

FIN 3080

投资分析与投资组合管理

项目一

杨博宇 / 119020065

李东明 / 119020023

王崇锐 / 119020049

2021 年 3 月 8 日



香港中文大學(深圳)

The Chinese University of Hong Kong, Shenzhen

目录

| | |
|---------------------------|----|
| 案例一：沪深主板的 CAPM 实证检验 | 3 |
| 1 CAPM 模型介绍 | 4 |
| 1.1 模型假设 | 4 |
| 1.2 基本公式 | 4 |
| 2 检验思路与原理 | 4 |
| 2.1 计算个股贝塔系数 | 5 |
| 2.2 构造股票投资组合 | 5 |
| 2.3 CAPM 的横截面回归 | 6 |
| 3 样本数据准备工作 | 6 |
| 3.1 个股收益率 | 6 |
| 3.2 市场收益率 | 6 |
| 3.3 无风险利率 | 7 |
| 3.4 样本数据选择 | 7 |
| 4 实证检验结果 | 8 |
| 4.1 个股贝塔系数回归结果 | 8 |
| 4.2 投资组合贝塔系数回归结果 | 8 |
| 4.3 CAPM 横截面回归结果 | 9 |
| 5 案例一结论 | 10 |
| 案例二：有关股票高回报率的传言探究 | 12 |
| 6 小市值股票能否带来高回报率？ | 13 |
| 6.1 样本数据处理 | 13 |
| 6.1.1 基本数据 | 13 |
| 6.1.2 数据筛选 | 13 |
| 6.2 模型建构 | 13 |
| 6.3 回报率可视化结果 | 13 |
| 7 “追涨杀跌”能否带来高回报率？ | 17 |
| 7.1 样本数据 | 17 |
| 7.2 模型建构 | 17 |
| 7.3 回报率可视化结果 | 17 |
| 8 低市净率股票能否带来高回报率？ | 19 |

| | |
|--------------------|----|
| 8.1 样本数据 | 19 |
| 8.2 模型建构 | 19 |
| 8.3 回报率可视化结果 | 19 |
| 9 案例二总结 | 21 |
| 参考文献 | 22 |
| 附录 | 23 |

案例一：沪深主板的 CAPM 实证检验

摘要

作为资产定价模型中最经典的模型之一，CAPM 能否在现实市场中得到实证早已成为学者研究的焦点。为了探究 CAPM 模型是否适用于尚未成熟的中国市场，本篇报告将以沪深主板（包括中小板）全部股票为研究对象，采用 2017 年 1 月 4 日至 2020 年 12 月 27 日的股票数据进行 CAPM 的实证检验。

在得到所需贝塔值的基础上，对收益与风险进行截面回归可以完成对 CAPM 的实证检验。结果表明收益与风险存在显著的正相关性，符合 CAPM 模型。但线性拟合程度较低，说明了 CAPM 并不能完全适用于我国市场。CAPM 本身较为严苛的假设、我国尚未成熟的市场和以及投资观念尚未成熟的投资者体系是导致这一结论的主要因素。

关键词: CAPM, 投资组合, 贝塔系数, 系统性风险, 周回报率, 横截面回归, 时间序列回归

1 CAPM 模型介绍

1.1 模型假设

CAPM，中文称作资本资产定价模型，揭示了股票收益与风险之间严格的定量关系。基于现代资产投资理论（MPT），CAPM 的主要模型假设如下：

- 投资者基于均值-方差模型进行投资选择
- 资本市场具有有效性，即每一时刻的价格能够充分反应已有市场信息
- 股票的价格应在时间序列上线性无关

1.2 基本公式

在假设条件下，证券市场的有效前沿线和从无风险利率出发的资本市场线（CML）相切，切点定义为最佳投资组合。每个投资者的区别在于其投资组合中无风险资产和市场组合的占比。若以股票的预期收益率为 y 轴，股票的贝塔系数为 x 轴建立坐标系，CAPM 的关系式将为坐标系中的一条直线（证券市场线），其关系式为：

$$E(\tilde{r}_j) = r_f + \beta_j [E(\tilde{r}_M) - r_f] \quad (1)$$

其中， r_f 是无风险收益率， $E(\tilde{r}_j)$ 是股票 j 的预期收益率， $E(\tilde{r}_M)$ 是市场组合的预期收益率。 β_j 是股票 j 的贝塔系数，是衡量股票 j 的系统性风险的指标，定义为股票 j 与市场模型预期收益率的协方差与市场模型方差的比值，即：

$$\beta_j = \frac{\sigma_{jM}}{\sigma_M^2} \quad (2)$$

2 检验思路与原理

根据 Fama 和 MacBeth (1973)、陈奕帆等 (2019) 的研究方法，CAPM 的实证检验方法可按照如下步骤进行：

1. 将全体样本数据按时间分为三段，每段包含若干周。
2. 利用第一段的全部数据进行时间序列回归，得到所有个股的贝塔系数回归结果 $\hat{\beta}_i$ 。
3. 根据第一段的 β_i 结果从大到小排序分组，构建投资组合。利用第二段数据对所有投资组合再次进行时间序列回归得到投资组合贝塔系数 $\hat{\beta}_p$ 。
4. 根据第二段的 $\hat{\beta}_p$ 结果作为自变量，利用第三段数据进行横截面回归，进而检验 CAPM 模型。

2.1 计算个股贝塔系数

检验的第一步是利用第一时间段的全部数据进行时间序列回归，得到所有个股的贝塔系数回归结果 $\hat{\beta}_i$ 。

其中，回归因变量为每只样本股在不同时期的周回报率，回归自变量为市场投资组合在不同时期的周回报率。采用的单因子市场模型为：

$$r_{i,t} = \alpha_i + \beta_i \cdot r_{m,t} + \varepsilon_{i,t} \quad (3)$$

$r_{i,t}$ 是股票 i 在 t 周的收益率， $r_{m,t}$ 是市场投资组合在 t 周的收益率， α_i 为状态固定效益参数， ε_i 为回归误差。由于模型中并未加入时间固定效益参数，该模型要求变量数据在时间序列上线性无关。假设 $E(\varepsilon_{i,t})=0$ ，并对(3)式取期望值，得到：

$$E(r_{i,t}) = \alpha_i + \beta_i \cdot E(r_{m,t}) \quad (4)$$

在此模型下，贝塔系数的回归结果是真实贝塔系数的无偏估计值：

$$\hat{\beta}_i = \frac{S_{\tilde{r}_i, \tilde{r}_m}}{\hat{\sigma}_{\tilde{r}_m}^2} \quad (5)$$

2.2 构造股票投资组合

在上一步的市场因子模型中，个股贝塔系数的真实值为个股与市场投资组合收益率的协方差与市场投资组合收益率方差的比值，因此衡量了个股的风险中与市场相关联的部分。误差项 ε_i 的方差 $\sigma_{\varepsilon_i}^2$ 则反应了个股风险中与市场不相关的部分，即非系统性风险。

由于单只股票的非系统性风险较大，若直接利用个股的贝塔系数进行 CAPM 检验，将导致得出的股票收益与风险的相关关系产生较大偏差。因此第二步需要将得到的 $\hat{\beta}_i$ 进行排序，并构建投资组合以分散个股的非系统性风险。利用算术平均法计算得投资组合超额收益，并以市场组合超额收益为自变量做时间序列回归，进而得到投资组合的贝塔系数回归结果 $\hat{\beta}_p$ 。回归模型为：

$$r_{p,t} - r_{f,t} = \alpha_p + \beta_p (r_{m,t} - r_{f,t}) + \varepsilon_{i,t} \quad (6)$$

假设 $E(\varepsilon_{i,t})=0$ ，对(6)式两端取期望，得到：

$$E(r_{p,t} - r_{f,t}) = \alpha_i + \beta_p \cdot E(r_{m,t} - r_{f,t}) \quad (7)$$

其中， $r_{p,t}$ 是组合收益率， $r_{m,t}$ 与 $\varepsilon_{i,t}$ 为市场组合收益率和回归残差。根据 Fama 和 MacBeth (1973) 的研究，这样的投资组合分类保留了第一步回归中的绝大部分数据信息，组合贝塔系数估计 $\hat{\beta}_p$ 相较个股的贝塔系数 $\hat{\beta}_i$ 也更加准确。但同时，该研究分析了 $\hat{\beta}_i$ 的截面数据后发现：拥有高 $\hat{\beta}_i$ 的个股，其真实贝塔系数倾向于被高估；而拥有较低 $\hat{\beta}_i$ 的个股，其真实贝

