

# 中国股票市场收益率的可预测性研究<sup>①</sup>

蒋志强, 田婧雯, 周炜星\*

(华东理工大学商学院金融学系, 上海 200237)

摘要: 构造了包括中国 A 股市场组合、行业组合、账面市值比组合和市值组合在内的 31 个投资组合, 选取了 8 个预测因子(账面市值比、股利分配率、股息价格比、股息收益率、每股收益价格比、现金收益价格比、通货膨胀率、股票波动率), 运用了可行拟广义最小二乘法对各因子对各投资组合收益率的可预测性进行了样本内和样本外检验. 研究发现: 1) 中国股市收益率是可预测的, 但是各投资组合收益率的可预测性在样本内、样本外、熊市和牛市均存在差异; 2) 绝大部分投资组合收益率的可预测性可由条件 CAPM 模型捕捉的时变系统风险溢价充分解释; 3) 行业投资组合收益率的可预测性与行业集中度显著负相关, 可由信息摩擦理论解释.

关键词: 可行拟广义最小二乘法; 样本内预测; 样本外预测; 条件 CAPM; 行业集中度

中图分类号: F830.91 文献标识码: A 文章编号: 1007-9807(2019)04-0092-18

## 0 引言

近几年中国的资本市场大起大落, 特别是 2016 年 A 股市场剧烈震荡, 出现了股灾 2.0 版、千股再度跌停等问题, 两次熔断一度动摇了市场的信心, 大盘长期浮动在 3 000 点上下, 在收益整体低迷的情况下, 量化投资却逆势突围、大放异彩, 凭借亮眼的业绩成为投资领域的新宠. 2016 年各大量化策略平台纷纷崛起, 公募量化基金和私募量化基金规模大幅度提升, 全年有 44 只新基金相继成立.

根据上海证券报报道<sup>②</sup>, 截至 2017 年初, 万德数据显示: 在上证综指和创业板指近 1 年涨幅只有 3.40% 和 -8.37% 的情况下, 公募主动量化基金获得了 16.05% 的平均收益, 近 3 年平均回报率为 105.07%, 收益率接近同期权益类基金的两倍.

我国量化投资的迅猛发展离不开股市非完全有效的前提. 众多的研究已发现我国股市收益率

存在经典线性相关之外的非线性相关, 不满足随机游走的假设. 股价的波动复杂表面的背后, 存在可预测成分. 因此, 本文在中国量化投资的发展前提下, 提出其理论背景的研究, 即中国股市收益率的可预测性问题.

## 1 国内外研究现状

从有效市场假说到 Markowitz 均值-方差理论、从 Tobin 两基金分离定理到 Sharpe 等人在 20 世纪 60 年代初提出的资本资产定价模型 (CAPM), 现代投资理论的探索由规范理论转入实证研究, 资产收益率的可预测性也经历了从无至有的转变. CAPM 模型认为股票的期望收益率是由系统风险  $\beta$  系数唯一决定的, 而后越来越多的定量研究发现, 传统定价理论很难充分揭示资产价格的变动. 除了  $\beta$  外还有许多基础因子能对资产收益率进行解释. 从 1992 年 Fama 和 French

① 收稿日期: 2017-05-01; 修订日期: 2018-06-02.

基金项目: 国家自然科学基金资助项目 (U1811462; 71532009; 71571121); 中央高校基本科研业务费资助项目 (222201718006).

通讯作者: 周炜星 (1974—), 男, 浙江诸暨人, 博士, 教授. Email: wxzhou@ecust.edu.cn

② 报道链接: <http://news.cnstock.com/paper,2017-01-23,774836.htm>

提出的三因子模型<sup>[1]</sup>到2015年改进的五因子模型<sup>[2]</sup>,因子的探索之路从未停歇。

关于因子的研究主要可分为如下三类:1) 基本面因子以个股财务指标为主,主要包括股息价格比、股票规模、财务杠杆、账面市值比、市盈率、股息率、市净率、投资资本、企业支付和融资活动等<sup>[3-17]</sup>。2) 市场及宏观经济因子包括股市波动率、通货膨胀率、长短期利率差和优质债券与劣质债券的收益差、消费财富比率等<sup>[18-21]</sup>。3) 技术分析因子的研究主要以检验交易策略的有效性为主,方法更为多元化,包括统计工具、计量方法、理论建模等<sup>[22-23]</sup>。涉及的因子包括相对强弱指标RSI、随机指标KD、趋向指标DMI、平滑异同平均线MACD、能量潮OBV等<sup>[24-26]</sup>。

前两类研究的分析方法主要为最小二乘回归法,其中最常用的是Fama-MacBech方法。Fama-MacBech方法分别取单因子或多因子,根据两阶段回归结果取回归系数代数平均,重新计算系数的标准差及相应的 $t$ 统计量,即可得到判断因子显著性的指标<sup>[1]</sup>。第三类研究因其检验方式的多样性和因子的特殊性,使得技术有效性分析成为一个单独的研究领域。因此,本文的研究仅涉及前两类因子。

对于中国股票市场,关于基本面、市场及宏观经济因子的研究尚存在如下特点:

1) 实证检验主要为定价模型的直接检验或特定因子的检验,但局限于样本内检验<sup>[27-28]</sup>。

2) 受样本区间选择、样本容量、因子计算差异等影响,相同因子的预测性存在研究结果不一致的问题,如市盈率因子<sup>[29-30]</sup>、市净率因子<sup>[31-32]</sup>。

3) 绝大多数研究直接运用普通最小二乘法(OLS)进行收益率预测性分析<sup>[33-34]</sup>。然而,因子变量(特别是估计因子)往往具有高度持续性、异方差性、内生性问题,这些特性容易导致回归系数的偏误,使传统的 $t$ 统计量不再有效<sup>[35-37]</sup>,说明研究的计量方法还存在改进的空间。

4) 只给出了因子的经济学解释,缺少对投资组合预测性的解释,未给出投资组合预测性不同的原因<sup>[38-39]</sup>。

因此,本文将针对现有研究的不足,通过构建投资组合,对比分析了样本内外预测性的表现、不同经济周期下预测性的差异,实证分析较为系统

和全面。

在研究方法上,已存在的Newey和West提出的方法使存在异方差和自相关的情况下仍能对方差矩阵进行一致性估计,但根据研究表明<sup>[40-41]</sup>,HAC法本身存在某些缺陷,导致经HAC法修正的 $t$ 统计量仍然存在伪回归,只有在样本容量趋于 $+\infty$ 时,HAC法才能获得长期方差一致的估计量,且即使变量个数、样本容量增加,统计量仍具有较高伪回归的概率。

相比之下,可行拟广义最小二乘法不仅可以解决异方差、自相关,还可以消除内生性因素、因子变量高度持续性问题带来的影响,计算过程也较为简便,而且目前没有研究显示该方法对样本容量有较为苛刻的要求。因此本文基于模型改进的角度利用Westerlund和Narayan<sup>[42-43]</sup>提出的可行拟广义最小二乘法(FQGLS)进行中国股市收益率的可预测性研究,以克服OLS方法的缺陷。另外,还将因子均值的检验结果拆分成理性条件CAPM模型可预测性和Alpha可预测性,并分析了行业集中度是否能解释行业样本外预测性的差异,最终给出投资组合可预测性的经济学解释。

## 2 检验方法

### 2.1 样本内检验

参考Westerlund和Narayan<sup>[42-43]</sup>进行FQGLS方法的推导。

通常传统收益率预测的回归模型为

$$r_t = \alpha + \beta x_{t-1} + \varepsilon_{r,t} \quad (1)$$

其中 $r_t$ 代表超额收益率; $x_{t-1}$ 代表预测因子; $\alpha$ 代表截距项; $\beta$ 代表预测因子对收益率的影响系数,若 $\beta=0$ ,即因子对收益率无预测性; $\varepsilon_{r,t}$ 为残差项。需要说明的是,预测因子的构造方法不同于FF三因子模型。三因子模型的因子构造采取分组取差值形式,本文直接运用因子值,具体计算方式见第三节。

Lewellen<sup>[44]</sup>已经证明对式(1)直接估计存在偏误。因此,本文采用FQGLS方法克服数据本身的持续性、异方差和内生性问题,具体实施步骤如下:

假设式(1)中预测因子遵循一阶自回归过程 (AR(1)) ,即

$$x_t = \mu(1 - \rho) + \rho x_{t-1} + \varepsilon_{x,t} \quad (2)$$

其中  $\mu$  为截距项中的参数  $\rho$  为预测因子与其滞后一阶的相关系数,  $\varepsilon_{x,t}$  为残差项. 根据 Westerglund 和 Narayan<sup>[42,43]</sup>, 假设式(1)、式(2)的残差项负相关是合理的. 例如, 当  $x_{t-1}$  表示股息率, 那么股价上升会降低股息, 使得收益率提高.

