

FIN 3080

投资分析与投资组合管理

项目二

杨博宇 / 119020065

李东明 / 119020023

王崇锐 / 119020049

2021 年 4 月 12 日



香港中文大學(深圳)

The Chinese University of Hong Kong, Shenzhen

目录

案例一：中国股市价值因子复刻与相关性检验	3
1 背景简介：何为价值因子?	4
1.1 Fama-French 三因子模型	4
1.2 价值因子	4
2 复刻思路与原理	4
2.1 依据市值划分投资组合	5
2.2 依据账面市值比划分投资组合	5
2.3 计算价值因子	5
3 样本数据准备工作	5
3.1 基本数据	5
3.2 数据加工与筛选	6
4 因子复刻结果	6
4.1 组合划分结果	6
4.2 价值因子计算	7
4.3 可视化与相关性检验	7
5 案例一结论	8
案例二：中国股市“特质波动率之谜”探究	10
6 检验思路与原理	11
6.1 计算回归残差标准差	11
6.2 建立股票投资组合	12
6.3 构建市场组合的追踪组合	12
6.4 多空策略收益的显著性检验	13
7 样本数据清洗	13
8 探究结果	14
8.1 个股回归残差标准差结果	14
8.2 投资组合贝塔系数回归结果	14
8.3 追踪组合收益率可视化结果	15
8.4 多空策略收益显著性检验	18
9 案例二总结	18

案例一：中国股市价值因子复刻与相关性检验

摘要

作为 CAPM 模型的延伸, Fama-French (1993) 引入了市场因子以外的市值因子 (SMB) 与账面市值比因子 (HML), 以进一步解释投资组合的收益率。账面市值比因子, 又称价值因子, 描述了价值股在长期表现与增长股相比趋优的现象。

在国泰安数据库中, HML 因子的月度数据定义为高账面市值比组合与低账面市值比组合的月收益率之差。其中组合划分基于 FAMA2*3 组合划分。本文将采用同样的 P9715 市场类型 (综合 AB 股、创业板以及科创板), 以 2006 年 1 月至 2020 年 12 月的账面市值比 A 为基准进行组合划分。关于组合的划分标准, 同样采取在每年 7 月 1 日进行更新, 依据所有股票在前一年 6 月 30 日的市值与 12 月 31 日的账面市值比划分为 2*3 六个组合。最后计算 HML 月度因子序列, 并与国泰安数据库的 HML 因子进行相关性检验。

结果表明所复刻的价值因子与国泰安价值因子具有显著的相关性 (相关系数为 0.8576)。账面价值比计算方式 (是否考量递延所得税负债、递延所得税资产因素等) 不同是造成偏差的因素之一。此外, 划分投资组合时流通市值范围的选取, H、M、L 组合的更迭时间、缺失值的不同处理方法等均是造成偏差的可能因素。

关键词: Fama-French 三因子模型, 账面市值比, 市净率倒数, 相关性检验, 价值因子

1 背景简介：何为价值因子？

1.1 Fama-French 三因子模型

Fama-French 三因子模型，揭示了股票收益与市场因子、价值因子（HML）与市值因子（SMB）之间严格的线性定量关系。作为无套利（APT）理论下的定价模型，三因子模型有相比资本资产定价模型较少的模型假设，例如：

- 回报率与因子之间的关系可使用多元线性回归模型解释
- 个股的非系统性风险互不关联，投资组合可以有效地稀释所有个股的非系统性风险
- 市场的投资者通过行动消除所有的套利机会

在假设条件下，资产回报率与三因子之间的多元线性关系可以表达为：

$$\tilde{r}_{i,t} = \alpha_j + \beta_{j,m} \cdot \tilde{r}_{M,t} + \beta_{j,s} \cdot \tilde{r}_{SMB,t} + \beta_{j,h} \cdot \tilde{r}_{HML,t} + \varepsilon_{i,t} \quad (1)$$

其中， $\tilde{r}_{M,t}$ 为市场因子， $\tilde{r}_{SMB,t}$ 代表市值因子， $\tilde{r}_{HML,t}$ 代表价值因子， $\varepsilon_{i,t}$ 为回归残差。

1.2 价值因子

价值因子的计算定义为高账面市值比的组合与低账面市值比组合的回报率之差。其揭示了长期持有下价值股表现优于成长股的普遍现象。其中，账面市值比较高的股票便是价值股，其价值倾向于被低估，可能具有较高的投资价值。账面市值比低的股票是成长股，倾向于被高估。

值得注意的是，此处提到的价值股与成长股与市场热词里含义并不相同。在后者中，价值股更多指股价具有上涨潜力的股票，表现在重组预期、分红预期等等许多方面；而成长股则是具有行业红利或盈利能力不断提高的公司股票。

价值股与成长股之间并非一成不变，而是时刻动态联系的。股票的账面价值变化较慢，而市值却是每日波动的，股价的迅速波动便可能造成“价值”与“成长”股之间的来回切换。因此，计算价值因子时需定期更新构建的组合，以确保数据的有效性。