塔系数倾向于被低估。因此,根据 $\hat{\beta}_i$ 大小排序选择投资组合将导致 $\hat{\beta}_p$ 较高的组合估计值偏大,而 $\hat{\beta}_p$ 较低的组合估计值偏小。

这一问题最终被证实可以通过更新回归数据来解决。通过更新回归 $\hat{\beta}_p$ 的数据时间, $\hat{\beta}_i$ 就有了较大的随机性,估计值的准确度也得到了极大的提高。这也是本文的CAPM实证方法中将数据按时间分段的主要原因。

2.3 CAPM 的横截面回归

最后一步需将第二步得到的所有组合的 $\hat{\beta}_p$ 值作为该组合贝塔系数的估计值。为进行CAPM的横截面回归,最后时间段的数据需通过算术平均法将不同组合周超额收益的面板数据转化为截面数据:

$$\overline{r_{p,t} - r_{f,t}} = \frac{\sum_{t=1}^T (r_{p,t} - r_{f,t})}{T} \quad (8)$$

其中T为最后一段的总周数。为验证组合超额收益与其系统性风险的线性关系,截面回归模型如下:

$$\overline{r_{p,t} - r_{f,t}} = \gamma_0 + \gamma_1 \cdot \beta_p + \varepsilon_p \quad (9)$$

最后对 γ_1 进行显著性检验可判断收益与系统性风险的相关性,并检验是否高系统性风险能够带来高收益,即:

$$E(\gamma_1) = E(r_{m,t}) - E(r_{f,t}) > 0 \quad (10)$$

R^2 值可以判断系统性风险 $\hat{\beta}_p$ 对收益波动情况的解释程度以及模型的拟合程度。对 γ_0 进行显著性检验可以判断是否有系统性风险以外的因素影响收益率。

3 样本数据准备工作

3.1 个股收益率

CAPM实证研究大多选用月收益数据,考虑到此次样本选择时间较短,以下实证研究将采用股票的周收益率进行。其中,股票的现金红利再投资也被考虑在内。计算公式为:

$$r_{i,t} = \frac{P_{i,t} - P_{i,t-1} + D_{i,t}}{P_{i,t-1}} \quad (11)$$

其中 $P_{i,t}$ 为股票 i 在第 t 周的收盘价, $D_{i,t}$ 为对应时期的红利, $r_{i,t}$ 是股票 i 在第 t 周的收益率。

3.2 市场收益率

市场投资组合为所有沪深主板(包括中小板)的股票组成的投资组合。市场收益率将

采用 CSMAR 数据库中的综合周市场回报率。其中，股票的现金红利再投资同样被考虑在内，市场回报率通过等权平均法计算而得，即：

$$r_{m,t} = \frac{\sum_{i=1}^n r_{i,t}}{n} \quad (12)$$

其中， $r_{m,t}$ 表示第 t 周的市场收益率， n 为市场组合中股票的总数。

3.3 无风险利率

无风险利率 $r_{f,t}$ 采用 CSMAR 数据库中的周度化无风险利率。据数据库数据显示，从 2017 年 1 月 4 日至 2020 年 12 月 27 日之间，周度化无风险利率恒为 0.0286%。

3.4 样本数据选择

3.4.1 基本数据

根据单因子模型的回归假设，市场保持有效性且股票价格应在时间序列上线性无关。选择的研究时间为 2017 年 1 月 4 日至 2020 年 12 月 27 日，期间无股灾等事件发生，股票数据较为符合模型假设。研究对象为沪深主板所有股票，样本数量大，符合 CAPM 中市场投资组合构造的要求。数据选取上将股票的周收盘价作为观测值计算收益率，所有数据均来自 CSMAR 数据库。数据按时间等分为三个阶段，平均一个阶段约有 68 周。

3.4.2 数据加工与筛选

1. CSMAR 的沪深股票周回报率数据中共有 4269 只股票。由于所需样本为沪深主板所有股票，需将 CSMAR 的股票周收益率文件中以编号 688 开头的科创板以及以 300 开头的创业板剔除。剔除后，剩余主板 A 股和中小板（60 或 00 开头的股票）与主板 B 股（900 或 200 开头的股票）共 3160 只。
2. 需将 2017 年 1 月 4 日之后上市的股票剔除，以确保参与 CAPM 验证的股票有较多的准确周收益率数据。剔除后，剩余股票 2416 只。
3. 将个股的周收益率与 CSMAR 库中的综合市场收益率对齐后存在两种数据缺失。一种为收益率缺失，但缺失对应周仍有交易（例如 605500 森林包装在 2020 年 12 月 25 日的收益率缺失），此时使用该个股平均收益填充缺失值；另一种为数据中未出现个股某一整条数据，表明存在停牌或退市现象。

对于第二种情况，在第三阶段（2019 年 8 月 26 日至 2020 年 12 月 27 日）中，为确保将面板数据通过算术平均转换为截面数据的准确性，需将在该阶段停牌超过一周的股票全部过滤。经过检验，最后剩余的 2240 只股票在第二阶段计算组合贝塔系数时均有完备的数据。

4 实证检验结果

4.1 个股贝塔系数回归结果

经过数据筛选，第一阶段共有 2416 只股票参与时间序列回归。利用第一阶段（前 68 周）数据并采用市场回归模型(3)，可通过 R 语言回归得出所有样本个股的 $\hat{\beta}_i$ 值，取其中股票代码前十位的回归结果如下表 1（编号 1 代表股票代码 000001，以此类推）：

表 1. 第一期样本个股时间序列回归结果

| 股票编号 | alpha | beta | alpha_t | beta_t | alpha_p | beta_p | R^2 |
|------|------------|----------|------------|-----------|-----------|-------------|-----------|
| 1 | 0.004904 | 0.162501 | 0.856474 | 0.773935 | 0.394882 | 0.441775291 | 0.009131 |
| 2 | 0.008711 | 0.563686 | 1.360806 | 2.401181 | 0.178275 | 0.019212074 | 0.081476 |
| 4 | -0.01082 | 0.672482 | -1.59175 | 2.650358 | 0.117172 | 0.010479667 | 0.113252 |
| 5 | -0.00723 | 0.981727 | -2.49983 | 9.252425 | 0.014959 | 1.78E-13 | 0.568414 |
| 6 | -0.00547 | 0.958579 | -0.7715 | 3.344817 | 0.444834 | 0.001769127 | 0.214376 |
| 7 | -0.01116 | 0.325309 | -1.19193 | 0.870014 | 0.248767 | 0.395747246 | 0.040354 |
| 8 | -0.00528 | 0.420479 | -1.55835 | 3.381936 | 0.124007 | 0.001222547 | 0.149632 |
| 9 | -0.00523 | 1.085623 | -1.50787 | 8.541757 | 0.136434 | 3.19E-12 | 0.528854 |
| 10 | -0.0072801 | 1.483479 | -1.2086855 | 6.7163287 | 0.2311623 | 5.48439E-09 | 0.4096763 |

其中 alpha 与 beta 为市场模型中的常数项与 $\hat{\beta}_i$ 值，三四列为对应的 t 检验值，五六列为对应的 P 值，最后一列为 R^2 数值。下图 1 展现了所有个股回归结果的整体情况，分别记录了所有个股回归结果中常数、 $\hat{\beta}_i$ 、对应检验值、P 值与 R^2 的分布情况（包括最小值、最大值、四分位值、均值等）：

图 1. 第一期个股时间序列回归总体情况

| | alpha | beta | alpha_t | beta_t | alpha_p | beta_p | R^2 |
|-------|-------------|-------------|-------------|-------------|-------------|--------------|-------------|
| type | | | | | | | |
| count | 2416.000000 | 2416.000000 | 2416.000000 | 2416.000000 | 2416.000000 | 2.416000e+03 | 2416.000000 |
| mean | -0.001361 | 0.952687 | -0.399159 | 5.525353 | 0.434237 | 2.166475e-02 | 0.321008 |
| std | 0.006290 | 0.384057 | 1.132349 | 2.553612 | 0.292168 | 1.046247e-01 | 0.167426 |
| min | -0.039662 | -0.622298 | -4.363622 | -2.816119 | 0.000047 | 2.252312e-26 | 0.000002 |
| 25% | -0.004992 | 0.697460 | -1.135307 | 3.848919 | 0.174980 | 2.939909e-09 | 0.198721 |
| 50% | -0.002090 | 0.947585 | -0.476202 | 5.309126 | 0.410155 | 1.611279e-06 | 0.310210 |
| 75% | 0.001677 | 1.199180 | 0.341897 | 6.901513 | 0.674367 | 2.922936e-04 | 0.430700 |
| max | 0.062462 | 2.681249 | 3.345816 | 17.565359 | 0.998824 | 9.939805e-01 | 0.825990 |

