	0.97***	1.05***	1.04***	1.01***	0.97***	49.99	59.34	41.64	57.55	38.01
	s					$t(s)$				
Small	1.14***	1.13***	1.28***	1.32***	1.24***	14.48	18.30	36.14	25.45	21.92
2	1.00***	1.02***	1.04***	1.02***	1.03***	15.65	23.27	32.05	20.95	21.13
3	0.76***	0.79***	0.81***	0.85***	0.69***	12.75	18.38	18.96	19.35	13.94
4	0.57***	0.63***	0.54***	0.55***	0.42***	9.76	13.3	9.88	8.87	10.07
Big	-0.22***	-0.12***	-0.03	-0.07	-0.33***	-4.59	-3.21	-0.49	-1.32	-7.74
	h					$t(h)$				
Small	-0.27***	-0.41***	-0.18***	-0.02	0.16*	-3.42	-6.33	-3.43	-0.42	1.80
2	-0.25**	-0.26***	-0.15***	0.09	0.41***	-2.32	-4.42	-3.33	1.15	5.37
3	-0.5***	-0.25***	-0.13*	0.19**	0.40***	-6.02	-4.21	-1.77	2.31	5.75
4	-0.6***	-0.25***	-0.07	0.08	0.47***	-7.09	-3.55	-0.96	0.99	9.28
Big	-1.02***	-0.45***	0.13	0.46***	0.59***	-13.96	-7.29	1.47	5.24	9.07
	R^2					$s(e)$				
Small	0.88	0.94	0.97	0.96	0.96	4.00	2.76	1.98	2.09	2.05
2	0.94	0.97	0.97	0.97	0.96	2.50	1.79	1.76	1.90	2.08
3	0.95	0.96	0.96	0.95	0.95	2.24	1.91	1.99	2.13	2.15
4	0.94	0.95	0.95	0.95	0.95	2.25	2.05	2.18	2.22	2.03
Big	0.95	0.95	0.94	0.93	0.94	1.96	1.94	2.16	2.21	2.02

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

当然本文也注意到了在回归中加入 *SMB* 和 *HML* 对股票的 β 产生了和 Fama 和 French (1993) 发现的一致影响。在表 9 的单因子回归中，最小规模和最低账面市值比的投资组合的 β 是 1.13；在另一个极端，最大规模和最高账面市值比的投资组合的 β 是 0.94。而在表 11 的三因素回归中，这两个组合的 β 分别是 1.02 和 0.97，在回归中加入 *SMB* 和 *HML* 会使股票的 β 向 1.0 塌陷：低 β 向 1.0 移动，高 β 则向下移动。

4.4 对 Fama-French 三因子模型验证的结论

在本文数据上对 Fama-French 三因子模型的验证的结果，几乎和 Fama 和 French (1993) 的结论如出一辙。表 9、表 10、表 11 充分说明了模型中三因子的确可以捕捉到 A 股市场收益率的绝大部分变动。

五、 中国版三因子模型

和大多数基于中国股市数据研究三因子模型的文章一样，本文的结果也存在一定的异常。上文提及的表 8 中体现的“平均回报率没有呈现随着 BE/ME 的增加而增加的趋势”这一结果让本文进行了后续更加深入的思考。

刘等人 (2019) 认为规模因子旨在捕捉与规模有关的股票风险和回报的差异，这些差异来自于与规模有关的相关业务的差异。然而，在中国小型上市企业的股票价格通常反映了与企业基础业务无关的大量价值，而与中国的首次公开发行(IPO)过程有关。在中国，IPO 市场受到严格的监管，公司对公开上市的需求不断增长，但监管局批准 IPO 的处理能力却很低。因此，私营企业寻求另一种方法：反向合并，以便及时成为上市公司。在反向合并中，私营企业以一家上市公司，即所谓的“壳”为目标，通过收购其股份获得控制权，然后，壳公司购买私人公司的资产以换取新发行的股票。虽然逆向并购在其他地方也有发生，但中国的 IPO 限制非常严格，以至于主要交易所中最小的公司成为有吸引力的借壳目标，也就造成了所谓的“壳价值污染现象”。