2 复刻思路与原理

根据国泰安数据库“Fama-French 使用指南”中的构造方法，价值因子的计算可分为以下步骤：

1. 根据每年 6 月 30 日的市值划分 S, B 投资组合。组合划分于每年 7 月 1 日更新。
2. 根据前一年的年终综合报表与去年年末的流通市值计算账面市值比（BE/ME），筛选负市净率股票，划分 H, M, L 组合。组合划分于每年 7 月 1 日更新。
3. 取 S, B 组合与 H, M, L 组合交集组合，通过加权计算股票市场的市值因子。

2.1 依据市值划分投资组合

计算价值因子的第一步是根据市值中位数划分两个投资组合。根据国泰安用户手册说明，采用 P9715 市场类型时，市值中位数应选择**主板市场**所有股票流通市值的中位数，而并非对应 AB 股、创业板以及科创板的全部股票对应的市值中位数。大于中位数的记为大市值公司（B），小于中位数的记为小市值公司（S）。市值选取时间为每年 6 月 30 日，组合划分于每年 7 月 1 日进行更新。

2.2 依据账面市值比划分投资组合

第二步是依据账面市值比（BE/ME）划分三个投资组合。使用前一年年终流通市值作为 ME 的度量，使用前一年年终报表中的“所有者权益+递延所得税负债-递延所得税资产-优先股账面值”作为 BE 的度量。

其次，剔除所有账面市值比小于 0 的股票。对 t 年 7 月至 t+1 年 6 月的因子计算，采用 t-1 年 12 月 31 日的账面市值比进行排序，取较小 30 分位的股票为 L 组合，30 至 70 分位的股票为 M 组合，较大 30 分位的股票为 H 组合。分位值的划分同样使用所有主板股票账面市值比的分位数。组合划分于每年 7 月 1 日进行更新。

2.3 计算价值因子

最后一步需将得到的组合划分取交集。依据划分，可将全体目标个股分为六个组合：

Table 1. Portfolio Sorting of Value Factor

	H (30%)	M (40%)	L (30%)
S (50%)	SH	SM	SL
B (50%)	BH	BM	BL

其中 SH、BH、SM、BM、SL、BL 代表不同划分的组合取交集后，经过加权的股票组合回报率。进而得到价值因子的计算公式：

$$HML_t = \frac{(SH_t + BH_t)}{2} - \frac{(SL_t + BL_t)}{2} \quad (2)$$

3 样本数据准备工作

3.1 基本数据

复刻使用综合 AB 股以及创业板和科创板的全部股票自 2006 年 1 月至 2020 年 12 月的月回报率数据。划分 S、B 组合时所用市值采用国泰安数据库中的市场流通市值数据。划

分 H、M、L 组合时所用的账面市值比则采用国泰安数据库中的“账面市值比 A”作为衡量标准。市值数据挑选所有个股在每年 6 月 31 日的流通市值，账面市值比数据则采用每年 12 月 31 日的定点数据。同时使用个股流通市值以计算划分后组合的加权月收益。所有数据均来自 CSMAR 数据库。

此后，以每年 7 月 1 日为基准，使用上一年 6 月 30 日的所有股票流通市值作为划分 S、B 组合的标准；使用上一年 12 月 31 日的个股市净率倒数作为划分 H、M、L 组合的标准。划分后取交集，产生的六个投资组合。计算 HML 因子时采用组合的加权收益率之差，其中加权标准为组合内个股的流通市值。

3.2 数据加工与筛选

仿照国泰安价值因子的构建，剔除所有账面市值比为负数的股票。剔除在每年 6 月 30 日的流通市值数据缺失的股票。计算 HML 因子数值时，月回报率缺失的股票将被剔除，不参与组合加权收益率的计算。

基础处理后，可得到全体样本个股自 2006 年 1 月至 2020 年 12 月的月回报率、流通市值以及组合归属情况（例如：股票 300352 属于 BL 组合），共约 453978 条总数据。经检验，其中有 66942 条个股存在收益率缺失或是组合归属情况缺失。其原因如下：

以股票 300264（佳创视讯）为例。其缺失 2011 年 7 月至 2012 年 6 月的 H、M、L 归属情况。根据国泰安的组合划分规则，佳创视讯在 2011 年 7 月至 2012 年 6 月中 HML 划分需依据 2010 年年底的其账面市值比值。而该股票在 2011 年 12 月 30 日上市，不具有 2010 年的账面市值比数值，因此该归属情况表现缺失。在划分节点间隔处停牌的股票同样会产生组合归属情况缺失。由此可见，对于每只 2006 年后上市的股票，其月度组合归属情况中必定存在缺失值现象，因为其前一年账面市值比值的缺失。因此，需将这类归属缺失的股票从当月的组合划分中剔除。

清洗处理后，仍有 387036 条有效股票数据。交集产生的六个投资组合仍有较大数目的个股组成，可以全部参与 HML 因子的权重计算。

4 因子复刻结果

4.1 组合划分结果

经过数据筛选，可进行自 2006 年至 2020 年的投资组合划分。组合于每年 7 月 1 日进行更迭。例如，对于 2006 年 1 月至 2006 年 7 月的投资组合，需使用 2005 年 6 月 30 日的流动市值与 2004 年 12 月 31 日的账面市值比 A 值进行划分。