4.2 投资组合贝塔系数回归结果

将第一阶段所得个股 $\hat{\beta}_i$ 值按照大小排序后，划分为 100 个投资组合。每个投资组合均

有不少于 20 个不同个股。计算投资组合与市场组合超额收益，并按照模型(6)利用二阶段的数据进行时间序列回归，可得出 100 个投资组合的贝塔系数 $\hat{\beta}_p$ 。取其中前十组组合的回归结果如表 2:

表 2. 第二期股票组合时间序列回归结果

| 组合编号 | alpha | beta | alpha_t | beta_t | alpha_p | beta_p | R^2 |
|------|------------|-----------|------------|-----------|-----------|-------------|-----------|
| 0 | -0.0023119 | 1.1462214 | -1.2399217 | 22.446831 | 0.219462 | 2.5196E-32 | 0.8857364 |
| 1 | 0.0008184 | 1.0903027 | 0.5099751 | 24.809525 | 0.6117963 | 7.36955E-35 | 0.9044838 |
| 2 | -0.0008027 | 1.1416163 | -0.4799263 | 24.921819 | 0.6328909 | 5.64812E-35 | 0.9052612 |
| 3 | 0.0009469 | 1.1818793 | 0.539425 | 24.585628 | 0.5914367 | 1.25625E-34 | 0.9029058 |
| 4 | 0.0023026 | 1.1775339 | 1.4161773 | 26.444197 | 0.161498 | 1.68346E-36 | 0.9149544 |
| 5 | 0.0010383 | 1.1617012 | 0.6403488 | 26.161335 | 0.5241969 | 3.19227E-36 | 0.9132657 |
| 6 | 0.000832 | 1.1516627 | 0.6211781 | 31.394951 | 0.5366548 | 5.36595E-41 | 0.9381331 |
| 7 | 0.0029265 | 1.0693963 | 1.7906469 | 23.893 | 0.0780066 | 6.70788E-34 | 0.8977787 |
| 8 | 0.0006155 | 1.2036183 | 0.4943002 | 35.294423 | 0.6227607 | 4.03076E-44 | 0.950408 |
| 9 | -0.0004801 | 1.0409609 | -0.3331575 | 26.378939 | 0.7400877 | 1.95027E-36 | 0.9145691 |
| 10 | -0.0002012 | 1.1892086 | -0.1203449 | 25.972952 | 0.9045812 | 4.90412E-36 | 0.912114 |

由表 2 可见,所有组合 $\hat{\beta}_p$ 数据显著不为零,表明组合收益与市场收益的关联十分显著。在 5%的显著性水平下,100 个组合中超过 99%的组合的 alpha 值并不显著。这与 CAPM 模型中除了无风险利率外无其余常数项相符。 R^2 大多超过 0.9,表明模型拟合度较高。下图 2 展现了所有组合回归的整体结果,分别记录了所有组合回归结果中常数、 $\hat{\beta}_p$ 、对应检验值、P 值与 R^2 的分布情况(包括最小值、最大值、四分位值、均值等):

图 2. 第二期组合时间序列回归整体情况

| | alpha | beta | alpha_t | beta_t | alpha_p | beta_p | R^2 |
|-------|------------|------------|------------|------------|------------|--------------|------------|
| type | | | | | | | |
| count | 100.000000 | 100.000000 | 100.000000 | 100.000000 | 100.000000 | 1.000000e+02 | 100.000000 |
| mean | -0.000059 | 0.992459 | -0.037860 | 26.405010 | 0.538523 | 2.290025e-20 | 0.905073 |
| std | 0.001272 | 0.134550 | 0.861776 | 4.991626 | 0.270391 | 1.567761e-19 | 0.045977 |
| min | -0.003944 | 0.617645 | -2.153080 | 12.260945 | 0.035028 | 1.102551e-46 | 0.698139 |
| 25% | -0.000846 | 0.896715 | -0.642406 | 23.869734 | 0.308572 | 6.818007e-40 | 0.897599 |
| 50% | 0.000114 | 1.002993 | 0.087455 | 26.411568 | 0.545650 | 1.816868e-36 | 0.914762 |
| 75% | 0.000819 | 1.097707 | 0.555888 | 30.111022 | 0.746431 | 7.135270e-34 | 0.933105 |
| max | 0.002926 | 1.255665 | 1.790647 | 38.813334 | 0.986132 | 1.454849e-18 | 0.958638 |

4.3 CAPM 横截面回归结果

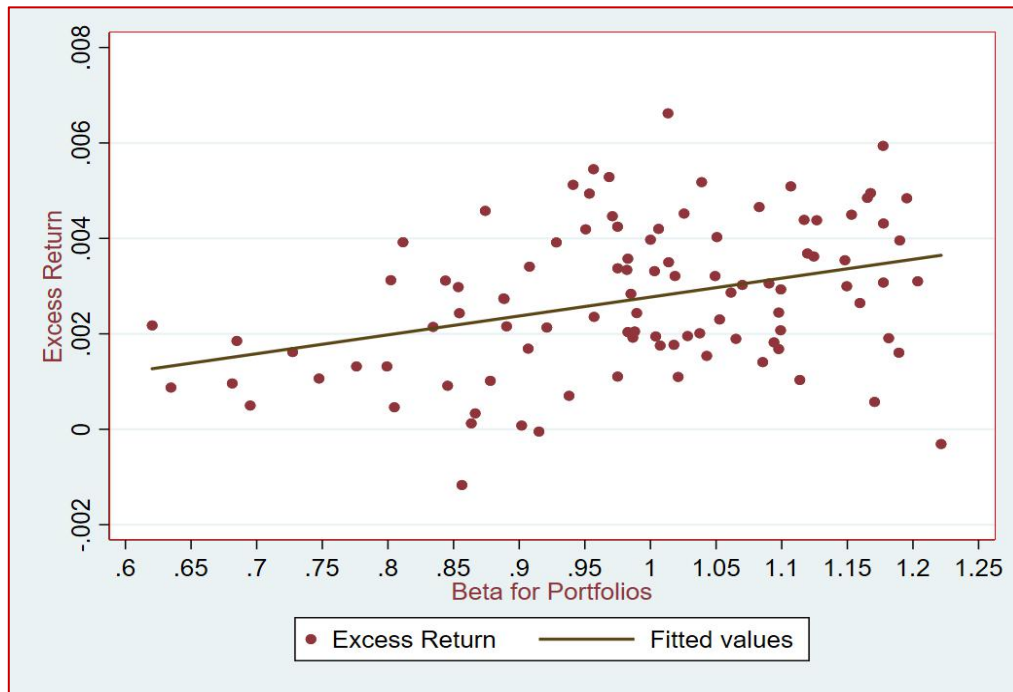
利用第二步投资组合的 $\hat{\beta}_p$ 值作为自变量,组合超额收益率为因变量,回归结果如下表 3:

表 3. 第三期股票组合横截面回归结果

| | γ_0 | γ_1 | R^2 | F 检验值 | 概率值($>F$) |
|---------|------------|------------|--------|---------|-------------|
| 系数/数值 | -0.001185 | 0.003955 | 0.1250 | 16.11 | 0.0001 |
| t 检验值 | -1.23 | 4.01 | | | |
| P 值 | 0.221 | 0.000 | | | |

由该表数据可知, R^2 值仅为 0.125, 说明系统性风险只能解释较小一部分的回报率波动, 模型拟合度一般。而在 1% 的显著性水平下, γ_1 仍与收益率显著正相关, 表明收益随系统性风险增大而增大, 符合 CAPM 模型。常数 γ_0 的 P 值较大, 不显著不为零, 说明系统性风险较好地决定了收益率的高低, 也印证了本文实证方法中将股票构建为投资组合以分散个股非系统性风险的有效性。最终的回归散点拟合图如下:

图 3. 第三期股票组合贝塔系数与组合超额周收益率的线性拟合图



5 案例一结论

实证结果表明, 收益与风险存在显著的正相关性, 同时系统性风险是影响股票收益的主要因素, 与 CAPM 模型较为吻合。但同时, 超额收益与风险具有较弱的线性关系, 即存在着其他的因素影响收益的波动情况。因此, 不能排除非系统性风险对回报率和定价的影响。

相较于无套利理论下的模型, 作为均衡理论下的 CAPM 模型有着较为严苛的模型假设。市场价格可能并不能充分反应市场信息, 投资者的效用函数也千差万别。因此 CAPM 模型本身使得其在现实市场中的应用遭受了种种限制。此外, 我国证券市场起步较晚, 股市规模较小, 证券市场信息公开化程度较低, 与发达国家证券市场还有较大差距。中小投资者占据投资者的绝大部分, 其投资观念不成熟, 也会导致模型检验中的种种差异。

实际上, CAPM 模型最核心的结果, 在于揭示了资产回报率与整体市场回报率的关联, 而并非在于是否可以准确的估计资产收益率。同时可以相信, 随着我国机构投资者的主体化, 多层次市场体系建设的推进以及新股发行体制改革的深化, 我国证券市场将会更加成熟, 证券投资中风险与回报的均衡关系将更加明显。

案例二：有关股票高回报率的市場传言探究

摘要

本案例将对三种有关股票高回报率的市場传言进行探究。包括小市值股票是否可带来较高的平均回报率、过去月份的高回报率股票是否仍能在下月维持高回报率、低市净率的股票是否能带来较高的平均回报率。

以下探究将以沪深主板（包括中小板）所有股票自 2006 年以来的月度数据为基础，视月市值、过去月回报率、市净率为因子，并依据因子构建投资组合。通过考察不同投资组合在未来时期的真实回报率，可检验因子的数值大小对组合回报率影响的显著性。其次，通过每月对投资组合进行调仓重整并记录其未来时期的真实回报率，可将自 2006 年至今的组合回报率可视化，得到更为直观的结果。

分析结果否定了“追涨杀跌”策略可提高股票回报率的可行性。选择低市净率的股票并长期持有是较为优质的投资选择。多数情况下，长期持有低市值股票将带来较优的收益。但是否该策略可以沿用还需经过未来数据与宏观经济结果的检验。

关键词：高回报率，市值大小，追涨杀跌，市净率，投资组合回报，假设检验，数据可视化

6 小市值股票能否带来高回报率？

6.1 样本数据处理

6.1.1 基本数据

采用沪深主板（及中小板）所有股票自 2006 年以来的月回报率进行检验。股票市值使用 CSMAR 数据库中月个股总市值。其中个股月回报率将现金红利再投资考虑在内。

6.1.2 数据筛选

CSMAR 库中目标个股的月回报率数据有较少量的缺失。由于缺失量较小以及股票数量庞大，缺失数据的股票将被剔除，不参与到投资组合的构建之中。构建股票组合时，采用个股的月回报率平均值作为组合的月回报率。

6.2 模型建构

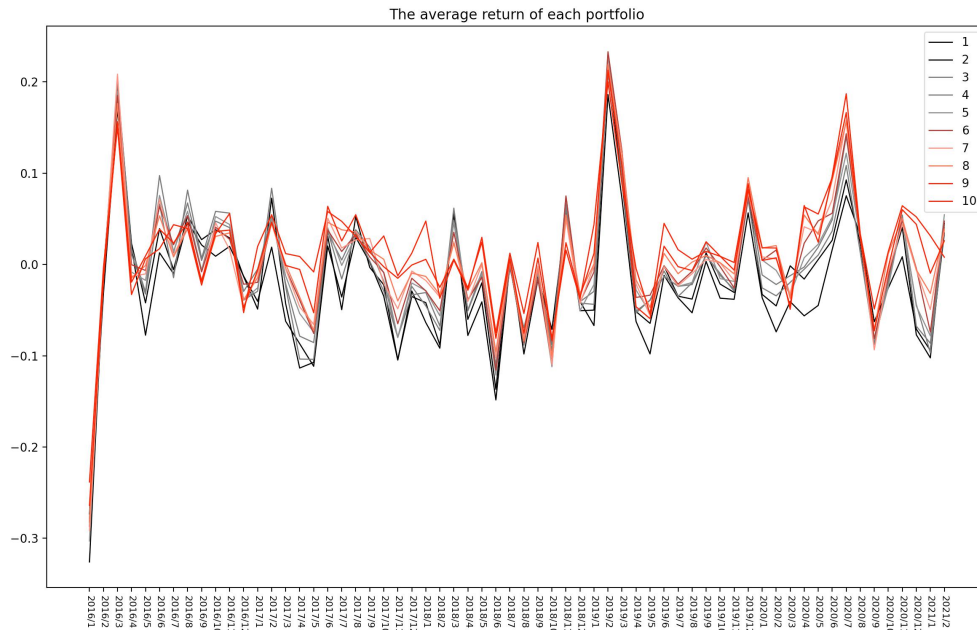
自 2016 年 1 月起，将所有目标个股依据上月总市值进行排序。随后，将个股月市值位于最小十分位的股票纳入投资组合一，市值位于总体样本最小十分位至二十分位的股票纳入投资组合二，以此类推可构建十个投资组合。

对该十个组合在本月或持有一段时间的平均回报率进行记录，作为纵坐标。此后，每月根据市值更新情况调仓，重新构建投资组合。重新记录组合在该月或持有一段时间的平均回报率，作为纵坐标的更新。以时间序列为横坐标，可对组合收益率的结果进行可视化。

6.3 回报率可视化结果

若以每月构建的十个投资组合为基点，记录持有一月的组合回报率，可构建所得收益率的时间序列折线图。取 2016 年 1 月至 2021 年 2 月的时间序列如下（其余时间图参考附录）：

图 4. 依据市值持有一月下的组合回报率 (2016-2021)



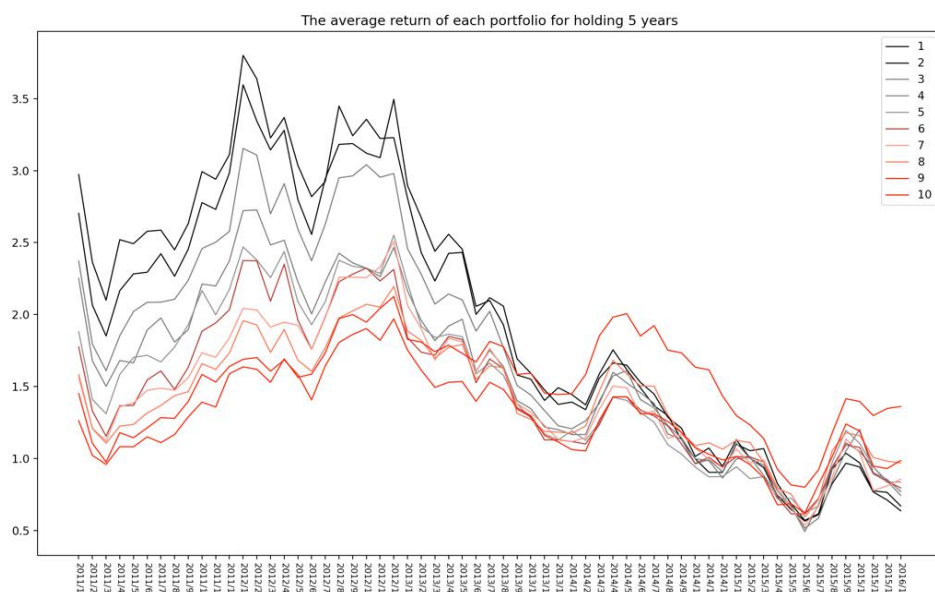
其中，使用红线的五组组合在构建时拥有较高市值，而使用黑线的五组组合市值较低。后五组中，市值越大对应的红线越亮；前五组中，市值越小对应的黑线越亮。由图 4 可直观看，高市值股票组合在短期内会有较高的回报率。利用 Stata 进行显著性检验可以得到，在 1% 的显著性水平下，高市值对应的股票组合的回报率的期望值均显著大于低市值对应的五组股票组合。以组合 3 与 7 对应的 t 检验结果为例：