由于内生性问题可导致式(1)的估计有偏, FQGLS 中将内生性形式定义为

$$\varepsilon_{x,t} = \gamma \varepsilon_{x,t} + \eta_t \quad (3)$$

将式(2)、式(3)代入式(1)的残差项, 可得

$$r_t = \theta + \beta x_{t-1} + \gamma(x_t - \rho x_{t-1}) + \eta_t \quad (4)$$

参考 Westerglund 和 Narayan<sup>[42,43]</sup>, ARCH 模型是学术界常用的描述异方差形式的模型. 因此 FQGLS 中将异方差性定义为

$$var(\eta_t | I_{t-1}) = \sigma_{\eta_t}^2 = \eta_0 + \sum_{k=1}^q \psi_k \eta_{t-k}^2 \quad (5)$$

其中  $I_t$  表示时刻  $t$  的信息. 式(5)需同时满足如下假设

$$\begin{aligned} \psi_0 > 0 \quad \psi_1, \dots, \psi_q \geq 0, \\ \sum_{k=1}^q \psi_k < 1 \quad par(\varepsilon_{x,t} | I_{t-1}) = \sigma_{\varepsilon_{x,t}}^2 \end{aligned} \quad (6)$$

进而得到  $t$  统计量

$$t_{FQGLS} = \frac{\sum_{t=q_m+2}^T \pi_t^2 x_{t-1}^d r_t^d}{\sqrt{\sum_{t=q_m+2}^T \pi_t^2 (x_{t-1}^d)^2}} \quad (7)$$

其中  $x_t^d = x_t - \sum_{s=2}^T x_s / T$ ,  $r_t^d = r_t - \sum_{s=2}^T r_s / T$ ,  $\pi_t = 1 / \sigma_{\eta_t}$ . 使用该形式的目的在于处理残差项中可能存在的 ARCH 结构. 特别地,  $\pi_t$  的计算可参考如下步骤:

1) 对式(4)直接进行 OLS 回归, 得到  $\eta_t$  的拟合值  $\hat{\eta}_t$ ;

2) 将拟合值  $\hat{\eta}_t$  的平方对其滞后项进行 OLS 回归

$$\hat{\eta}_t^2 = \lambda_0 + \sum_{j=1}^q \lambda_j \hat{\eta}_{t-j}^2 \quad (8)$$

回归至滞后  $q$  阶的系数显著. 此  $q$  即为式(7)中的  $q_m$ ;

3) 式(8)的拟合值即为  $\sigma_{\eta_t}^2$ , 进而计算  $\pi_t$ .

### 2.2 样本外检验

参考 Rapach 等<sup>[45]</sup> 和 Rapach 等<sup>[46]</sup> 本文利用扩张移动窗口分析法进行样本外检验. 假设样本长度共  $2t$  期, 将整个样本分为估计样本(前  $t$  期)和预测样本(后  $t$  期). 利用最小二乘法对前  $t$  期估计样本估算模型参数, 进而预测  $t+1$  期的收益率, 不断迭代直至计算出整个预测样本. 参考 Campbell 和 Thompson<sup>[47]</sup> 本文利用相对平均绝对误差、相对均方根误差和样本外  $R^2$  作为样本外预测效果的指标, 具体计算方法如下:

1) 相对平均绝对误差 表示为  $RMAE$

$$RMAE = \frac{\sum_{t+1}^{2t} |r_t - \hat{r}_t| / t}{\sum_{t+1}^{2t} |r_t - \bar{r}_t| / t} \quad (9)$$

2) 相对均方根误差 表示为  $RRMSE$

$$RRMSE = \frac{\sqrt{\sum_{t+1}^{2t} (r_t - \hat{r}_t)^2 / t}}{\sqrt{\sum_{t+1}^{2t} (r_t - \bar{r}_t)^2 / t}} \quad (10)$$

3) 样本外  $R^2$  表示为  $OR^2$

$$OR^2 = 1 - \frac{\sum_{i=t+1}^{2t} (r_i - \hat{r}_i)^2}{\sum_{i=t+1}^{2t} (r_i - \bar{r}_i)^2} \quad (11)$$

式(9)~式(11)中,  $\bar{r}_i$  为历史平均收益, 表示用第  $i$  期之前所有收益率的均值预测第  $i$  期的收益率.  $\hat{r}_i$  为因子模型式(1)的预测结果. 当  $RMAE$ 、 $RRMSE$  小于 1,  $OR^2$  大于 0, 表示选取的因子预测效果优于历史均值模型, 该因子样本外预测能力较好. 本文还根据 Rapach 等<sup>[45]</sup>, 取所有因子预测收益率的平均值检验了综合预测收益率的预测效果.

### 2.3 样本外预测性的分解

参考姜富伟等<sup>[48]</sup> 使用条件 CAPM 资产定价模型把样本外统计量  $OR^2$  分解成基于条件资产定价模型的可预测性  $OR_R^2$  (即理性可预测性) 和 Alpha 可预测性  $OR_\alpha^2$ . 其中, 理性可预测性源于暴露的时变宏观经济风险的溢价, 而 Alpha 可预测性衡量了源于理性可预测性之外的可预测程度, 这种预测性可能与投资者行为有关.

对于投资组合  $i$ , 超额收益率可以表示为

$$r_{i,t} = \alpha_i(x_{t-1}) + \beta_i' f_t + u_{i,t} \quad (12)$$

其中  $x_{t-1}$  表示  $J$  维滞后预测因子,  $f_t$  是  $K$  维基于投资组合的系统性风险因子,  $\beta_i'$  表示  $K$  维投资组合  $i$  的  $\beta$  系数,  $u_{i,t}$  表示随机扰动项. 若假设系统性风险因子  $f_t$  满足

$$f_t = \lambda(x_{t-1}) + \varepsilon_t \quad (13)$$

其中  $\lambda(x_{t-1})$  是基于滞后预测变量  $x_{t-1}$  的系统性风险因子  $f_t$  的期望值. 该条件期望值代表时变的风险溢价, 即  $\lambda$  受  $x_{t-1}$  的影响是随时间变化的.  $\varepsilon_t$  是独立于  $u_{i,t}$  的随机扰动项.

根据姜富伟等<sup>[48]</sup>, 从资产定价模型的条件变形式可以得  $r_{i,t}$  的条件期望收益

$$E(r_{i,t} | x_{t-1}) = \beta_i' E(f_t | x_{t-1}) = \beta_i' \lambda(x_{t-1}) \quad (14)$$

当维数  $K=1$  时, 式(14)就是条件 CAPM 模型.  $f_t$  是一个标量, 代表总体市场的超额收益率.  $\lambda(x_{t-1})$  是预期的股票溢价. 根据式(14)可知, 在条件 CAPM 形式的定价模型下, 任何组合收益率  $r_{i,t}$  的预测性仅包含风险因子的预测性和投资组合收益率  $r_{i,t}$  对市场组合的敏感性, 即风险暴露  $\beta_i' \lambda(x_{t-1})$ . 此时  $\alpha_i(x_{t-1})$  为 0. 而当  $\alpha_i(x_{t-1})$  不为 0 则表示存在 Alpha 可预测性, 即不能被  $\beta_i' \lambda(x_{t-1})$  解释的定价误差.

$r_{i,t}$  基于条件 CAPM 理性定价的综合预测收益率  $\hat{r}_{i,t}^R$  为

$$\hat{r}_{i,t}^R = \hat{\beta}_{i,t}' \hat{f}_t^c \quad (15)$$

其中  $\hat{f}_t^c$  是总体市场的综合预测收益率, 取所有因子预测的总市场收益率的平均值.  $\hat{\beta}_{i,t}'$  由每个投资组合超额收益率对总市场超额收益  $\hat{f}_t^c$  的无截距回归获得, 即  $\{r_{i,t}\}_1^{t-1}$  对  $\{\hat{f}_t^c\}_1^{t-1}$  做回归估计出  $\hat{\beta}_{i,t}'$ . 将  $\hat{f}_t^c$  与  $\hat{\beta}_{i,t}'$  相乘, 即得到  $r_{i,t}$  基于条件 CAPM 理性定价的平均预测收益率  $\hat{r}_{i,t}^R$ .

根据姜富伟等<sup>[48]</sup>, 可将样本外统计量  $OR^2$  分解成理性可预测性  $OR_R^2$  和 Alpha 可预测性  $OR_\alpha^2$ .  $OR_R^2$  衡量了条件资产定价模型约束的综合预测收益率相对于历史平均收益的均方预测误差减少量, 反应了投资组合条件资产定价模型的预测性.  $OR_\alpha^2$  测度了无约束综合预测收益率相对于条件资产定价模型约束的综合预测收益率的均方预测误差减少量, 代表了投资组合的 Alpha 可预测性.

$OR_R^2$  和  $OR_\alpha^2$  的具体计算公式如下

$$OR_R^2 = 1 - \frac{\sum_{i=t+1}^{2t} (r_i - \hat{r}_i^R)^2}{\sum_{i=t+1}^{2t} (r_i - \bar{r})^2} \quad (16)$$

$$OR_\alpha^2 = 1 - \frac{\sum_{i=t+1}^{2t} (r_i - \hat{r}_i)^2}{\sum_{i=t+1}^{2t} (r_i - \hat{r}_i^R)^2} \quad (17)$$

其中  $\hat{r}_i$  为无条件平均预测收益率, 同时包含了理性可预测性和 Alpha 可预测性.