以 2006 年 1 月至 7 月的投资组合划分结果为例，所得到的六个组合的个股数目：

Table 2. Number of Stocks in Sorted Portfolios for the Year 2006

	H (30%)	M (40%)	L (30%)
S (50%)	163	288	187
B (50%)	232	247	202

由表可知，筛除了缺失数据后，划分的六个投资组合仍有较可观的平均个股数目。下依据个股的流通市值进行加权计算，得到六个投资组合的月收益率。取 2006 年 1 月的六个投资组合的月收益率，展示如下：

Table 2. Returns for the Sorted Portfolios of 2006.1

	H (30%)	M (40%)	L (30%)
S (50%)	0.102742	0.082226	0.074130
B (50%)	0.109209	0.115968	0.105405

4.2 价值因子计算

得到每年更迭的六个投资组合后，通过月流通市值加权可计算得每月的投资组合收益率。利用 HML 因子计算公式 (2) 可计算得每月的 HML 因子数值。以 2006 年 1 月为例，利用上表 2，该月的 HML 因子数值便为：

$$HML_{2006,1} = \frac{0.102742 + 0.109209}{2} - \frac{0.074130 + 0.105405}{2} = 0.01623$$

依次计算每月对应组合划分下的 HML 因子数值，并将结果整理为时间序列。使用 describe 命令，可得到复刻的价值因子自 2006 年 1 月至 2020 年 12 月的统计分布：

Table 3. Descriptive Statistics about Calculated HML and HML from CSMAR

HML	count	mean	std	min	25%	50%	75%	max
HML_ours	180	-0.00578	0.03724	-0.19211	-0.02491	-0.00504	0.01582	0.17340
HML_CSMAR	180	-0.00267	0.03424	-0.14187	-0.01951	-0.00128	0.01539	0.15614

其中，“HML_ours”代表本文复刻价值因子得到的统计数据，“HML_CSMAR”代表国泰安数据库中的价值因子数据。

4.3 可视化与相关性检验

将所复刻的价值因子的时间序列与国泰安数据库的价值因子时间序列进行可视化对比，结果如下：

Figure 1. Time Series Comparison for Value Factors

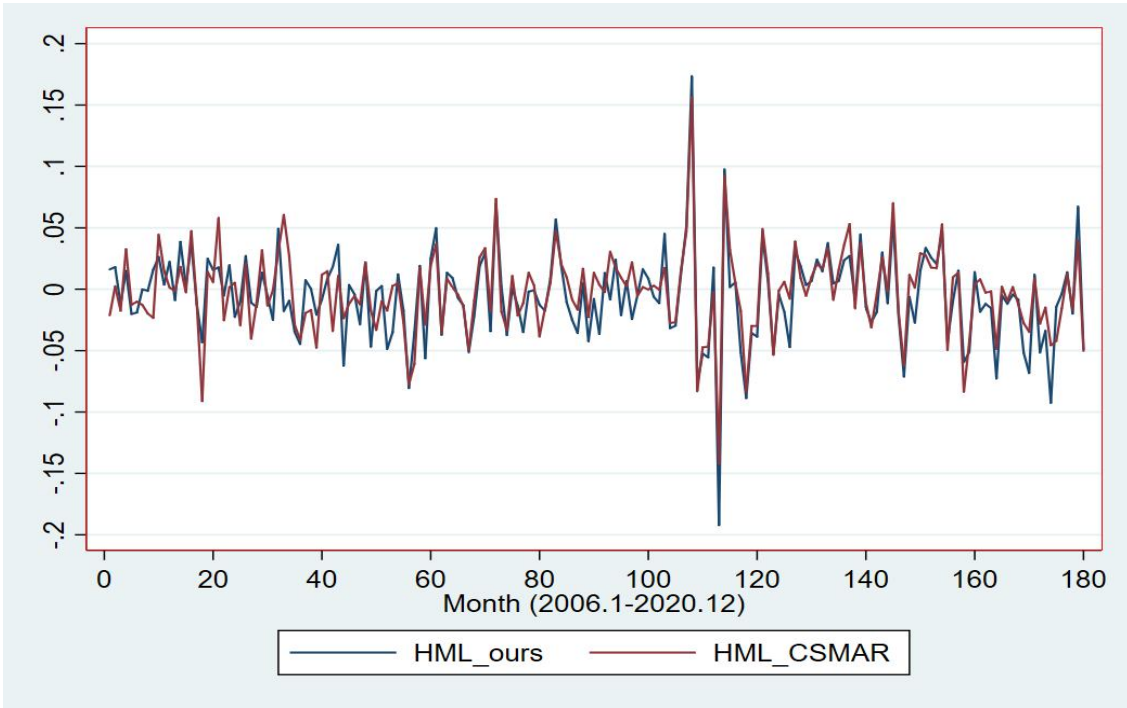


图 1 中，红线代表国泰安数据库的价值因子序列，蓝线代表本文复刻的价值因子时间序列。可以看出，二者拥有较强的相关性与高度相似的数值波动情况。下使用皮尔逊 (Pearson) 相关性检验计算相关系数，结果如表 2：