图 5. 组合 3 与组合 7 的 t 检验结果

| Paired t test | | | | | | |
|----------------------------|-----|------------------------|-----------|--------------------------|----------------------|-----------|
| Variable | Obs | Mean | Std. Err. | Std. Dev. | [95% Conf. Interval] | |
| p3 | 182 | .0155789 | .0078713 | .1061901 | .0000475 | .0311102 |
| p7 | 182 | .0249934 | .0075467 | .1018102 | .0101027 | .0398842 |
| diff | 182 | -.0094146 | .0019196 | .0258974 | -.0132023 | -.0056268 |
| mean(diff) = mean(p3 - p7) | | | | t = -4.9043 | | |
| Ho: mean(diff) = 0 | | | | degrees of freedom = 181 | | |
| Ha: mean(diff) < 0 | | Ha: mean(diff) != 0 | | Ha: mean(diff) > 0 | | |
| Pr(T < t) = 0.0000 | | Pr(T > t) = 0.0000 | | Pr(T > t) = 1.0000 | | |

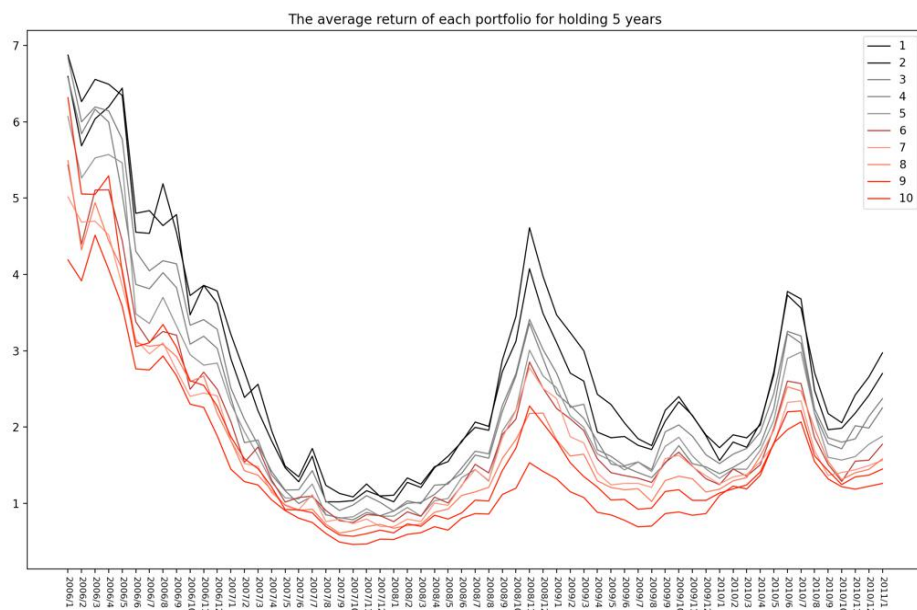
而若考虑每个组合基于未来长期的回报率，可视化结果发生了显著的变化。以每月构建的十个投资组合为基点，记录其未来五年的平均组合回报率，取其中部分时间时间序列图如下：

图 6. 依据市值持有五年下的组合回报率 (2011-2016)



由图可看出，在多数情况下，若长期持有低市值股票组合，其回报率明显高于低市值股票组合。2006 年至 2011 年的情况也是如此：

图 7. 依据市值持有五年下的组合回报率 (2006-2011)



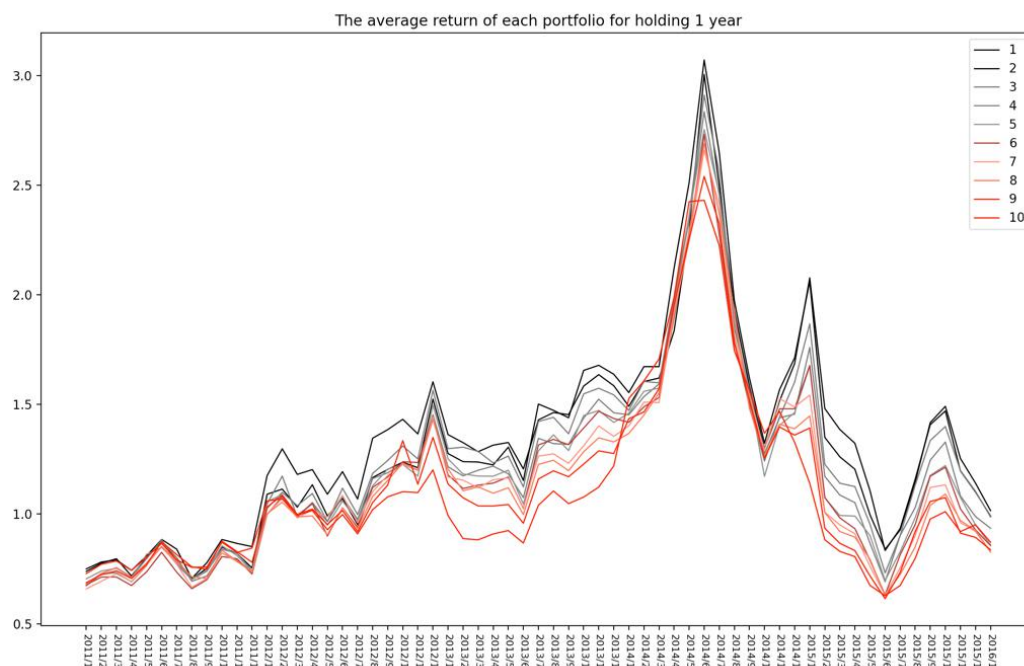
经 Stata 检验，红线代表的五只组合，拥有显著低于黑线代表五只组合的平均收益。以组合 5 与组合 6 的 t 检验为例：

图 8. 组合 3 与组合 7 的 t 检验结果

| Paired t test | | | | | | |
|----------------------------|-----|------------------------|-----------|--------------------------|----------------------|----------|
| Variable | Obs | Mean | Std. Err. | Std. Dev. | [95% Conf. Interval] | |
| p5 | 121 | 1.824673 | .0935139 | 1.028653 | 1.639522 | 2.009824 |
| p6 | 121 | 1.707629 | .0815329 | .8968618 | 1.546199 | 1.869058 |
| diff | 121 | .1170448 | .0161643 | .177807 | .0850406 | .1490489 |
| mean(diff) = mean(p5 - p6) | | | | t = 7.2410 | | |
| Ho: mean(diff) = 0 | | | | degrees of freedom = 120 | | |
| Ha: mean(diff) < 0 | | Ha: mean(diff) != 0 | | Ha: mean(diff) > 0 | | |
| Pr(T < t) = 1.0000 | | Pr(T > t) = 0.0000 | | Pr(T > t) = 0.0000 | | |

而若将持有时间重新减少至一年，图像变为：

图 9. 依据市值持有一年下的组合回报率（2011-2016）



由图可看出，市值低的股票仍在多数时间占有优势，但持有时间越短，优势越不明显。其余时间段的图像以及持有期 3 年的组合回报率图像列在附录之中，可供参考。

因此，市值高的股票组合在短期内平均表现较优。这类股票价格较高，往往也颇受投资者的青睐。但随持仓时间加长，高低市值股票的平均收益率差距逐渐缩小。当持仓时间较长（超过 1 年）时，低市值股票的平均收益表现较优。自 2006 以来的绝大多数时间，长期持有低市值股票往往是不错的选择。而近年来，即使持有期较长，高市值股票与低市值股票平均收益表现也势均力敌，并无明显的优劣之分。

7 “追涨杀跌”能否带来高回报率？

7.1 样本数据

采用沪深主板（及中小板）所有股票自 2006 年以来的月回报率进行检验。其中个股月回报率将现金红利再投资考虑在内。与上问同样，缺失回报率数据的股票将被剔除，不参与投资组合的构建之中。构建股票组合时，采用个股的月收益率平均值作为组合的月回报率。所有数据均来自 CSMAR。