他们发现中国 83% 的反向并购涉及最小的 30% 的股票，最小的 30% 的公司的回报与经营基本面的关系较小，但与 IPO 活动的关系更大。因此，为了避免在构建因子时受到壳价值的污染，他们剔除了市值底部 30% 的股票。并对 Fama-French 三因子模型中的个别因子基于数据做了相应地调整，得到了“中国版三因子模型。”

于是本文也效仿刘等人 (2018) 通过剔除市值最低 30% 部分的股票的方式，尽量削减“壳价值污染”带来的不利影响，但由于时间限制，本文并未对模型做调整。剔除数据后得到了表 12 中的结果，与表 8 相比，该表中的结果明显能够更好地支持“平均回报率也没有呈现随着 BE/ME 的增加而增加的趋势”这一结论，最大 BE/ME 组和最小 BE/ME 组平均收益率之差没有出现负值，而且部分正值在 10% 的显著性水平上都是显著的。

表 12. 基于规模和账面市值比构建的 25 个投资组合的月平均超额收益率
(去除市值底部 30%的股票后)

规模	账面市值比					
	Low	2	3	4	High	High-Low
Small	1.17	1.43	1.49	1.53	1.20	0.035
2	0.74	1.13	1.32	1.30	1.19	0.448*
3	0.69	0.98	1.10	1.15	0.64	0.464*
4	0.64	0.90	1.00	1.08	0.99	0.352
Big	0.53	0.61	0.77	0.75	0.75	0.220
Big-Small	-0.638	-0.815**	-0.724**	-0.789**	-0.454	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

另外本文也进一步回归得到了表 13. 中的结果, 和表 10. 相比, 可以发现调整 R^2 由最低的 0.88 提升到了 0.93。这都说明了在中国 A 股市场中, 采用剔除低市值股票的方式来减少“壳价值污染”是可行的。

表 13. 25 个组合的超额股票收益率对三因子的回归结果（去除市值底部 30%的股票后）
2000 年 1 月至 2019 年 12 月，240 个月

$$R(t) - RF(t) = a + b[RM(t) - RF(t)] + sSMB(t) + hHML(t) + e(t)$$

相关回归参数值										
账面市值比										
Size quintiles	Low	2	3	4	High	Low	2	3	4	High
<i>b</i>						<i>t(b)</i>				
Small	0.99***	0.98***	1.00***	1.00***	1.01***	36.21	39.52	55.92	48.97	41.98
2	1.01***	1.00***	0.99***	1.02***	1.01***	32.93	55.95	51.61	50.58	55.98
3	0.98***	0.99***	1.03***	1.05***	1.03***	42.33	55.85	57.34	55.05	52.68
4	0.96***	1.06***	1.05***	1.08***	1.06***	34.56	48.79	45.44	51.70	49.66
Big	0.99***	1.03***	1.07***	1.01***	0.95***	42.12	53.68	36.81	39.02	40.35
<i>s</i>						<i>t(s)</i>				
Small	1.06***	1.07***	1.12***	1.06***	1.06***	19.16	27.14	21.07	18.16	1.06***
2	0.97***	0.89***	0.87***	1.05***	0.84***	15.54	12.82	21.98	21.79	0.97***
3	0.71***	0.83***	0.74***	0.81***	0.64***	15.80	17.11	15.18	12.25	0.71***
4	0.45***	0.51***	0.53***	0.55***	0.38***	7.77	8.59	10.52	7.35	0.45***
Big	-0.35***	-0.22***	-0.05	-0.12	-0.41***	-4.30	-0.51	-1.50	-8.25	-0.35***
<i>h</i>										
Small	-0.33***	-0.27***	-0.16***	-0.07	0.25***	-5.18	-4.13	-1.41	4.00	-0.33***
2	-0.42***	-0.30***	-0.17**	0.08	0.34***	-4.88	-2.28	1.50	7.31	-0.42***
3	-0.63***	-0.35***	-0.13**	0.02	0.35***	-6.19	-2.53	0.41	7.63	-0.63***
4	-0.73***	-0.33***	-0.1	0.19***	0.45***	-4.62	-1.64	3.74	8.29	-0.73***
Big	-0.97***	-0.42***	0.06	0.34***	0.54***	-7.82	0.58	4.33	8.60	-0.97***
<i>R</i> ²										
Small	0.93	0.96	0.97	0.97	0.96	2.01	1.82	1.89	2.13	0.93
2	0.94	0.96	0.95	0.96	0.96	2.05	2.08	1.92	1.87	0.94
3	0.95	0.96	0.96	0.96	0.95	1.96	2.01	1.98	2.09	0.95
4	0.93	0.94	0.94	0.95	0.95	2.41	2.26	2.17	2.04	0.93
Big	0.94	0.93	0.92	0.91	0.94	2.24	2.54	2.58	2.00	0.94