Figure 2. Pearson Correlation Coefficient Result

```
. pwcorr HML_ours HML_CSMAR
```

	HML_ours	HML_CS~R
HML_ours	1.0000	
HML_CSMAR	0.8576	1.0000

检验所得的相关系数约为 0.8576，接近于 1。说明两个价值因子序列拥有显著的相关性。本文复刻所得价值因子与数据库中的因子度量较为相符。

5 案例一结论

本文使用“国泰安操作指南”提出的价值因子构建方法复刻价值因子。即根据上一年 6 月 30 日的个股流通市值将样本个股（P9715 型市场，综合 AB 股、创业板与科创板）分为 S、B 组合；依据上一年 12 月 31 日的账面市值比 A 将样本个股分为 H、M、L 组合并取交集计算价值因子月度数值。

复刻结果与国泰安的价值因子相关性较强，数值波动相似，所得相关系数高达 0.8576。

说明构建的价值因子与数据库中较为吻合，可作为 Fama-French 三因子模型中价值因子的月度数据，对中国股票市场价值因子作用的显著性进行进一步的检验与考量。

同时，复刻所得价值因子仍并非与国泰安价值因子呈现完全线性关系。其中，账面市值比的选取（是否考量递延所得税负债、递延所得税资产因素等），以及流通市值的计算（国泰安数据库中市值 A、市值 B）均是影响组合划分与造成偏差的因素。此外，大量组合归属缺失值的不同处理方法也将影响二者之间的相关水平。

案例二：中国股市“特质波动率之谜”探究

摘要

在经典的因子定价模型中，股票收益只受到系统性风险的影响。在有效市场、投资者是理性、信息完全等假设下，股票的非系统性风险可通过构建投资组合得到充分的稀释与分散。然而，Ang et al. (2006) 借助 Fama-French (1993) 的三因子模型进行实证检验，提出特质波动率 (idiosyncratic volatility, 下称 IVOL) 与股票收益率存在显著的负相关性，与传统模型假设相悖。IVOL 对于中国股市股票回报率是否有显著影响同样存疑。

为探究特质波动率对中国股市股票回报率的影响，本文将以沪深主板以及创业板的全部股票自 2006 年以来的月回报率为基础，使用市场单因子模型残差的标准差衡量股票的特质波动率。在此基础上，依据个股特质波动率大小构建投资组合，并使用投资组合与无风险组合构建针对市场组合的追踪证券组合 (tracking portfolio)。最后，依据构建的多空组合 (long-short portfolio) 计算月回报率差值，对所得时间序列进行分析，检验该“套利收益”的显著性水平。

检验结果表明，通过多空策略所获收益多数在 10% 显著性水平下显著，部分在 5% 的显著性水平下显著。说明低特质波动率的投资组合平均拥有比高特质波动率投资组合更高的平均收益率，中国市场中特质波动率与平均回报率存在着较为显著的负相关性。

关键词：因子定价模型，特质波动率，残差标准差，月回报率，追踪证券组合，多空组合，统计显著性

6 检验思路与原理

Ang et al. (2006) 等采用 Fama-French (1993) 三因子模型进行时间序列回归以获得回归残差。本文仿照其研究方法，采用市场单因子模型进行特质波动率的衡量。具体检验过程如下：

1. 以每月为基准点，使用最近两年的月度回报率数据进行时间序列回归，得到所有目标个股的回归残差估计值 $\hat{\varepsilon}_i$ 与其标准差 $SE(\hat{\varepsilon}_i)$ 。
2. 使用 $SE(\hat{\varepsilon}_i)$ 衡量个股的特质波动性，依据波动性将所有个股排序并划分为投资组合。
3. 采用算数平均法计算投资组合超额收益，并以市场组合超额收益为自变量做时间序列回归，得到投资组合的贝塔系数回归结果 $\hat{\beta}_p$ 。
4. 依据两个投资组合贝塔系数回归结果，结合无风险组合并调整组合权重，构建针对市场风险组合的追踪组合（tracking portfolio），并计算追踪组合月回报率的差值。
5. 对每月数据均重复以上过程，得到每月回报率差值的时间序列，检验该“套利值”的显著性水平。

6.1 计算回归残差标准差

第一步是对个股进行时间序列回归，得到回归残差及其标准差。

其中，回归因变量为每只样本股在最近两年的月回报率，回归自变量为市场投资组合在对应时期的月回报率。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 月的收益率， $r_{m,t}$ 是市场投资组合在 t 月的收益率， α_i 为状态固定效益参数， ε_i 为回归误差。利用所得回归结果 $\hat{\alpha}_i$ 与 $\hat{\beta}_i$ ，计算残差估计值：

$$\hat{\varepsilon}_{i,t} = r_{i,t} - \hat{\alpha}_i - \hat{\beta}_i \cdot r_{m,t} \quad (4)$$

由于 $\hat{\varepsilon}_{i,t}$ 的表达式存在两个估计值，其标准差的计算对应丢失两个自由度：

$$SE(\hat{\varepsilon}_i) = \sqrt{\frac{1}{T-2} \sum_{t=1}^T \hat{\varepsilon}_{i,t}^2} \quad (5)$$