7.2 模型建构

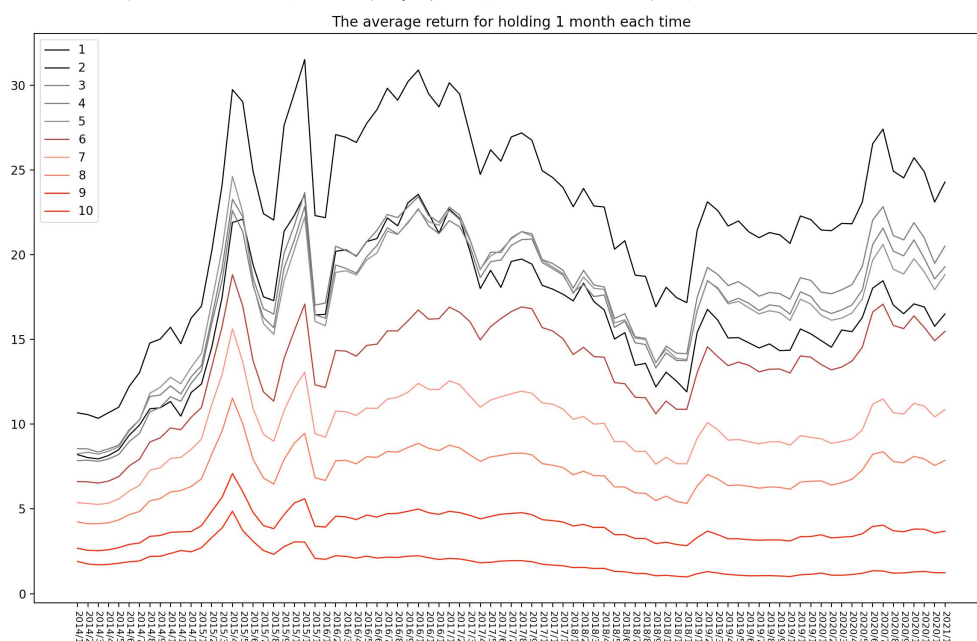
自 2016 年 1 月起，将所有目标个股依据上月的回报率进行排序。随后，将回报率位于最小十分位的股票纳入投资组合一，位于总体样本最小十分位至二十分位的股票纳入投资组合二，以此类推可构建十个投资组合。

对该十个组合持有一月后的平均回报率进行记录，作为纵坐标。此后，每月根据市值更新情况调仓，重新构建投资组合。重新记录组合在持有一月后的平均回报率，作为纵坐标的更新。以时间序列为横坐标，对组合收益率的结果进行可视化。

7.3 回报率可视化结果

若以每月构建的十个投资组合为基点，记录的持有一月组合回报率，可构建所得收益率的时间序列折线图。取 2014 年 1 月至 2021 年 1 月的时间序列如下（其余时间图参考附录）：

图 10. 依据上月回报率持有一月后组合回报率 (2014-2011)



其中，使用红线的五组组合在构建时拥有较高上月回报率，而使用黑线的五组组合上月回报率较低。后五组中，历史回报率越大对应的红线越亮；前五组中，历史回报率越小对应的黑线越亮。由图 6 可直观看出，上月高回报率对应的股票组合（红线）在下月的回报率显著低于上月低回报率对应的股票组合（黑线）。经 Stata 检验，在 1% 的显著性水平下，红线代表的组合的下月回报期望值显著低于黑线代表的组合，说明了“追涨杀跌”策略几乎不具备提高回报的可行性。以组合 1 与组合 10 为例：

图 11. 组合 1 与组合 11 的回报率的 t 检验结果

| Paired t test | | | | | | |
|-----------------------------|-----|------------------------|-----------|--------------------------|----------------------|----------|
| Variable | Obs | Mean | Std. Err. | Std. Dev. | [95% Conf. Interval] | |
| p1 | 181 | 1.020999 | .0077828 | .1047075 | 1.005641 | 1.036356 |
| p10 | 181 | 1.006147 | .007394 | .0994762 | .9915569 | 1.020737 |
| diff | 181 | .0148516 | .0039544 | .053201 | .0070487 | .0226546 |
| mean(diff) = mean(p1 - p10) | | | | t = 3.7557 | | |
| Ho: mean(diff) = 0 | | | | degrees of freedom = 180 | | |
| Ha: mean(diff) < 0 | | Ha: mean(diff) != 0 | | Ha: mean(diff) > 0 | | |
| Pr(T < t) = 0.9999 | | Pr(T > t) = 0.0002 | | Pr(T > t) = 0.0001 | | |

反之，“追跌杀涨”策略在特定时期的回报率表现优异。但同时，该策略的波动性极大，而“追涨杀跌”策略的长期收益较为稳定。只是极低的收益率使得该策略较为无用。

8 低市净率股票能否带来高回报率？

8.1 样本数据

采用沪深主板（及中小板）所有股票自 2006 年以来的月回报率进行检验。其中个股月回报率将现金红利再投资考虑在内。市净率取自 CSMAR 库中的市净率的季度值。与上问同样，缺失回报率数据的股票将被剔除，不参与到投资组合的构建之中。同样地，剔除市净率数据中存在的较少缺失值对应的股票。构建股票组合时，采用个股的月收益率平均值作为组合的月回报率。

8.2 模型建构

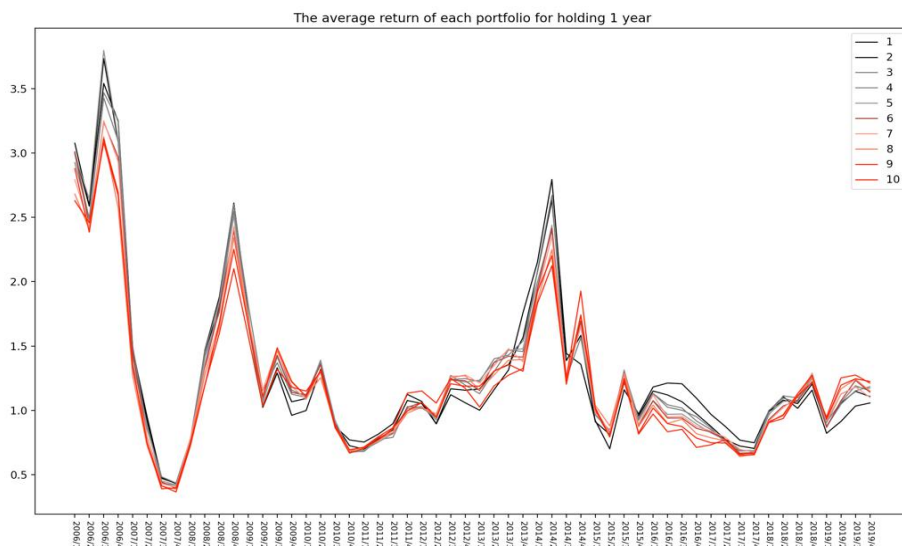
自 2016 年第一季度起，将所有目标个股依据该季度市净率进行排序。随后，将市净率位于最小十分位的股票纳入投资组合一，位于总体样本最小十分位至二十分位的股票纳入投资组合二，以此类推可构建十个投资组合。

对该十个组合持有一年后的平均回报率进行记录，作为纵坐标。此后，每月根据市值更新情况调仓，重新构建投资组合。重新记录组合在持有一年后的平均回报率，作为纵坐标的更新。以时间序列为横坐标，对组合收益率的结果进行可视化。

8.3 回报率可视化结果

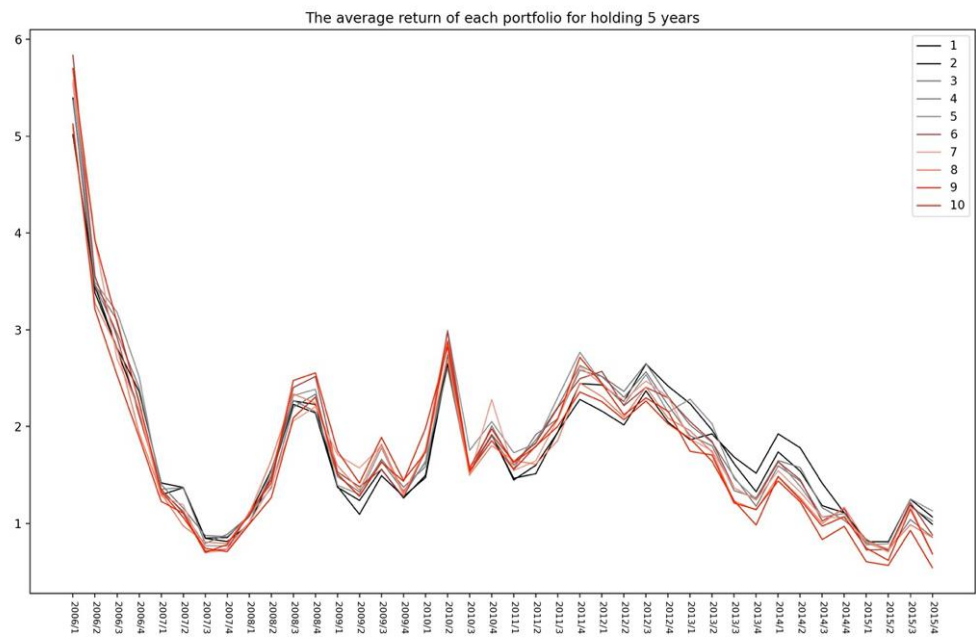
若以每季度构建的十个投资组合为基点，记录持有一年的组合回报率，可构建所得收益率的时间序列折线图。结果如下：