根据  $OR^2$  的综合预测收益率表达式以及式(16)和式(17)可以得到

$$OR_\alpha^2 = 1 - \frac{\sum_{i=t+1}^{2t} (r_i - \hat{r}_i)^2}{\sum_{i=t+1}^{2t} (r_i - \bar{r})^2} \times \frac{\sum_{i=t+1}^{2t} (r_i - \bar{r})^2}{\sum_{i=t+1}^{2t} (r_i - \hat{r}_i^R)^2} \quad (18)$$

$$= 1 - \frac{1 - OR^2}{1 - OR_R^2}$$

化简式(18)可以发现如下关系

$$OR^2 = OR_\alpha^2 + OR_R^2 - OR_\alpha^2 \times OR_R^2 \quad (19)$$

在  $OR_\alpha^2$  与  $OR_R^2$  都很小的情况下, 两者的乘积项  $OR_\alpha^2 \times OR_R^2$  近似为 0. 因此认为, 样本外预测性统计量  $OR^2$  可近似分解成基于条件资产定价模型的可预测性  $OR_R^2$  和 Alpha 可预测性  $OR_\alpha^2$  两部分, 即

$$OR^2 \approx OR_\alpha^2 + OR_R^2 \quad (20)$$

### 3 样本与数据

本文选取 2006 年 7 月至 2016 年 6 月沪深 A 股的全部股票超额收益率为研究对象, 剔除缺失数据、ST 和 \* ST 股票后, 共计选取样本股共 1 190 只. 收益率和因子数据均来自 WIND 数据库, 部分财务指标来自国泰安数据库. 本文构造四类投资组合, 其中行业投资组合按照 WIND 一级行业分类标准分为: 材料、工业、能源、可选消费、日常消费、医疗保健、金融、信息技术、电信服务、公用事业. 账面市值比投资组合按照每年 6 月的账面市值比将所有股票分成 10 等分组,  $BM1, BM2, \dots, BM10$  表示账面市值比排名前 10%, 10% ~ 20%,  $\dots$ , 最后 10% 的股票. 由于我国没有账面

市值比的数据,因此  $t+1$  年 6 月的账面市值比等于  $t-1$  年账面价值比  $t+1$  年 6 月的市值,其中账面价值由每股净资产与流通股的乘积计算. 市值投资组合按照每年 6 月的市值将所有

股票分成 10 等分组,  $SIZE1, SIZE2, \dots, SIZE10$  表示账面市值比排名前 10%, 10% ~ 20%,  $\dots$ , 最后 10% 的股票.  $t$  年 6 月的分组标准即为当月的市值.

表 1 因子名称及其计算方法

Table 1 Construction of predicting factors

	因子名称	计算方法
<i>BM</i>	账面市值比( book-to-market)	$t-1$ 年每股净资产比当期股价
<i>DE</i>	股利分配率( dividend-payout)	$t-1$ 年每股股利与 $t-1$ 年每股收益对数差
<i>DP</i>	股息价格比( dividend-price)	$t-1$ 年每股股利与当期股价对数差
<i>DY</i>	股息收益率( dividend-yield)	$t-1$ 年每股股利与滞后一期股价对数差
<i>EP</i>	每股收益价格比( earnings-price)	$t-1$ 年每股收益与当期股价对数差
<i>CFP</i>	现金收益价格比( cash flow-price)	$t-1$ 年每股现金收益比当期股价
<i>INF</i>	通货膨胀率( inflation)	当期的 CPI 增长率
<i>SVAR</i>	股票波动率( stock variance)	沪深 300 日收益率的平方和

注: 每股股息、每股收益均为一年 12 个月的总股息、总收益; 每股净资产、每股现金收益为每年 12 月末的值.

表 2 各投资组合收益率的描述性统计量

Table 2 Descriptive statistics of the returns of different portfolios

	A: 市场及行业投资组合					B: 账面市值比分组投资组合					C: 市值分组投资组合						
	均值 $\times 100$	中值 $\times 100$	标准 差	偏度	峰度		均值 $\times 100$	中值 $\times 100$	标准 差	偏度	峰度		均值 $\times 100$	中值 $\times 100$	标准 差	偏度	峰度
总市场	1.70	2.20	0.10	-0.23	3.20	<i>BM1</i>	-0.30	0.40	0.11	-0.19	3.55	<i>SIZE1</i>	1.10	0.80	0.10	-0.14	3.34
能源	0.80	1.30	0.12	-0.13	3.85	<i>BM2</i>	0.60	0.20	0.11	0.10	4.04	<i>SIZE2</i>	1.80	1.90	0.11	0.20	4.00
材料	1.70	2.00	0.12	-0.13	3.02	<i>BM3</i>	0.90	1.40	0.11	-0.16	3.26	<i>SIZE3</i>	1.90	2.20	0.11	-0.05	3.27
工业	1.70	1.70	0.11	-0.21	3.14	<i>BM4</i>	1.70	1.10	0.11	0.02	3.85	<i>SIZE4</i>	2.30	2.50	0.11	-0.11	3.60
可选 消费	1.90	2.00	0.11	-0.06	3.37	<i>BM5</i>	1.70	1.60	0.11	-0.06	3.50	<i>SIZE5</i>	2.60	1.40	0.12	-0.03	3.16
日常 消费	1.70	1.60	0.09	0.01	3.80	<i>BM6</i>	1.70	1.90	0.11	-0.41	3.02	<i>SIZE6</i>	2.70	2.50	0.12	0.03	3.26
医疗	1.90	1.90	0.10	0.33	4.71	<i>BM7</i>	2.00	3.20	0.11	-0.26	3.06	<i>SIZE7</i>	3.00	2.40	0.12	-0.02	3.23
公共 事业	1.20	0.60	0.10	0.41	5.18	<i>BM8</i>	2.50	2.70	0.11	-0.16	3.16	<i>SIZE8</i>	3.20	2.60	0.12	0.08	3.58
金融	2.00	0.30	0.12	0.21	3.41	<i>BM9</i>	3.50	3.40	0.10	-0.10	3.18	<i>SIZE9</i>	4.00	4.10	0.12	0.00	3.14
电信	1.10	0.50	0.12	0.95	8.47	<i>BM10</i>	4.40	3.70	0.11	0.16	3.42	<i>SIZE10</i>	4.70	3.70	0.12	0.48	4.40
信息 技术	2.00	2.00	0.11	-0.20	3.03												

### 3.1 因子选择

本文的因子选择参考 Welch 和 Goyal<sup>[49]</sup>、Narayan 和 Bannigidadmath<sup>[50]</sup> 两篇文献. Welch 和 Goyal<sup>[49]</sup> 选择了 14 个因子, 侧重市场宏观因子的

可预测性研究; Narayan 和 Bannigidadmath<sup>[50]</sup> 的研究对象是印度股票市场, 因此选取的因子是 Welch 和 Goyal<sup>[49]</sup> 的子集, 包含 6 个基本面因子和 2 个宏观市场因子. 考虑到本文侧重中国股市

基本面因子的分析、中国宏观因子数据的可得性、中国和印度一样同属于发展中国家,选取与 Narayan 和 Bannigidadmath<sup>[50]</sup> 一样的 8 个因子,包含 6 个基本面因子(账面市值比、股利分配率、股息价格比、股息收益率、每股收益价格比、现金收益价格比)、1 个市场因子(沪深 300 指数波动率)、1 个宏观因子(通货膨胀率)。因子的计算方法见表 1。值得注意的是  $t$  年 7 月至  $t+1$  年 6 月的收益率要与日历会计年度  $t-1$  年的财务指标相对应。这样可以保证预测投资组合收益率时,所有因子已知,使得实证结果更加具有现实的预测意义。

### 3.2 统计性检验

表 2 对各投资组合的收益率进行了统计描述。整体市场的平均超额收益率为 1.74%,左偏分布,均值受低收益股票影响较大。行业投资组合的平均超额收益率从能源行业的 0.81% 到金融行业的 1.99%,标准差分别为 11.48% 和 11.50%。公共事业、金融、电信行业的收益率均值远大于中值,相差 2 倍以上,特别是金融行业达到了 7 倍,且峰度均为正,表现为右偏,说明这三个行业高收益趋势显著,对均值影响更大。

账面市值比投资组合的平均超额收益率从 -0.26% 到 4.35%,从均值来看并不符合 Fama 和 French 提出的账面市值比效应<sup>[1]</sup>。但从  $BM1$  到  $BM2$ 、 $BM3$  到  $BM4$ 、 $BM7$  到  $BM8$  存在一个中位数的递减趋势,另外在高账面市值比组合( $BM1 \sim BM5$ )中,偏度为正的组合更多,说明账面市值比高的投资组合倾向高收益,且  $BM1$  至  $BM10$  的标准差由 11.04% 到 10.46%,账面市值比与收益率波动性成正比。

由于样本数量的限制、时间跨度的选择等因素,本文并没有发现明显的账面市值比效应,但仍发现了相应的趋势,即高账面市值比投资组合更倾向于右偏分布,收益率波动性也更大。

市值投资组合平均超额收益由 1.06% 至 4.66%。除了第一组例外,大市值组合( $SIZE1 \sim SIZE4$ )的均值都小于中值,偏度大多为负,左偏趋势明显,小市值组合( $SIZE5 \sim SIZE10$ )刚好相反,说明大市值组合倾向低收益而小市值组合倾向高收益,且标准差随着市值的减小而增大。综上所述,满足 Fama 与 French 提出的市值效应<sup>[1]</sup>,即

市值越小的投资组合平均收益越大,波动性也更强。另外,所有投资组合收益率的峰度都大于 3,表现为厚尾非正态分布。

### 3.3 计量检验

此部分目的是检验收益率及因子序列是否存在高度持续性、异方差和内生性问题。首先,通过收益率和因子序列的一阶自回归系数进行检验,发现除了收益率序列、 $INF$  和  $SVAR$  因子,其他所有序列  $AR(1)$  系数均接近 1,表明存在高度的持续性。

运用针对自回归条件异方差(ARCH)过程提出的 LM 检验法,本文对收益率和因子序列是否存在异方差进行了检验。在最大阶数为 12 的情况下,计算 LM 检验的  $F$  统计量,发现所有因子序列在 1% 的显著性水平下拒绝“不存在 ARCH 效应”的原假设。收益率序列的 ARCH 效应在市场、6 个行业组合(材料、工业、日常消费、医疗、公共事业、信息技术)、5 个  $BM$  组合( $BM5$ 、 $BM6$ 、 $BM8$ 、 $BM9$ 、 $BM10$ )、3 个  $SIZE$  组合( $SIZE1$ 、 $SIZE2$ 、 $SIZE9$ )中较为明显,在 1% 的水平下显著。