其中， T 为回归时间段的总月数。

6.2 建立股票投资组合

得到回归结果后，使用回归残差的标准差衡量特质波动率大小。依据特质波动率大小将个股划分为投资组合，有效地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：

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

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。得到所有投资组合的贝塔系数估计值 $\hat{\beta}_p$ ，并记录用于构建追踪组合。

6.3 构建市场组合的追踪组合

在经典的 APT 多因子模型中，资产回报率与因子之间的多元线性关系可表示为：

$$\tilde{r}_{i,t} = \alpha_i + \sum_{j=1}^K \beta_{i,j} \cdot \tilde{F}_{j,t} + \tilde{\varepsilon}_{i,t} \quad (7)$$

其中 K 为因子总数。对式 (7) 两边取期望，结合 APT 模型的假设条件，可得：

$$E(\tilde{r}_i) = \alpha_i = r_f + \sum_{j=1}^K \beta_{i,j} (E(\tilde{F}_i) - r_f) \quad (8)$$

在 6.2 中利用得到的五个投资组合贝塔系数回归结果，若不处理直接选取两个组合实行多空策略，则所获收益的表达式为：

$$\Delta R = E(\tilde{r}_L) - E(\tilde{r}_S) = \sum_{j=1}^K E(\tilde{F}_i) \cdot (\hat{\beta}_{L,j} - \hat{\beta}_{S,j}) \quad (9)$$

其中， \tilde{r}_L 与 \tilde{r}_S 分别表示做多与做空的投资组合。由于多空组合的贝塔系数回归结果 $\hat{\beta}_{L,j}$ 与 $\hat{\beta}_{S,j}$ 并不相同，若得出超额收益 ΔR 在时间序列上较为显著，则该显著性可能由特质波动率的影响造成，也可能由组合不同的系统性风险差 $\hat{\beta}_{L,j} - \hat{\beta}_{S,j}$ 所造成。

因此，为控制变量，这一步需将所得到的五个组合分别与无风险组合结合，调整权重，重新构建特定因子载荷 (factor loading) 的追踪组合。本文选取市场组合 (市场因子的载荷为 1，其余因子的载荷为 0) 作为追踪组合的目标。

经计算，追踪组合的权重调整公式如下：

$$\tilde{r}_{track} = \frac{1}{\hat{\beta}_p} \cdot \tilde{r}_p + (1 - \frac{1}{\hat{\beta}_p}) \cdot \tilde{r}_f \quad (10)$$

\tilde{r}_{track} 代表追踪组合， \tilde{r}_f 代表无风险组合， \tilde{r}_p 代表上一步得到的任意投资组合。该权重分配使得构建的追踪组合拥有与市场组合完全相同的因子载荷 (factor loading)。追踪组合

控制了系统性风险的变量影响，确保下一步实行多空策略时所获超额收益均来自组合特质波动率的影响。

6.4 多空策略收益的显著性检验

最后一步需对所得的追踪组合实行多空策略。记 $\tilde{r}_{track,1}$ 与 $\tilde{r}_{track,0}$ 别为特质波动率相异的组合所构建的追踪组合。则该月采用多空策略获取的收益为：

$$\tilde{R}_t = \tilde{r}_{track,1} - \tilde{r}_{track,0} \quad (11)$$

本文的关键策略在于每月重复以上流程并更新不同特质收益率的组合划分。事实上，由于个股的特质波动率会随时间而变化，仅使用单一时间的历史数据划分所得的特质波动率组合有失偏颇。因此本文选择**每月更新**特质波动率组合的划分。相应地每月构造追踪组合，采用多空策略记录超额收益。

最后对不同追踪组合的回报率进行可视化展现，描绘超额收益的时间序列图，并检验超额收益的统计显著性。

7 样本数据清洗

采用沪深主板及创业板所有个股自 2006 年以来的月回报率进行检验。其中月个股回报率的计算不考虑现金红利再投资。无风险利率 $r_{f,t}$ 采用 CSMAR 数据库中的月度化无风险利率。

缺失回报率数据的股票将在进行市场模型回归时被剔除，不参与到投资组合的构建之中。构建股票组合时，采用个股的月收益率平均值作为组合的月回报率。

本文实证方法的关键在于时间点的更迭。自 2008 年 1 月起，每月均利用过去两年的月回报率数据进行时间序列回归，得到个股的特质波动率并分类进行投资组合构造。因此，从 2008 年 1 月至 2020 年 12 月，每月均需对最近两年的股票数据进行清洗，筛出收益率缺失的个股。经计算，纳入考量的 156 个月份中，筛选后平均剩余的股票均大于 1000 只，且数量在逐年增长。因此可以确保依据特质波动率划分投资组合时，每个组合均有足够的股票数目以分散非系统性风险。

8 探究结果

8.1 个股回归残差标准差结果

经过数据筛选，使用市场回归模型对最近两年收益率进行时间序列回归。得到回归残差以及其标准差。以 2008 年 1 月至 2008 年 10 月为例，所得全部个股的残差标准差的统计分布情况如下：