图 12. 依据市净率持有一年后组合回报率



其中，使用红线的五组组合在构建时拥有较高市净率，而使用黑线的五组组合市净率较低。后五组中，市净率越大对应的红线越亮；前五组中，市净率越小对应的黑线越亮。若将组合回报率拉长至五年，所得序列图如下：

图 13. 依据市净率持有五年后组合回报率



由图可知，在 2006-2012 年期间，无论持有期长短，高低市净率组合回报率表现十分相近，并无明显的优劣之分。以组合 5 与组合 6 为例进行 t 检验：

图 14. 组合 5 与组合 6 的 t 检验数据

| Paired t test | | | | | | |
|----------------------------|-----|----------|-----------|------------------------|----------------------|----------|
| Variable | Obs | Mean | Std. Err. | Std. Dev. | [95% Conf. Interval] | |
| p5 | 40 | 1.857854 | .1448964 | .916405 | 1.564773 | 2.150934 |
| p6 | 40 | 1.816239 | .1488392 | .9413417 | 1.515184 | 2.117295 |
| diff | 40 | .0416147 | .0222056 | .1404405 | -.0033003 | .0865298 |
| mean(diff) = mean(p5 - p6) | | | | t = | 1.8741 | |
| Ho: mean(diff) = 0 | | | | degrees of freedom = | 39 | |
| Ha: mean(diff) < 0 | | | | Ha: mean(diff) != 0 | Ha: mean(diff) > 0 | |
| Pr(T < t) = 0.9658 | | | | Pr(T > t) = 0.0684 | Pr(T > t) = 0.0342 | |

可以发现，P 值较大，因此并无显著证据表明两组回报率的均值不同。而图像结果表明，在 2012 年后，市净率低的股票组合表现明显较优。因此持有低市净率股票对于提高回报率仍具有一定的指导意义。

9 案例二总结

通过对沪深主板（以及中小板）股票数据的研究，本篇报告检验了三类有关高回报率市场传闻的真实性，构建出对应的时间序列模型，对股票的选取有一定的指导意义。

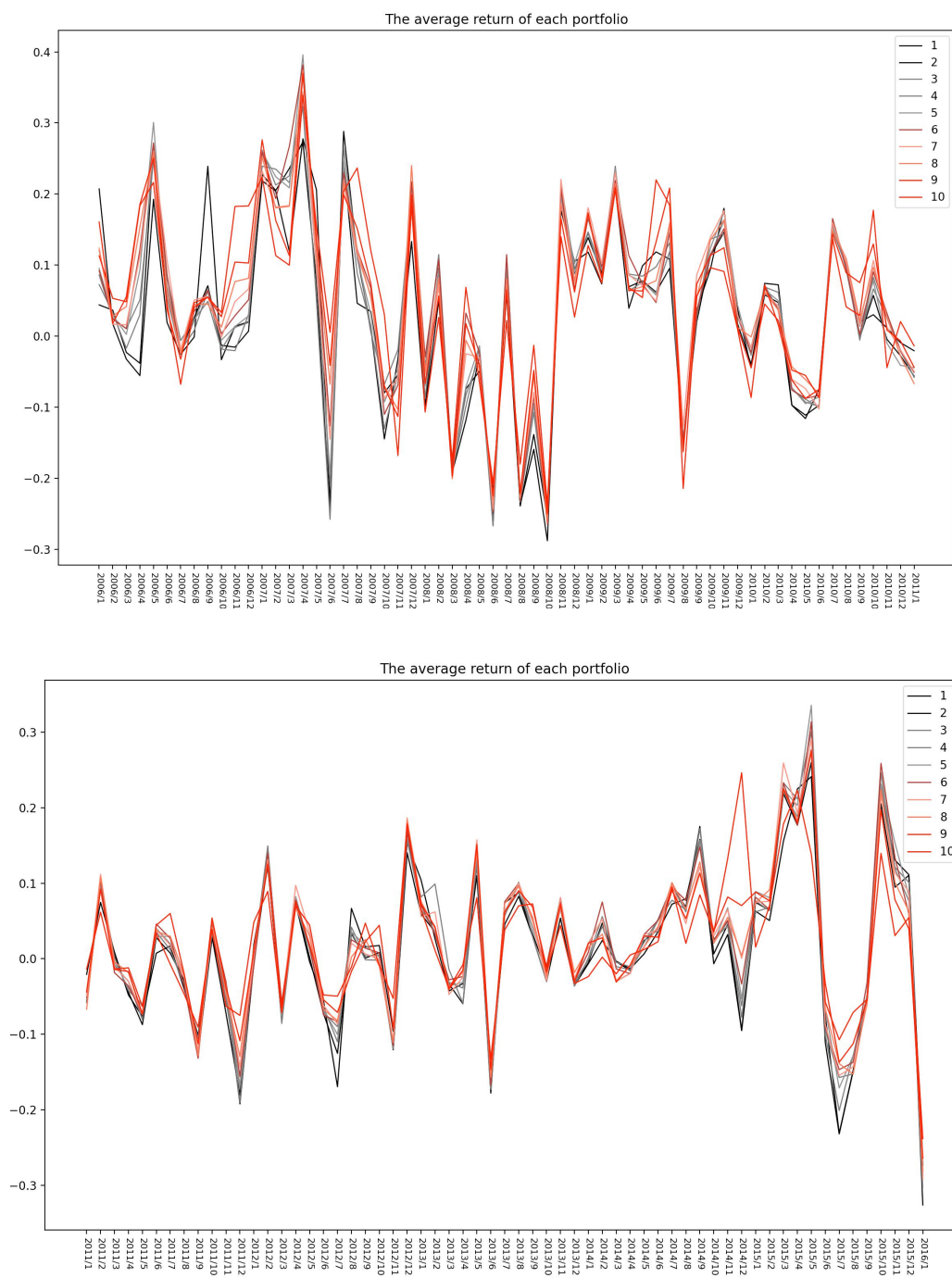
1. 短期时间内，高市值的股票更可能带来高回报率。这类股票价格较高，也更受广大投资者的青睐。而当持有时间加长，低市值股票的回报率表现趋优。当持有时间达到5年时，在多数情况下，低市值股票的期望回报率显著优于高市值股票的回报率。而近年来，无论持有长期或短期，高低市值股票的回报率表现势均力敌，并无明显优劣之分。在较远的将来是否可持有低市值股票进而获取高回报率，还将依赖于未来股票的数据以及相应的宏观经济研究结果。
2. 可视化结果表明“追涨杀跌”的策略极不可取。从长期来看，该策略为投资者产生的收益微乎其微。相反地，“追跌杀涨”在一些特定时期却有着令人惊异的收益表现。图表显示，“追涨杀跌”策略的稳定性远强于“追跌杀涨”，但其过低的收益率对投资者而言形同鸡肋。
3. 在2006到2012年之间，无论持有期长短，高低市净率组合回报率表现十分相近，并无明显的优劣之分。而可视化结果显示，在2012年后的多数时间，低市净率的股票组合的回报表现趋优。选择市净率小的股票并长期持有将可能成为较好的投资选择。