参考 Narayan 和 Bannigidadmath<sup>[50]</sup> 对因子序列的内生性进行了检验。具体步骤如下:第一步,每个投资组合对式(1)进行最小二乘估计,得残差项。第二步,每个预测因子对式(2)进行自回归估计,得残差项。第三步,根据式(3)检验两个残差项是否存在相关关系,得到系数  $\gamma$  及其  $t$  统计量。若系数显著不为 0,则存在内生性问题。通过检验,发现除了  $DY$ 、 $INF$ ,其他预测因子都存在显著的内生问题,至少在 1% 的显著性水平下拒绝“ $\gamma=0$ ”的原假设。

通过检验分析,发现大部分因子序列确实存在高度持续性、异方差和内生性问题,直接用最小二乘法进行预测性研究会造成偏误,无法准确度量因子的预测性。因此,采用可行拟广义最小二乘法(FQGLS)进行样本内检验,以期克服上述问题。

## 4 检验结果

### 4.1 样本内检验结果

总体市场和行业投资组合的实证结果如表 3

的 A 部分所示. 从因子对总体市场收益率的预测能力来看: *EP*、*BM*、*DE* 和 *SVAR* 四个因子能显著预测总体市场的收益率, 其余因子均不显著. 从因子对行业投资组合收益率的预测能力来看: *EP* 和 *BM* 可显著预测 8 个行业组合的收益率, 是预测效果最好的两个因子; 预测效果其次的是 *CFP* 和 *DP* 因子, 可预测 4 个行业组合的收益率; 而其余因子最多可预测 3 个行业组合的收益率. 从行业投资组合收益率的被预测程度来看: 日常消费、医疗和金融是最能被预测的行业组合, 分别被 6、6 和 4 个因子预测; 能源行业是最不能被预测的行业, 仅有 1 个因子预测显著; 而其他行业最多可被 3 个因子预测. 从因子的总体表现来看: *EP* 和 *BM* 既可以预测 8 个行业投资组合的收益率, 也可以显著预测总体市场的收益率, 是预测性能最好的因子; *DE*、*INF* 和 *SVAR* 是预测性能较差的因子, 仅能预测 2 个行业组合.

账面市值比投资组合的实证结果如表 3 的 B 部分所示. 从因子的预测能力来看: *EP* 和 *BM* 的预测能力最好, 分别能够预测 10 和 9 个组合的收益率; *CFP*、*DP*、*DY* 的预测能力其次, 能够预测 4 个组合的收益率; 而其余因子 *DE*、*SVAR* 和 *INF* 预测能力较差, 分别可预测 3、2 和 1 个组合的收益率. 综上所述, 因子预测能力的结果与市场及行业组合结果一致. 从投资组合收益率的被预测程度看: *BM1*、*BM3*、*BM5*、*BM8*、*BM9* 组合的收益率有非常好的被预测性, 可被 4 个以上的因子预测; *BM2*、*BM6*、*BM10* 组合收益率的被预测性最差, 仅能被 2 个因子预测.

市值投资组合的实证结果如表 3 的 C 部分所示. 从因子的预测能力来看: *EP* 和 *BM* 仍是预测能力最好的因子, 能够预测 7 个组合的收益率; *DE* 和 *CFP* 的预测能力其次, 分别可预测 5 和 4 个组合的收益率; 而其余因子预测能力较差, 最多可预测 3 个组合的收益率, 特别是因子 *INF* 和 *SVAR* 仅可预测 1 个和 0 个组合的收益率. 与行业组合和账面市值比组合结果不同的是, *DE* 因子的预测性能从 3/10 提升为 5/10. 从投资组合收益率的被预测程度看: *SIZE1*、*SIZE3*、*SIZE4*、*SIZE9* 组合有非常好的被预测性, 可被超过半数的因子预测; *SIZE6*、*SIZE7* 组合的被预测性最差, 仅能被 1 个因子预测.

综上所述, 样本内检验结果可总结为: 1) *EP* 和 *BM* 因子预测能力最强, 不仅能够成功预测市场收益率, 还可分别预测 26 和 25 个投资组合的收益率; *CFP*、*DP*、*DE* 因子的预测能力其次, 至少可预测 11 个组合的收益率, 其中 *CFP* 和 *DP* 在各个投资组合中表现稳定, *DE* 仅在市值组合中表现较好; *INF* 和 *SVAR* 最多预测 5 个组合的收益率, 预测能力最差. 2) 总体市场, 行业组合 (日常消费、医疗、金融), 账面市值比组合 (*BM1*、*BM3*、*BM5*、*BM8*、*BM9*) 和市值组合 (*SIZE1*、*SIZE3*、*SIZE4*、*SIZE9*) 的被预测程度最强, 可被 4 个以上因子预测; 而行业组合 (能源、电信、信息技术), 账面市值比组合 (*BM4*、*BM6*、*BM10*), 市值组合 (*SIZE5*、*SIZE6*、*SIZE7*、*SIZE8*、*SIZE10*) 的被预测性最差, 最多可被 2 个因子预测.

Narayan 和 Bannigidadmath<sup>[50]</sup> 也用同样的方法对印度股市进行了检验, 对比本文的研究结果可以发现, 相同点是 *EP* 都很好的预测了大部分投资组合, *DP*、*DY* 次之, 而 *SVAR*、*INF* 预测性都较差. 不同点是 *BM* 在我国股市中预测性很好而在印度市场表现一般; *CFP* 因子表现较好而在印度市场表现很差; *DE* 在我国股市表现一般却成为印度市场预测性最好的因子.

## 4.2 样本外检验结果

### 4.2.1 全体样本外区间检验

将研究区间 2006 年 7 月至 2016 年 6 月平均分成两部分, 前半部分 2006 年 7 月至 2011 年 6 月为估计样本, 后半部分 2011 年 7 月至 2016 年 6 月为预测样本. 本文认为, 三个指标中有两个及以上通过检验即认为因子具有预测性. 单因子的样本外检验结果如表 4 ~ 表 6 所示. 表 4 表明: *INF*、*SVAR*、*CFP* 可对市场收益率进行有效地样本外预测, 且有 2 个以上指标支持该结果. 在行业投资组合中 *INF*、*SVAR* 因子样本外预测能力最强, 分别预测了 9 和 7 个行业的收益率; *DE*、*EP* 的样本外预测能力其次, 预测了 4 个行业的收益率; *BM* 的样本外预测能力最差, 仅预测了 1 个行业的收益率; 而其余因子最多预测 3 个行业. 日常消费、金融、能源是样本外预测最好的行业, 能被 5 个以上因子预测; 其余行业表现中庸, 最多被 3 个因子预测.

表 3 样本内预测检验结果 (统计量  $t_{FQGLS}$ )  
Table 3 Lists of statistic  $t_{FQGLS}$  from in-sample tests

BZ]	A：市场及分行业投资组合							B：账面市值比分组投资组合							C：市值分组投资组合												
	BM	CFP	DE	DP	DY	EP	INFL	SVAR	摇	BM	CFP	DE	DP	DY	EP	INFL	SVAR	摇	BM	CFP	DE	DP	DY	EP	INFL	SVAR	
市场	3.10***	-1.19	2.28**	1.57	0.89	3.79***	0.89	2.09**	BM1	5.08***	13.78***	0.96	7.15***	1.95*	2.24**	0.97	15.47***	SIZE1	2.11**	13.39***	1.70*	2.54**	1.59	15.44***	1.09	1.15	
能源	0.13	-1.50	1.12	0.74	0.14	6.46***	1.40	0.53	BM2	2.95***	1.04	1.43	1.17	2.58**	4.62***	0.51	1.12	SIZE2	0.91	0.90	4.14***	1.81*	0.51	1.75*	0.97	1.11	
材料	3.40***	0.55	1.04	10.64***	1.20	1.58	1.00	1.69*	BM3	5.72***	2.97***	3.19***	2.75***	2.80***	4.96***	0.55	1.36	SIZE3	29.50***	1.86*	1.82*	0.89	1.06	2.46**	1.23	1.19	
工业	2.09**	1.44	1.47	30.54***	0.06	2.19**	0.83	1.13	BM4	6.76***	1.48	0.37	0.80	0.80	1.93*	1.13	1.25	SIZE4	3.20***	1.43	3.85***	0.79	2.33**	7.13***	0.54	0.83	
可选消费	2.71***	2.33***	1.40	0.51	1.39	14.21***	1.12	1.28	BM5	9.70***	0.94	1.83*	1.83*	2.29**	2.22**	1.62	1.44	SIZE5	4.70***	0.63	1.13	0.53	0.32	6.99***	0.90	1.10	
日常消费	4.61***	0.55	2.81***	1.70*	1.71*	4.80***	2.00**	1.62	BM6	30.27***	0.62	0.23	0.96	0.82	3.36***	0.63	0.50	SIZE6	4.88***	0.91	1.63	0.75	1.11	1.14	1.55	0.21	
医疗	8.80***	1.45	2.00**	2.52**	2.86***	6.32***	1.95*	0.14	BM7	4.20***	0.59	1.09	18.17***	0.36	10.07***	1.32	0.70	SIZE7	0.19	0.36	1.24	1.33	0.67	8.50***	1.36	0.28	
公共事业	1.97*	1.81*	1.23	1.61	0.31	2.76***	0.57	0.77	BM8	15.11***	2.77***	0.95	0.94	0.85	6.33***	0.98	1.71*	SIZE8	13.43***	0.56	1.67*	1.32	1.57	1.47	1.40	0.21	
金融	1.22	7.21***	1.55	0.25	3.61***	1.80*	0.25	4.26***	BM9	4.46***	4.06***	0.30	0.75	1.11	5.19***	2.33**	0.83	SIZE9	0.08	2.92***	1.38	1.83*	1.85*	4.13***	2.03**	0.09	
电信	28.25***	1.92*	0.62	1.05	0.95	0.16	0.85	0.95	BM10	0.18	0.11	2.07**	0.01	0.37	2.60**	1.26	0.98	SIZE10	2.53**	1.78*	0.25	1.57	1.25	0.20	1.37	0.21	
信息技术	2.28**	0.46	0.01	0.90	0.47	18.32***	1.64	1.02																			