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

六、 结论

本文基于中国 A 股市场 2000 年到 2019 年的交易数据，对资本资产定价模型进行实证分析。整体上分为对标准 CAPM 的检验和 Fama-French 三因子模型的分析。

对标准 CAPM 的检验结果可以使本文得到如下结论：

- 1) 投资组合风险和其预期收益之间的关系的确是线性的；
- 2) 不能拒绝除了投资组合风险外，没有任何风险措施会系统地影响平均收益这一假设；
- 3) 无风险利率在标准 CAPM 模型近似充当常数项的作用。但是“收益和风险之间有正向的关系”这一重要结论并未被证实。

对 Fama-French 三因子模型的分析结果充分说明了：对于我国该阶段的股票交易数据而言，市场因子、规模因子 **SMB** 和账面市值比因子 **HML** 能够有效地解释股票超额收益率。但依然存在异常，本文采用“剔除最低市值 30%股票的方式来尽量削减壳价值污染”的方法对其进行了一定程度上的解释。

本文最大的亮点在于：对两模型的具体实证方法做了细致的呈现，不仅简洁明了地绘制了两模型的实证流程图，还采用实例以及数据结构呈现等方式详尽地展示了数据处理和最终回归方式。本文所有的数据和代码均已开源，这让笔者对资本资产定价模型有了更加深刻的理解，同时也为资本资产定价模型的基础学习提供了一套可以复现的开源资料。

参 考 文 献

- Banz (1981) Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of financial economics*, 9(1), 3-18.
- Black (1972) Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of business*, 45(3), 444-455.
- Black、Jensen 和 Scholes (1972) Black, F., Jensen, M. C., & Scholes, M. (1972). *The capital asset pricing model: Some empirical tests*.
- Blume (1970) Blume, M. E. (1970). Portfolio theory: a step toward its practical application. *The Journal of Business*, 43(2), 152-173.
- 刘等人(2019)Liu, J., Stambaugh, R. F., & Yuan, Y. (2019). Size and value in China. *Journal of financial economics*, 134(1), 48-69.
- Fama (1965a) Fama, E. F. (1965a). The behavior of stock-market prices. *The journal of Business*, 38(1), 34-105.
- Fama 和 Babiak (1968) Fama, E. F., & Babiak, H. (1968). Dividend policy: An empirical analysis. *Journal of the American statistical Association*, 63(324), 1132-1161.
- Fama 和 MacBeth (1973) Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy*, 81(3), 607-636.
- Fama 和 French(1992a)Fama, E. F., & French, K. R. (1992a). The cross-section of expected stock returns. *the Journal of Finance*, 47(2), 427-465.
- Fama 和 French (1992b) Fama, E. F., & French, K. R. (1992b). The economic fundamentals of size and book-to-market equity. *Unpublished working paper. University of Chicago*.
- Fama 和 French (1993) Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1), 3-56.

Lakonishok 和 Shapiro (1986) Lakonishok, J., & Shapiro, A. C. (1986). Systematic risk, total risk and size as determinants of stock market returns. *Journal of Banking & Finance*, 10(1), 115-132.

Lintner (1965) Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The journal of finance*, 20(4), 587-615.

Markowitz (1959) Markowitz, H. M. (1959). Portfolio Selection: Efficient Diversification of Investments. *Cowles Foundation Monograph*, 16.

Reinganum (1981) Reinganum, M. R. (1981). A new empirical perspective on the CAPM. *Journal of financial and quantitative analysis*, 16(4), 439-462.

Sharpe (1964) Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3), 425-442.

陈浪南, & 屈文洲. (2000). 资本资产定价模型的实证研究. *经济研究*, 4(61), 444-55.