Table 4. Descriptive Statistics of Idiosyncratic Volatility

Date	Count	Mean	Std. Dev.	Min	25%	50%	75%	Max
Jan-08	1009	0.16248	0.04356	0.06051	0.13128	0.15727	0.18535	0.38759
Feb-08	1019	0.16127	0.04291	0.05917	0.13049	0.15509	0.18582	0.39018
Mar-08	1029	0.16180	0.04290	0.05637	0.13187	0.15612	0.18549	0.40491
Apr-08	1040	0.16433	0.04372	0.06448	0.13374	0.15951	0.18866	0.40544
May-08	1051	0.16144	0.04355	0.06390	0.13150	0.15556	0.18631	0.40431
Jun-08	1073	0.15964	0.04166	0.06260	0.13077	0.15445	0.18329	0.37861
Jul-08	1088	0.15993	0.04444	0.06140	0.13155	0.15470	0.18315	0.69123
Aug-08	1099	0.16082	0.04437	0.06029	0.13174	0.15561	0.18383	0.69249
Sep-08	1130	0.16140	0.04401	0.06074	0.13257	0.15606	0.18469	0.69127
Oct-08	1140	0.16026	0.04295	0.06703	0.13207	0.15498	0.18303	0.69074

由表可知，筛选后每阶段均仍有较大的样本数量，且全体个股的百分位数分布较为接近，可确保依据特质波动率划分的投资组合能够充分分散个股的非系统性风险。

8.2 投资组合贝塔系数回归结果

得到个股的回归结果后，下依据所有个股特质波动率大小划分投资组合。根据样本数量大小，每月将样本个股等分为五个投资组合。其中最小二十分位的特质波动率股票构成组合 0，二十到四十分位的特质波动率股票构成组合 1，以此类推至组合 4。

以股票的算数平均收益率作为组合收益率，再次使用两年的月回报率数据进行时间序列回归，取 2008 年 1 月至 2008 年 10 月的五个组合的贝塔系数回归结果如下：

Table 5. Descriptive Statistics of Portfolio Beta Results

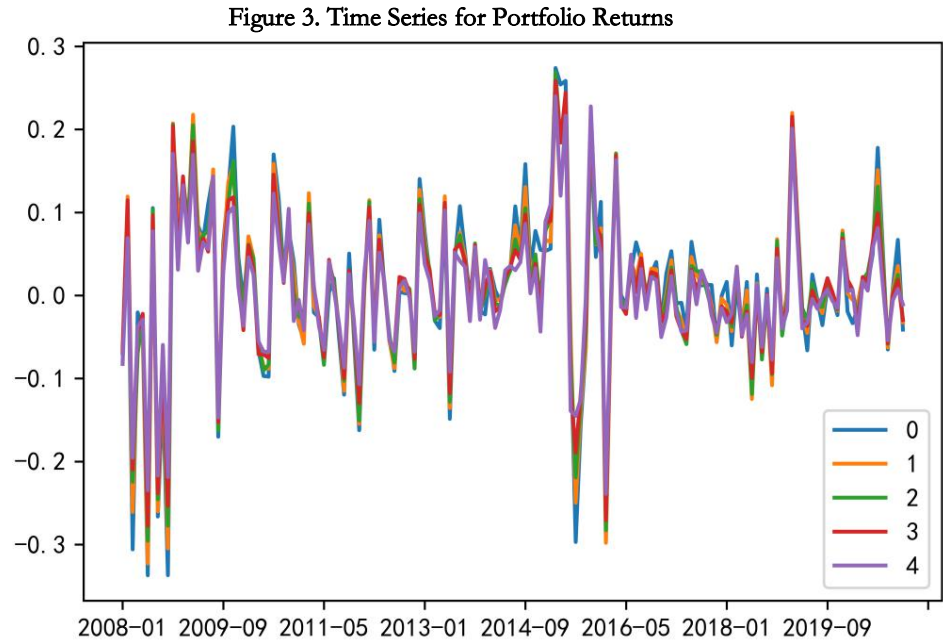
Date	P0	P1	P2	P3	P4
Jan-08	0.80527	0.87478	0.88779	0.99844	1.05971
Feb-08	0.62505	0.75122	0.84009	0.95529	1.10175

Mar-08	0.60314	0.73410	0.82843	0.92341	1.08585
Apr-08	0.66843	0.79090	0.87172	0.95588	1.13400
May-08	0.67880	0.79831	0.87710	0.96260	1.11324
Jun-08	0.63816	0.76942	0.82799	0.91898	1.08379
Jul-08	0.68298	0.82248	0.90959	0.99535	1.15676
Aug-08	0.69153	0.81618	0.90139	0.99468	1.14771
Sep-08	0.72514	0.86051	0.94627	1.04438	1.20775
Oct-08	0.73082	0.86772	0.94582	1.05646	1.21454