摇摇注：t 统计量由上文 FQGLS 方法计算，\*\*\*、\*\*、\* 分别代表 1%、5%、10% 的显著性水平。



表 4 市场及分行业投资组合的样本外检验结果  
Table 4 Results of out-of-sample tests for the market and industry portfolios

摇	BM			CFP			DE			DY			DP			EP			INF			SVAR		
	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>
总体市场	1.044	1.036	-0.090	1.010	1.016	-0.003	1.004	1.015	1.015	-0.008	1.006	1.015	1.011	1.002	1.012	1.002	1.012	-0.004	0.985	0.990	0.030	0.992	0.998	0.016
能源	1.080	1.061	-0.167	1.006	-0.062	0.041	1.003	0.984	0.984	-0.006	1.003	0.985	-0.006	1.011	0.992	1.011	0.992	-0.023	0.998	0.999	0.003	1.019	1.005	-0.039
材料	1.049	1.050	-0.100	1.045	1.063	-0.092	0.982	0.989	0.989	-0.037	1.019	1.001	-0.039	1.020	1.000	1.020	1.000	-0.041	0.982	0.987	0.035	1.006	1.004	-0.012
工业	1.059	1.041	-0.122	0.997	0.993	0.006	1.017	1.020	-0.034	1.020	1.021	1.026	-0.043	1.014	1.020	1.014	1.020	-0.028	0.994	0.995	0.013	0.994	1.000	0.012
可选消费	1.042	1.026	-0.085	0.990	0.998	0.020	1.014	1.017	-0.027	1.021	1.022	1.021	-0.044	1.020	1.018	1.020	1.018	-0.041	0.979	0.988	0.042	0.993	1.001	0.014
日常消费	1.087	1.040	-0.181	1.001	1.008	-0.003	0.992	0.999	0.015	0.993	0.998	0.997	0.011	0.989	0.990	0.989	0.990	0.022	0.960	0.966	0.078	0.996	1.008	0.008
医疗	0.991	0.985	0.019	1.028	1.010	-0.057	1.054	1.016	-0.110	1.048	1.013	1.013	-0.098	1.048	1.013	1.048	1.013	-0.095	0.969	0.973	0.062	1.002	1.007	-0.003
公共事业	1.038	1.017	-0.077	1.058	1.026	-0.119	1.073	1.036	-0.151	1.041	1.029	1.030	-0.087	1.043	1.030	1.043	1.030	-0.056	0.984	0.992	0.033	0.995	0.999	0.010
金融	1.010	1.009	-0.019	1.149	1.306	-0.320	0.990	1.020	0.020	0.989	1.021	1.022	0.019	0.987	1.019	0.987	1.019	0.025	1.001	1.001	-0.002	0.991	0.992	0.017
电信	1.004	1.004	-0.008	1.002	1.001	-0.004	1.183	1.238	-0.400	1.027	1.015	1.015	-0.055	1.028	1.015	1.014	1.007	-0.029	0.995	0.997	0.011	0.998	1.003	0.005
信息技术	1.031	1.017	-0.064	1.017	1.008	-0.034	1.036	1.023	-0.074	1.038	1.022	1.023	-0.079	1.039	1.021	1.039	1.021	-0.079	0.999	0.993	0.002	0.991	0.997	0.017

表 5 账面市值比分组投资组合的样本外检验结果  
Table 5 Results of out-of-sample tests for the BM portfolios

摇	BM			CFP			DE			DY			DP			EP			INF			SVAR		
	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>
BM1	1.071	1.068	-0.148	1.007	1.004	-0.013	1.016	1.010	-0.033	1.019	1.004	1.014	1.007	1.007	-0.029	1.014	1.011	-0.028	0.998	0.995	0.004	0.999	0.994	0.001
BM2	1.096	1.071	-0.201	1.022	1.010	-0.045	0.973	0.996	0.054	1.058	1.057	1.081	1.070	1.070	-0.169	1.083	1.066	-0.172	0.996	0.996	0.009	0.990	0.998	0.020
BM3	1.042	1.032	-0.085	0.985	0.992	0.029	0.967	0.979	0.065	1.018	1.028	1.021	1.033	1.033	-0.043	1.067	1.064	-0.139	0.989	0.996	0.022	0.996	0.995	0.009
BM4	1.030	1.028	-0.060	1.024	1.027	-0.048	0.990	1.002	0.019	0.979	1.017	0.990	1.016	1.016	0.021	1.021	1.029	-0.042	0.975	0.989	0.050	0.987	0.994	0.025
BM5	1.048	1.034	-0.098	1.026	1.018	-0.052	1.016	1.004	-0.032	1.080	1.044	1.101	1.068	1.068	-0.213	1.032	1.011	-0.065	0.967	0.972	0.064	0.994	1.002	0.013
BM6	1.044	1.015	-0.089	1.007	1.010	-0.013	1.097	1.061	-0.204	1.054	1.045	1.049	1.042	1.042	-0.100	1.002	1.004	-0.003	1.002	1.001	-0.004	0.997	1.003	0.006
BM7	1.002	1.000	-0.004	1.003	1.003	-0.006	0.996	1.002	0.009	1.002	1.008	1.009	1.014	1.014	-0.017	0.951	0.969	0.097	0.984	0.993	0.032	0.995	0.997	0.011
BM8	1.016	1.002	-0.032	1.008	1.006	-0.016	1.001	1.001	-0.002	1.018	1.021	1.007	1.013	1.013	-0.013	1.038	1.029	-0.078	0.981	0.988	0.038	1.001	1.000	-0.001
BM9	0.998	1.000	0.004	0.999	1.018	0.002	1.037	1.029	-0.076	0.994	0.994	1.000	1.000	1.000	-0.001	0.985	0.986	0.030	0.969	0.979	0.061	1.006	1.006	-0.013
BM10	1.021	1.012	-0.042	1.003	1.005	-0.007	0.991	0.984	0.018	1.007	1.006	1.009	1.009	1.009	-0.018	1.001	0.992	-0.002	0.973	0.987	0.054	1.002	1.002	-0.003

表6 播市值分组投资组合的样本外检验结果

Table 6 Results of out-of-sample tests for the size portfolios

播	BM			CFP			DE			DY			DP			EP			INF			SVAR		
	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>	RMAE	RRMSE	OR <sup>2</sup>
SIZE1	1.040	1.037	-0.083	1.008	1.001	-0.015	0.960	0.973	0.079	1.025	1.007	-0.051	1.021	1.010	-0.043	1.012	1.010	-0.023	0.989	0.994	0.022	0.992	0.996	0.017
SIZE2	1.040	1.032	-0.081	0.990	0.997	0.020	1.025	1.011	-0.051	1.042	1.034	-0.086	1.043	1.034	-0.088	1.000	0.999	0.000	0.988	0.992	0.023	0.995	1.001	0.010
SIZE3	1.042	1.019	-0.086	0.980	0.994	0.040	1.007	1.011	-0.014	1.025	1.033	-0.050	1.013	1.028	-0.026	0.973	0.974	0.054	0.985	0.991	0.030	0.990	1.000	0.020
SIZE4	1.030	1.021	-0.061	0.989	0.992	0.022	1.004	0.996	-0.009	1.007	0.997	-0.014	1.012	1.005	-0.024	1.027	1.015	-0.055	0.985	0.988	0.031	0.998	1.003	0.004
SIZE5	1.079	1.026	-0.164	1.010	1.000	-0.020	1.084	1.111	-0.176	1.047	1.033	-0.096	1.052	1.035	-0.107	1.018	1.013	-0.037	0.992	0.994	0.016	0.983	1.005	0.035
SIZE6	1.014	1.004	-0.029	0.998	0.997	0.003	1.075	1.051	-0.155	1.058	1.047	-0.120	1.060	1.045	-0.124	1.027	1.015	-0.054	0.982	0.985	0.036	0.997	1.009	0.006
SIZE7	1.031	1.013	-0.062	1.003	1.001	-0.006	1.052	1.037	-0.107	1.104	1.077	-0.219	1.112	1.082	-0.236	1.013	1.004	-0.026	0.978	0.987	0.043	0.991	1.007	0.018
SIZE8	0.994	0.999	0.012	1.009	1.006	-0.019	1.065	1.038	-0.134	1.031	1.046	-0.062	1.032	1.045	-0.065	1.036	1.015	-0.073	0.976	0.985	0.048	0.996	1.008	0.008
SIZE9	0.965	0.972	0.069	0.985	0.993	0.029	1.031	1.028	-0.062	1.043	1.034	-0.089	1.052	1.039	-0.107	0.925	0.980	0.144	0.981	0.979	0.037	1.004	1.008	-0.008
SIZE10	0.987	0.992	0.025	1.000	1.003	-0.001	1.011	1.023	-0.023	1.021	1.011	-0.043	1.026	1.011	-0.052	1.018	1.024	-0.036	0.967	0.980	0.065	1.001	1.006	-0.002

账面市值比投资组合的样本外预测结果如表5所示。可以发现,因子  $INF$  和  $SVAR$  样本外预测能力最强,分别预测了9和7个组合的收益率;因子  $DE$  样本外预测能力其次,可预测5个组合的收益率;因子  $BM$  样本外预测能力最差,仅可预测1个组合的收益率;其余因子最多可预测3个组合的收益率。该结果与行业组合样本外检验的结果一致。投资组合  $BM3$ 、 $BM4$ 、 $BM7$ 、 $BM9$  是样本外预测最好的组合,可被超过半数以上的因子预测;其余组合的样本外预测表现一般,最多可被3个因子预测。