参考文献

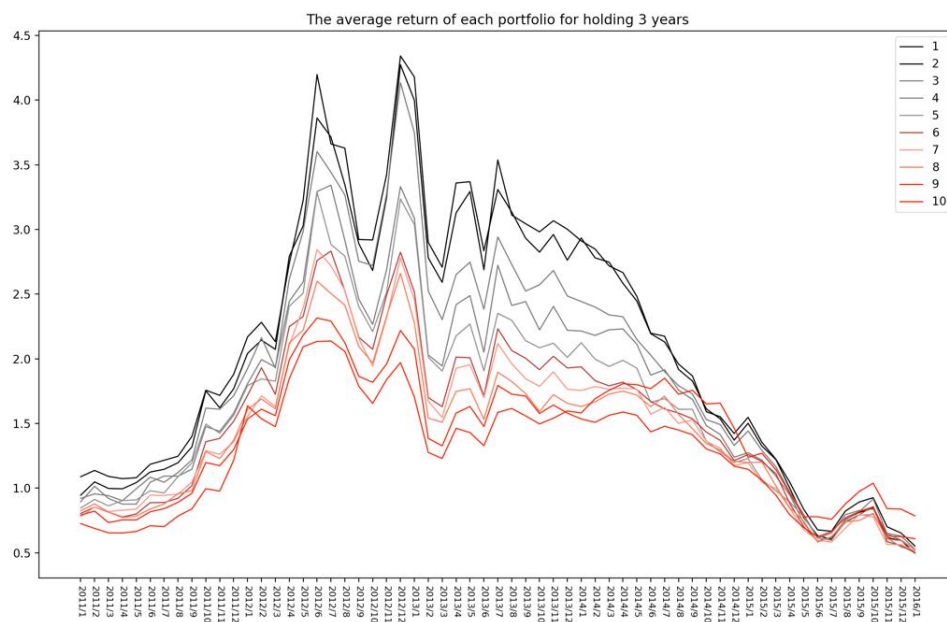
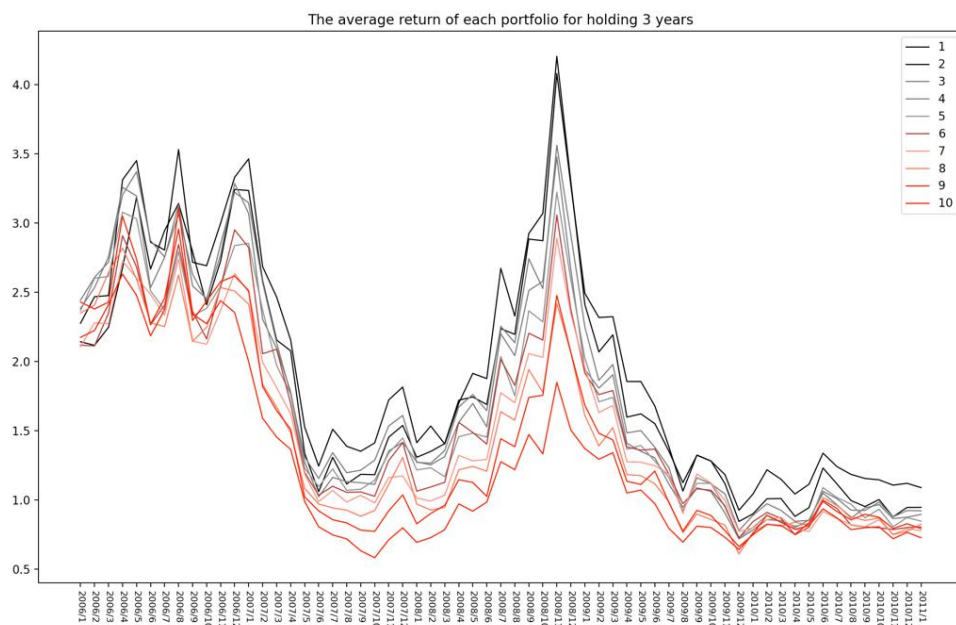
- Fama, E., & Macbeth, J. D. (1973). Risk, return and equilibrium: Empirical tests. *Journal of Political Economy*, 81, 607-636. <https://doi.org/10.1086/260061>
- Yifan, C., Jiayi, S., Wen, X., & Hui, J. (2019). Empirical test of CAPM in Shanghai securities market. *Journal of Finance*, 9(1), 28-33. <https://doi.org/10.12677/FIN.2019.91004>

附录

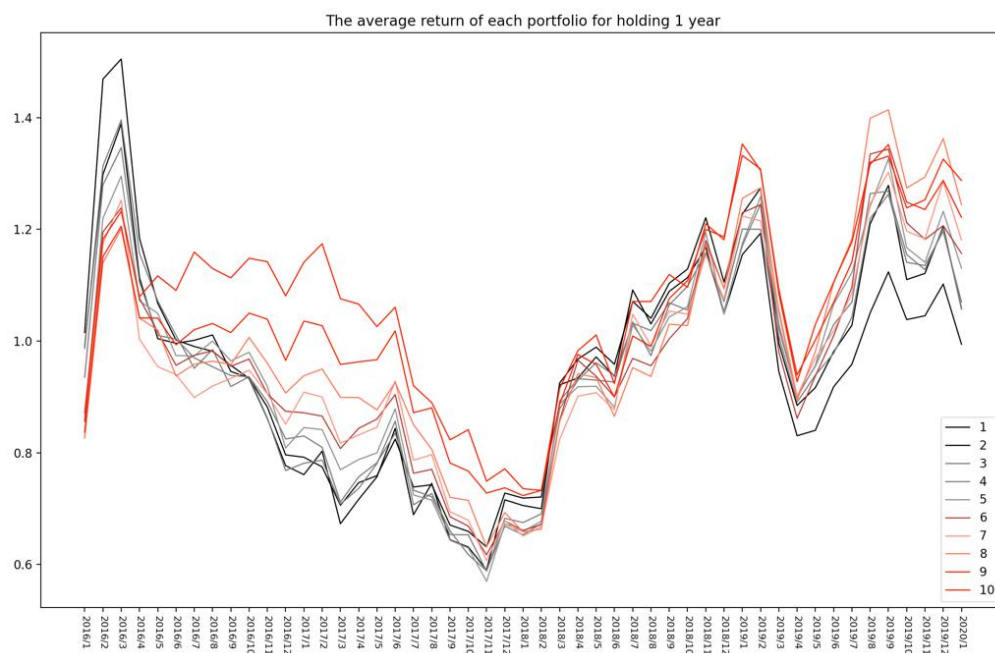
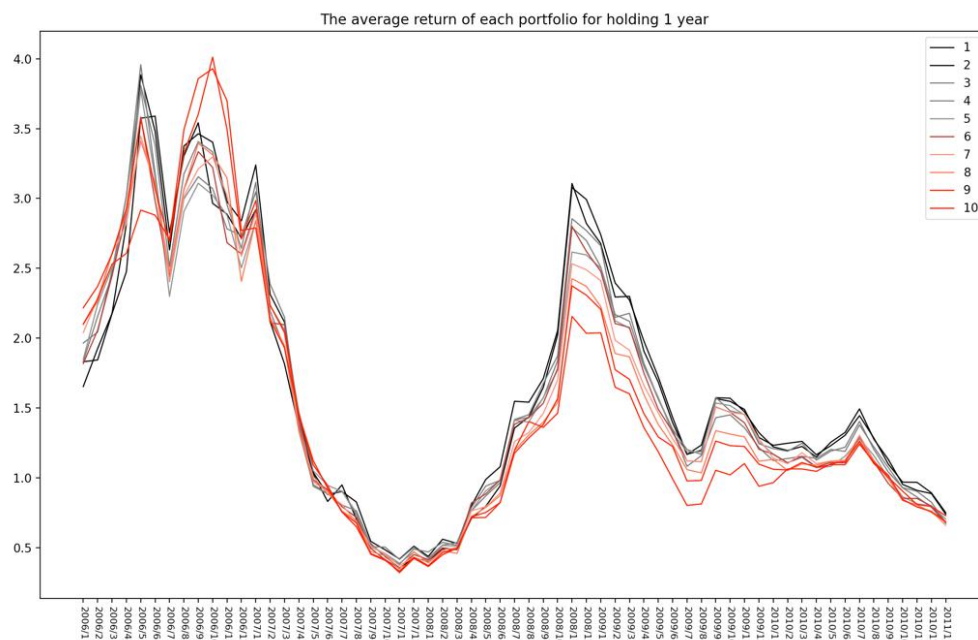
1. 依据市值持有一月的组合回报率（2006-2016）



2. 依据市值持有三年的组合回报率 (2006-2016)



3. 依据市值持有一年的组合回报率 (2006-2011; 2016-2020)



4. 依据上月回报持有一月后的组合回报率 (2006-2013)