其中，P0、P5 分别为特质波动率最小、最大的投资组合。得到所有投资组合的贝塔系数回归结果后，下分别利用无风险组合调整权重，构建市场组合的追踪组合。

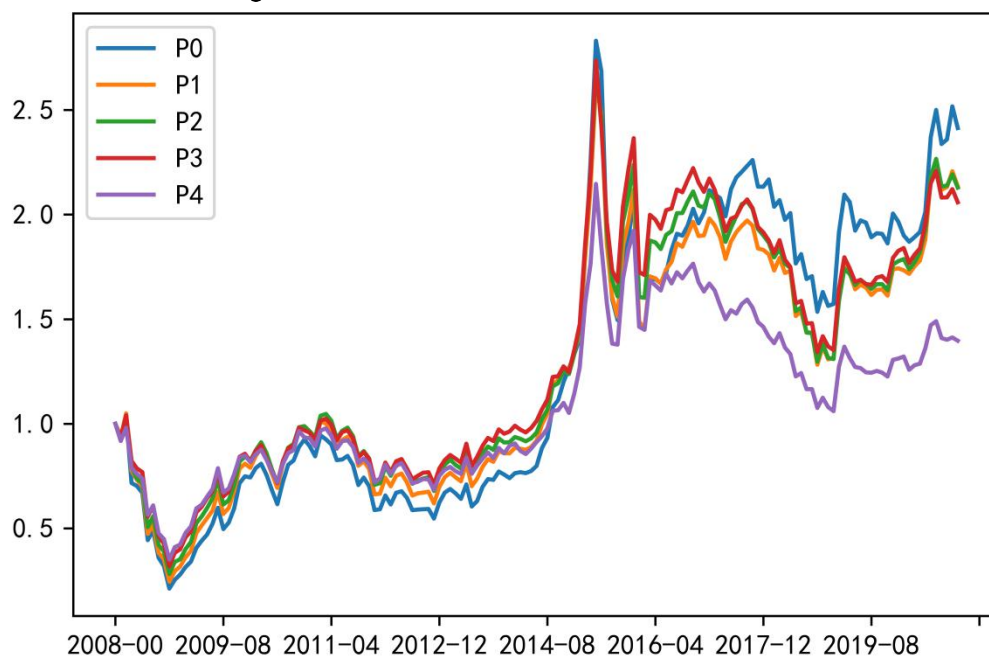
8.3 追踪组合收益率可视化结果

下将所得的五个投资组合的时间序列进行可视化展现。其中，横坐标为时间序列，纵坐标记录五个投资组合自 2008 年 1 月以来的月收益波动。结果如下：



可以看出，五个组合的基本波动趋势较为一致，但收益率震荡幅度仍有较大差异。下保持横坐标不变，纵坐标改为分别持有五个投资组合，自 2008 年 1 月以来可以得到的累积月收益率：

Figure 4. Time Series for Cumulative Portfolio Returns



由图可知，不同特质波动率的五个投资组合累计收益率在 2008 年至 2014 年波动接近，但自 2015 年起，五个组合的累计收益率差异逐渐变大。紫线代表的高特质波动率组合 4 在 2016 年后的累计收益率显著低于其他波动率的组合。相反地，蓝线代表的低特质波动率组合 0 在 2016 年后的累计收益率显著高于其他波动率组合。

据图可提出猜想，低特质波动率的组合累积收益率平均高于高特质波动率组合。下挑选不同的两个特质波动率组合，每月实行多空策略并记录超额收益，画出其时间序列图。以对组合 4 做多、组合 3 做空，组合 5 做多、组合 1 做空的超额收益率为例，其时间序列图展示如下：

Figure 5. Excess Return Time Series (L4-S3)