市值投资组合的样本外预测结果如表6所示。可以发现,  $INF$ 、 $SVAR$ 、 $CFP$  仍是样本外预测能力最强的因子,分别预测了10、8、6个组合的收益率。该结果与行业组合和账面市值比组合的结果一致。 $DP$  样本外预测能力最差,没有指标支持它对市值组合具备样本外预测性。投资组合  $SIZE2$ 、 $SIZE3$ 、 $SIZE4$ 、 $SIZE9$  的样本外被预测性最好,能被超过半数以上因子的预测;其余组合的样本外预测表现一般,最多可被3个因子预测。

综上所述,样本外检验结果可总结为:1)  $INF$  和  $SVAR$  是样本外预测能力最强的因子,  $OR^2$  从0.11%至7.81%,不但可对市场收益率进行样本外预测,还能分别对29和23个组合进行样本外预测,且有2个以上指标支持该结果;2) 因子  $CFP$ 、 $DE$ 、 $EP$  的样本外预测表现相对良好,能至少预测10个组合的收益率,其中  $CFP$  在总体市场和市值组合中表现较好,而  $DE$  和  $EP$  在行业、账面市值比和市值投资组合中表现稳定;3)  $BM$  和  $DP$  是样本外预测能力最差的因子,仅可预测5个投资组合的收益率。4) 样本外被预测效果较好的行业投资组合有日常消费、金融、能源,账面市值比组合有  $BM3$ 、 $BM4$ 、 $BM7$ 、 $BM9$  组合,市值组合有  $SIZE2$ 、 $SIZE3$ 、 $SIZE4$ 、 $SIZE9$  组合,均可被半数以上因子预测。

样本外检验的结果表明大部分因子的样本外预测效果较于样本内有所下降,但总体来看样本内外结果是基本一致的,如  $CFP$ 、 $DE$ 、 $DP$ 、 $DY$ ;也有一些因子样本内表现较好,样本外表现却很差,如  $BM$  因子。该结果与论文 Goyal 和 Welch 等<sup>[51]</sup>、Butler 等<sup>[52]</sup>一致,是由系统性问题导致的<sup>[41]</sup>。并

且样本外检验发现,市场宏观因子表现较为突出,该结论与姜富伟等<sup>[48]</sup>一致。

进一步对所有因子预测的收益率求平均,检验了综合预测收益率的样本外预测效果,结果如表7的A部分所示。可以发现,在大部分投资组合中,特别是市值投资组合中小盘股,综合预测收益率的样本外预测效果较单因子的预测效果有了显著地提升。而且,从样本外预测指标来看,对于市场组合,包括日常消费、材料、工业、可选消费在内的行业组合,包括  $BM1$  和  $BM10$  在内的账面市值比组合,包括  $SIZE6$ 、 $SIZE8$ 、 $SIZE10$  在内的市值组合,综合预测收益率的预测效果要优于历史平均收益率。其他涉及中国股市收益率样本外预测的研究<sup>[40,53,54]</sup>,也得出了综合预测收益率的样本外预测效果好于历史平均收益率预测的结论。

#### 4.2.2 不同时期的样本外检验

Fama 和 French 等学者曾经提出,伴随金融危机而产生的风险厌恶应有更高的风险溢价<sup>[48]</sup>,这将导致投资组合的收益率在危机期间有更强的可预测性。因此,本文进一步研究了不同经济周期各投资组合综合预测收益率的样本外预测效果。根据何兴强和周开国<sup>[53]</sup>股市周期的划分方法,将整个样本预测期2011年7月至2016年6月进行了牛市与熊市的区分,其中2011年7月至2012年11月、2015年6月至2016年6月为熊市,2012年11月至2015年6月为牛市。2011年7月至2012年11月期间由于存在IPO重启、紧缩的宏观政策、欧洲债务危机等问题,股市跌幅近40%;2012年11月至2015年6月波动较大,期初虽呈现小幅度下跌,但后期涨势迅猛,最高涨幅甚至达到了160%,整体表现为牛市;2015年6月至2016年6月仍为熊市,跌幅近40%。本文熊市的综合预测收益率  $OR^2$  取两次熊市的均值。

表7的B部分是市场及行业、账面市值比、市值投资组合的综合预测收益率样本外检验  $OR^2$  在熊市、牛市和全样本的样本外预测效果。不同时期的平均预测水平由10个行业、10组账面市值比及10组市值投资组合的均值  $OR^2$  表示。对于总体市场,牛市  $OR^2$  (11.76%) 约为熊市  $OR^2$  (1.25%) 的11倍,说明牛市市场收益率的预测性比熊市强。对于行业投资组合结论类似,牛市平

均  $OR^2(7.42\%)$  远大于熊市平均  $OR^2(-0.51\%)$ . 而行业组合的结果表明在整个样本预测期有较高预测性的三个行业是日常消费、材料和能源,而且这三个行业在牛市时期的  $OR^2$  也是最高的,说明预测性最强. 较整个样本预测期,牛市时期所有行业组合  $OR^2$  都得到显著提升,预测性增强;较熊市阶段平均结果,牛市阶段所有行业组合的  $OR^2$  更高,预测性更强. 值得指出的是,整个样本预测期包含两段熊市,不同行业在不同熊市的表現也会有不同,能源、材料、电信、信息技术行业在第一段熊市的预测性甚至远优于牛市,该发现与 Kadan 等<sup>[54]</sup> 的观点一致,即宏观经济周期对各行业的影响是不一样的,行业收益率有时可传达总体市场收益率而没有反映出的额外信息.

账面市值比组合和市值组合与市场及行业组合的结果完全相反. 熊市阶段的平均  $OR^2$  (分别为  $0.64\%$ 、 $0.9\%$ ) 远大于牛市时期的平均  $OR^2$  (分别为  $-2.66\%$ 、 $-1.28\%$ ),说明账面市场比组合和市值组合在熊市拥有更强的预测性. 账面市值比组合中,  $BM5$  和  $BM4$  在熊市阶段  $OR^2$  最高,分别为  $5.61\%$  和  $5.15\%$ ;  $BM2$ 、 $BM3$ 、 $BM4$ 、 $BM5$ 、 $BM7$ 、 $BM9$  在熊市阶段的预测性也较牛市阶段强. 市值组合中,  $SIZE8$  在熊市阶段  $OR^2$  最高,为  $14.05\%$ ; 而  $SIZE9$  在牛市时期的  $OR^2$  也高达  $8.23\%$ .

综上,总体市场和行业投资组合在牛市可预测性更强,而平均来看账面市值比和市值投资组合在熊市拥有更强的预测性,该结果与美国和印度市场的检验结果一致<sup>[47, 50]</sup>,并且根据样本内外预测结果发现高账面市值比组合、小市值组合(特别是样本外)更容易被预测,这一结论与 Kong 等<sup>[58]</sup> 和姜富伟等<sup>[48]</sup> 一致.

### 4.3 预测性的经济学解释

#### 4.3.1 样本外预测性分解

表7的C部分给出了基于条件CAPM模型的理性可预测性和Alpha可预测性. 在全部的投资组合中,共有20个投资组合(6个行业组合、7个账面市值比组合、7个市值组合)的理性可预测性  $OR_r^2$  为正,即这些投资组合可被条件CAPM约束模型成功预测. 而这20个投资组合中,扣除  $OR_a^2$  为负有的17个投资组合(5个行业组合、6个

账面市值比组合、6个市值组合),仅有能源行业、 $BM7$ 、 $SIZE4$  具有Alpha可预测性. 能源行业是唯一拥有Alpha可预测性的行业,可能是由于近几年该行业存在泡沫,导致股价的非理性波动,说明能源行业的收益率包含条件CAPM模型不能解释的投资者行为因素. 综上所述,大部分投资组合的可预测性可以被条件CAPM模型捕捉的时变系统风险溢价充分解释.

#### 4.3.2 样本外预测性与行业集中度

根据信息流动摩擦理论<sup>[55]</sup>,股票收益率的预测程度部分地受到信息摩擦的驱动,而信息摩擦的一个来源就是行业集中程度. 对于集中程度高的行业,投资者很容易获得该行业当中少数的几家大型公司的完整信息,因此可以推测集中程度高的行业收益率情况往往不易被预测. 而对于集中度较低的行业,市场上存在大量的中小型公司,投资者为获取这样的信息将付出更大的成本,信息流动摩擦高,因此可以推测集中度低的行业收益率情况更容易被预测. Kong 等<sup>[56]</sup>、Narayan 和 Bannigidmath<sup>[50]</sup> 已分别证实了美国、中国、印度股票市场上行业集中度与各行业收益率的可预测性负相关. 对于行业集中度,常用指标有  $CR_n$  指数、Herfindahl-Hirschman 指数.

对于  $CR_n$  指数,取各个行业内收益排名前4名的公司占该行业总收益的百分比. 通过计算整个样本预测期内行业集中度的均值,并将其与样本外预测性统计量  $OR^2$  进行线性回归,得

$$OR^2 = -0.07IC + 3.15 \tag{21}$$

其中线性系数的  $t$  统计量为  $-1.99$ ,在  $10\%$  的水平下显著,  $R^2$  统计量为  $33.04\%$ .