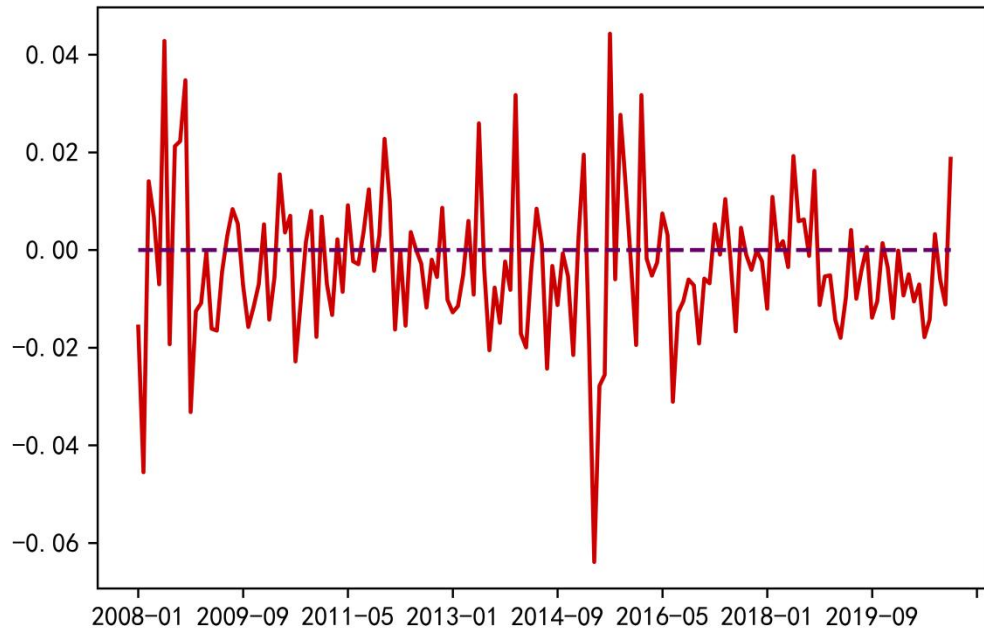
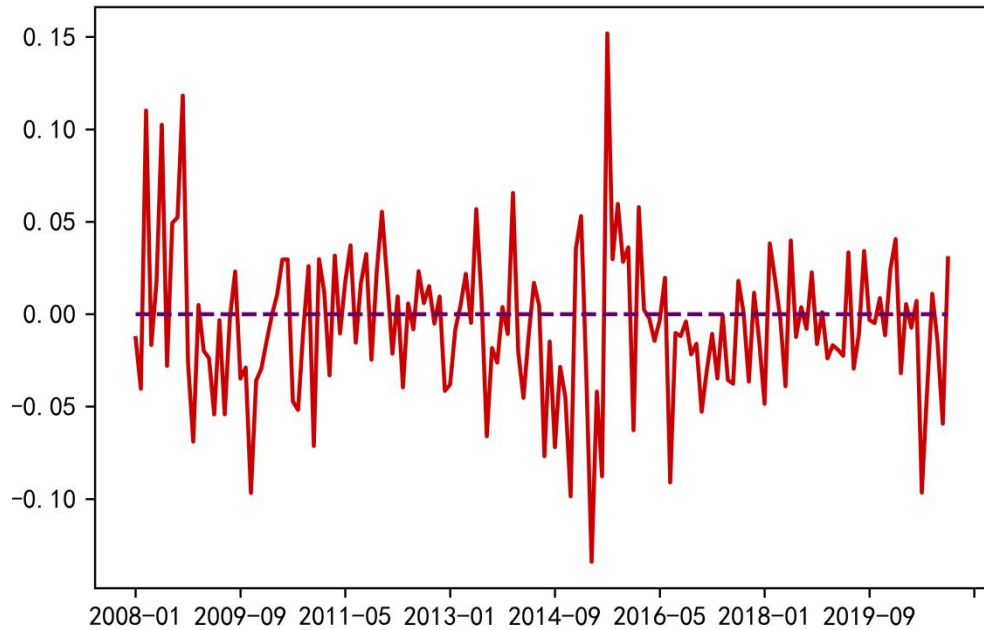


Figure 6. Excess Return Time Series (L5-S0)



可以看出，超额收益的波动较大，且多数时间为负值，与前面特质波动率与收益率呈现负相关的假设吻合。下对该猜想实行严格的统计检验。

8.4 多空策略收益显著性检验

现对不同组合使用多空策略并进行显著性 t 检验。基础假设为超额收益的回报均值为 0。若股票的特质波动率不影响组合收益率，则基础假设成立。下图展现出（部分）对五个组合中不同组合实行多空策略的超额收益率检验结果：

Table 6. T-test Results for Checking Volatility Significance

Results	L4-S0	L4-S1	L4-S2	L4-S3
T-statistics	-1.843	-1.933	-2.235	-2.762
P value	0.067	0.055	0.027	0.0064

其中，“L4-S0”代表对组合 4 实行做多，组合 0 实行做空的多空策略，其余同理。对所有可能的组合使用多空策略并进行 t 检验后，结果表明，多数超额收益的 P 值小于 0.1，部分的超额收益在 5% 的显著性水平下显著 ($P < 0.05$)，验证了低特质波动率组合平均收益高于高特质波动率组合的猜想。同时所得 T 检验值均为负值，说明特质波动率与组合收益之间存在较为显著的负相关性。

9 案例二总结

通过对沪深主板以及科创板股票数据的分析，本篇报告探究了中国股市“特质波动性之谜”、揭示了投资组合特质波动性与组合收益率的定量关系、并通过多空策略计算自 2008 年来每月的超额收益值，检验该“套利”收益的显著性。

结果表明，特质波动率最低的投资组合拥有比特质回报率最高的投资组合更优的平均回报，说明特质回报率对中国股市的平均回报率具有一定影响。部分特质波动率组合的收益差值所得检验 P 值小于 0.05，说明该影响具有一定的统计显著性。

因此，投资者需谨慎考量非系统性风险对投资组合收益率的影响。中国证券市场起步较晚，证券市场信息公开程度较低；中小型投资者居多，投资理念也尚不成熟。而市场频繁的震荡与不稳定性也是导致特质波动率的主要原因之一。使用传统经典因子模型 (CAPM, APT) 时，投资者需谨慎考量特质波动率的影响，检验模型假设是否成立。

参考文献

- Gu, M., George, J. J., & Xu, B. (2019). The role of analysts: An examination of the idiosyncratic volatility anomaly in the Chinese stock market. *Journal of Empirical Finance*, 52, 237-254. <https://doi.org/10.1016/j.jempfin.2019.03.007>
- Gu, M., Kang, W., & Xu, B. (2018). Limits of arbitrage and idiosyncratic volatility: Evidence from China stock market. *Journal of Banking & Finance*, 86, 240-258. <https://doi.org/10.1016/j.jbankfin.2015.08.016>