对于 Herfindahl-Hirschman 指数,取行业内各个公司收益占比的平方和. 将其与样本外预测性统计量  $OR^2$  进行线性回归,得

$$OR^2 = -1.98IC + 10.35 \tag{22}$$

其中线性系数的  $t$  统计量为  $-2.21$ ,在  $10\%$  的水平下显著,  $R^2$  统计量为  $37.82\%$ ,显著性相较于  $CR_n$  指数有所提升.

综上,两种行业集中度指标均说明我国行业样本外预测性与行业集中度确实呈现显著的负相关关系,不同行业样本外预测性的差异可由信息流动摩擦解释.

表 7 样本外检验结果及其经济学解释

Table 7 Results of out-of-sample prediction and their rational and Alpha components

	A: 综合预测收益率 样本外检验			B: 牛熊市样本外检验			C: 经济学解释	
	<i>RMAE</i>	<i>RRMSE</i>	<i>OR</i> <sup>2</sup>	熊市	牛市	全体	<i>OR</i> <sub>R</sub> <sup>2</sup> (%)	<i>OR</i> <sub>α</sub> <sup>2</sup> (%)
A: 市场及分行业投资组合								
总体市场	0.983	0.987	0.034	1.25	11.76	3.44		
能源	0.983	0.971	0.033	5.75	13.60	3.32	3.31	0.01
材料	0.975	0.964	0.049	-3.14	12.46	4.90	4.94	-0.04
工业	0.997	0.991	0.006	-0.22	9.20	0.63	1.75	-1.14
可选消费	0.991	0.991	0.019	-1.54	9.58	1.87	-0.07	1.94
日常消费	0.962	0.971	0.074	0.01	12.65	7.38	10.43	-3.40
医疗	1.012	0.985	-0.024	-11.69	-0.85	-2.35	-1.46	-0.88
公共事业	1.008	1.006	-0.017	-1.55	1.95	-1.70	-1.73	0.03
金融	0.993	1.008	0.015	0.09	3.37	1.49	5.45	-4.19
电信	1.001	1.001	-0.001	3.96	6.65	-0.13	3.27	-3.51
信息技术	1.006	1.001	-0.013	3.19	5.57	-1.28	-1.62	0.33
平均				-0.51	7.42	1.41		
B: 账面市值比分组投资组合								
<i>BM1</i>	1.000	0.999	0.001	-1.68	2.26	0.11	-3.98	3.93
<i>BM2</i>	1.027	1.018	-0.055	1.06	-12.83	-5.52	0.69	-6.25
<i>BM3</i>	0.998	1.001	0.004	2.10	-2.66	0.40	1.57	-1.18
<i>BM4</i>	0.990	0.998	0.020	5.15	-1.74	1.96	2.13	-0.17
<i>BM5</i>	1.010	0.992	-0.021	5.61	-16.21	-2.06	0.02	-2.08
<i>BM6</i>	1.027	1.018	-0.054	-7.38	-2.60	-5.38	2.53	-8.12
<i>BM7</i>	0.989	0.994	0.023	1.96	2.51	2.28	0.62	1.67
<i>BM8</i>	1.006	1.003	-0.012	-2.96	0.81	-1.23	-0.29	-0.94
<i>BM9</i>	0.989	0.992	0.023	3.21	1.10	2.29	6.04	-3.99
<i>BM10</i>	0.996	0.997	0.008	-0.58	2.70	0.76	-4.80	5.30
平均				0.65	-2.66	-0.64		
C: 市值分组投资组合								
<i>SIZE1</i>	1.008	1.002	-0.017	-3.12	0.07	-1.67	4.44	-6.39
<i>SIZE2</i>	1.009	1.007	-0.018	-0.48	-3.65	-1.76	3.98	-5.98
<i>SIZE3</i>	0.994	0.999	0.011	0.99	0.81	1.11	1.87	-0.78
<i>SIZE4</i>	0.996	0.994	0.009	3.08	-3.35	0.88	0.74	0.14
<i>SIZE5</i>	1.013	1.009	-0.027	-6.97	2.48	-2.69	6.71	-10.08
<i>SIZE6</i>	0.998	0.998	0.004	1.23	-8.73	0.40	1.28	-3.31
<i>SIZE7</i>	1.017	1.009	-0.034	-6.50	1.34	-3.41	4.51	-8.29
<i>SIZE8</i>	0.978	0.987	0.044	14.05	-12.71	4.41	-1.60	5.92
<i>SIZE9</i>	0.964	0.975	0.071	5.70	8.23	7.11	-4.72	11.30
<i>SIZE10</i>	0.993	0.997	0.014	1.02	2.67	1.44	-3.12	4.43
平均				0.90	-1.28	0.58		

## 5 结束语

基于中国沪深 A 股市场,本文构建了包括市场组合、行业组合、账面市值比组合和市值组合在内的 31 个投资组合,通过选取 8 个预测因子(账面市值比、股利分配率、股息价格比、股息收益率、每股收益价格比、现金收益价格比、通货膨胀率、股票波动率),分别对不同投资组合超额收益率的可预测性进行了样本内与样本外检验。通过比较检验结果,所选因子对不同投资组合的收益率具备一定的预测能力。*CFP*、*DE*、*DP*、*DY* 等大部分因子样本内外的预测结果具有一致性,而且样本外的预测精度一般要劣于样本内的预测精度。当然也存在一些例外,*BM* 因子样本内的预测结果要大大优于样本外的预测结果,而 *INF* 因子和 *SVAR* 因子则完全相反。在此基础上,本文还检验了不同经济周期下综合预测收益率的样本外预测性,发现总体市场和行业组合的收益率在牛市阶段的预测性远强于熊市,而账面市值比组合和市值组合的结果则完全相反。

本文还进一步研究了不同投资组合样本外预

测性的经济学解释。通过利用条件 CAPM 资产定价模型把样本外统计量  $OR^2$  分解为基于条件资产定价模型的理性可预测性  $OR_R^2$  和 Alpha 可预测性  $OR_\alpha^2$ ,发现大部分投资组合的可预测性可以被条件 CAPM 模型捕捉的时变系统风险溢价充分解释,只有能源行业、*BM7*、*SIZE4* 具有 Alpha 可预测性。在此基础上,根据信息流动摩擦理论,本文还研究了行业集中度是否能够解释行业组合样本外预测性的差异,发现我国行业组合样本外预测性与行业集中度确实呈现显著的负相关关系,行业组合样本外预测性的差异可由信息流动摩擦解释。

本文以因子定价模型为基础,从研究方法、因子构造、因子选择等方面拓展了 Fama 和 French 因子模型在中国的适用性,提升了投资者对我国股市发展现状的了解,为理性投资决策提供科学依据,同时也为促进我国量化投资的蓬勃发展乃至整个资本市场的良性运转提供了重要的实践指导。本文的研究不但得到了一些与美国、印度股市研究相一致的结论,而且还发现了某些因子(如 *BM*、*DE*、*CFP*) 在中国股市上的表现完全不同于其他市场,后续可进一步选取有中国特色的因子或新方法做更为深入研究。

## 参考文献:

- [1]Fama E F, French K R. Common risk factors in the returns on stocks and bonds[J]. Journal of Financial Economics, 1993, 33(1): 3-56.
- [2]Fama E F, French K R. A five-factor asset pricing model[J]. Journal of Financial Economics, 2015, 116(1): 1-22.
- [3]Fama E F, French K R. Business conditions and expected returns on stocks and bonds[J]. Journal of Financial Economics, 1989, 25(1): 23-49.
- [4]Campbell J Y, Shiller R J. Stock prices, earnings, and expected dividends[J]. The Journal of Finance, 1988, 43(3): 661-676.
- [5]Banz, Rolf W. The relationship between return and market value of common stocks[J]. Journal of Financial Economics, 1981, (9): 3-18.
- [6]Bhandari L C. Debt/equity ratio and expected common returns: Empirical and evidence[J]. Journal of Finance, 1988, 43: 507-528.
- [7]Rosenberg B, Reid K, Lanstein R. Persuasive evidence of market inefficiency[J]. Journal of Portfolio Management, 1985, 11: 9-16.
- [8]Basu S. The relationship between earnings yield, market value, and return for NYSE common stocks: Further evidence[J]. Journal of Financial Economics, 1983, (12): 129-156.
- [9]Hodrick R J. Dividend yields and expected stock returns: Alternative procedures for inference and measurement[J]. Review of Financial Studies, 1992, 5(3): 357-386.

- [10] Fama E, French K. Dividend yields and expected stock returns [J]. *Journal of Financial Economics*, 1988, (22): 3-25.
- [11] Pontiff J, Lawrence S. Book-to-market ratios as predictors of market returns [J]. *Journal of Financial Economics*, 1998, (49): 141-160.
- [12] Liu L X, Whited T M, Zhang L. Investment-based expected stock returns [J]. *Journal of Political Economy*, 2009, 117(6): 1105-1139.
- [13] Lamont O. Earnings and expected returns [J]. *The Journal of Finance*, 1998, 53(5): 1563-1587.
- [14] Zhang W, Wang G, Wang X, et al. Profitability of reversal strategies: A modified version of the Carhart model in China [J]. *Economic Modelling*, 2018, 69: 26-37.
- [15] 俞红海, 陈百助, 蒋振凯, 等. 融资融券交易行为及其收益率可预测性研究 [J]. *管理科学学报*, 2018, 21(1): 72-87.
- Yu Honghai, Chen Baizhu, Jiang Zhenkai, et al. Margin trading and stock return predictability [J]. *Journal of Management Sciences in China*, 2018, 21(1): 72-87. (in Chinese)
- [16] 罗进辉, 向元高, 金思静. 中国资本市场低价股的溢价之谜 [J]. *金融研究*, 2017, 439(1): 191-206.
- Luo Jinhui, Xiang Yuangao, Jin Sijing. The puzzle of low price premium effects: Evidence from China [J]. *Journal of Financial Research*, 2017, 439(1): 191-206. (in Chinese)
- [17] Lawrenz J, Zorn J. Predicting international stock returns with conditional price-to-fundamental ratios [J]. *Journal of Empirical Finance*, 2017, 43: 159-184.
- [18] French K R, Schwert G W, Stambaugh R F. Expected stock returns and volatility [J]. *Journal of Financial Economics*, 1987, 19(1): 3-29.
- [19] Campbell J Y. Stock returns and the term structure [J]. *Journal of Financial Economics*, 1987, 18(2): 373-399.
- [20] Lettau M, Ludvigson S. Consumption, aggregate wealth, and expected stock returns [J]. *The Journal of Finance*, 2001, 56(3): 815-849.
- [21] Chen J P, Xiong X, Zhu J, et al. Asset prices and economic fluctuations: The implications of stochastic volatility [J]. *Economic Modelling*, 2017, 64: 128-140.
- [22] Brock W, Lakonishok J, LeBaron B. Simple technical trading rules and the stochastic properties of stock returns [J]. *The Journal of Finance*, 1992, 47(5): 1731-1764.
- [23] Jensen M C. Random walks: Reality or myth: Comment [J]. *Financial Analysts Journal*, 1967, 23(6): 77-85.
- [24] Neely C J, Rapach D E, Tu J, et al. Forecasting the equity risk premium: The role of technical indicators [J]. *Management Science*, 2014, 60(7): 1772-1791.
- [25] Wang S, Jiang Z Q, Li S P, et al. Testing the performance of technical trading rules in the Chinese markets based on superior predictive test [J]. *Physica A*, 2015, 439: 114-123.
- [26] Lin Q. Technical analysis and stock return predictability: An aligned approach [J]. *Journal of Financial Markets*, 2018, 38: 103-123.
- [27] 吴世农, 许年行. 资产的理性定价模型和非理性定价模型比较研究——基于中国股市的实证分析 [J]. *经济研究*, 2004, (6): 105-116.
- Wu Shinong, Xu Nianhang. A Comparative study on the rational asset pricing model and irrational asset pricing model: Evidence from stock market in China [J]. *Economic Research Journal*, 2004, (6): 105-116. (in Chinese)
- [28] 邓长荣, 马永开. 三因素模型在中国证券市场的实证研究 [J]. *管理学报*, 2005, 2(5): 591-596.
- Deng Changrong, Ma Yongkai. Empirical research of the three-factor model in Chinese stock market [J]. *Chinese Journal of Management*, 2005, 2(5): 591-596. (in Chinese)
- [29] 陈信元, 张田余. 预期股票收益的横截面多因素分析: 来自中国证券市场的经验证据 [J]. *金融研究*, 2001, (6): 22-35.
- Chen Xinyuan, Zhang Tianyu. A cross-sectional multifactor analysis of expected stock returns: Empirical evidence from China stock market [J]. *Journal of Finance*, 2001, (6): 22-35. (in Chinese)



- [30]范龙振,王海涛. 上海股票市场股票收益率因素研究[J]. 管理科学学报,2003,6(1):60-67.  
Fan Longzhen, Wang Haitao. Study on the factors that affect average returns in SSE[J]. Journal of Management Sciences in China,2003,6(1):60-67. (in Chinese)
- [31]刘昱熙,宋旺. 股票回报与公司规模、市净率、市盈率之间的关系——基于中国A股市场的实证论证[J]. 财会通讯,2011,(6):85-86,96.  
Liu Yuxi, Song Wang. The relationship between stock returns and firm size, book value and P/E ratio: An empirical study based on China's A-share market[J]. Communication of Finance and Accounting,2011,(6):85-86,96. (in Chinese)
- [32]潘莉,徐建国. A股市场的风险与特征因子[J]. 金融研究,2011,(10):140-154.  
Pan Li, Xu Jianguo. Risk and characteristics factors in China's A-share stock returns[J]. Journal of Financial Research,2011,(10):140-154. (in Chinese)
- [33]苏宝通,陈炜,陈浪南. 公开信息与股票回报率相关性的实证研究[J]. 管理科学,2004,17(6):67-75.  
Su Baotong, Chen Wei, Chen Langnan. An empirical study of the relationship between public information and market activities[J]. Journal of Management Sciences,2004,17(6):67-75. (in Chinese)
- [34]王茵田,朱英姿. 中国股票市场风险溢价研究[J]. 金融研究,2011,(7):152-166.  
Wang Yintian, Zhu Yingzi. On the cross-sectional risk premium of Chinese stock market[J]. Journal of Financial Research,2011,(7):152-166. (in Chinese)
- [35]Nelson C R, Kim M J. Predictable stock returns: The role of small sample bias[J]. The Journal of Finance,1993,48(2):641-661.
- [36]Stambaugh R F. Predictive regressions[J]. Journal of Financial Economics,1999,54(3):375-421.
- [37]Cavanagh C L, Elliott G, Stock J H. Inference in models with nearly integrated regressors[J]. Econometric Theory,1995,11(5):1131-1147.
- [38]田利辉,王冠英. 我国股票定价五因素模型:交易量如何影响股票收益率?[J]. 南开经济研究,2014,(2):54-75.  
Tian Lihui, Wang Guanying. Asset pricing model of the Chinese stock market: How trading volumes influence the returns?[J]. Nankai Economic Studies,2014,(2):54-75. (in Chinese)
- [39]欧阳志刚,李飞. 四因子资产定价模型在中国股市的适用性研究[J]. 金融经济研究,2016,(2):84-96.  
Ouyang Zhigang, Li Fei. The applicability of the four-factor asset pricing model in Chinese stock market[J]. Journal of Finance and Economics,2016,(2):84-96. (in Chinese)
- [40]Clive W J, Granger I V, Namwon Hyung, et al. Spurious regressions with stationary series[J]. Applied Economics,2001,33(7):899-904.
- [41]刘汉中. 稳健的 HAC 法在多元平稳时间序列之间伪回归中的应用[J]. 数量经济技术经济研究,2011,(10):134-147.  
Liu Hanzhong. The application of robust HAC method in spurious regression under multivariate stationary time series[J]. The Journal of Quantitative and Technical Economics,2011,(10):134-147. (in Chinese)
- [42]Westerlund J, Narayan P K. Does the choice of estimator matter when forecasting returns?[J]. Financial Econometrics,2012,36(9):2632-2640.
- [43]Westerlund J, Narayan P. Testing for predictability in conditionally heteroskedastic stock returns[J]. Journal of Financial Econometrics,2014,13(2):342-375.
- [44]Lewellen J. Predicting returns with financial ratios[J]. Journal of Financial Economics,2004,74(2):209-235.
- [45]Rapach D E, Strauss J K, Zhou G. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy[J]. Review of Financial Studies,2010,23(2):821-862.
- [46]Rapach D E, Wohar M E. In-sample vs. out-of-sample tests of stock return predictability in the context of data mining[J]. Journal of Empirical Finance,2013,13(2):231-247.
- [47]Campbell J Y, Thompson S B. Predicting excess stock returns out of sample: Can anything beat the historical average?

- [J]. *Review of Financial Studies*, 2008, 21(4): 1509 – 1531.
- [48] 姜富伟, 徐俊, Rapach D E, 等. 中国股票市场可预测性的实证研究[J]. *金融研究*, 2011, (9): 107 – 121.  
Jiang Fuwei, Xu Jun, Rapach D E, et al. How predictable is the Chinese stock market? [J]. *Journal of Financial Research*, 2011, (9): 107 – 121. (in Chinese)
- [49] Welch I, Goyal A. A comprehensive look at the empirical performance of equity premium prediction [J]. *Review of Financial Studies*, 2008, 21(4): 1455 – 1508.
- [50] Narayan P K, Bannigidadmath D. Are Indian stock returns predictable? [J]. *Journal of Banking & Finance*, 2015, 58: 506 – 531.
- [51] Welch I, Goyal A. A comprehensive look at the empirical performance of equity premium prediction [J]. *Review of Financial Studies*, 2008, 21(4): 1455 – 1508.
- [52] Butler A W, Grullon G, Weston J P. Can managers forecast aggregate market returns? [J]. *The Journal of Finance*, 2005, 60(2): 963 – 986.
- [53] 何兴强, 周开国. 牛、熊市周期和股市间的周期协同性[J]. *管理世界*, 2006, (4): 35 – 40.  
He Xingqiang, Zhou Kaiguo. Bull, bear market cycle and the cyclical synergy between the stock market [J]. *Management World*, 2006, (4): 35 – 40. (in Chinese)
- [54] Kadan O, Madureira L, Wang R, et al. Analysts' industry expertise [J]. *Journal of Accounting and Economics*, 2012, 54(2): 95 – 120.
- [55] Hong H, Torous W, Valkanov R. Do industries lead stock markets? [J]. *Journal of Financial Economics*, 2007, 83(2): 367 – 396.
- [56] Kong A, Rapach D, Strauss J, et al. How predictable are components of the aggregate market portfolio [EB/OL]. Available at: <http://se.shufe.edu.cn/upload/htmleditor/File/090601014628.pdf>

## Return predictability in the Chinese stock markets

JIANG Zhi-qiang, TIAN Jing-wen, ZHOU Wei-xing\*

School of Business, East China University of Science and Technology, Shanghai 200237, China

**Abstract:** This paper investigates the return predictability to understand the theoretical foundation of quantitative investments in Chinese stock markets. Eight factors (book-to-market ratio, dividend-payout ratio, dividend-price ratio, dividend-yield ratio, earnings-price ratio, cash flow-price ratio, inflation rate, stock variance) are employed to carry out in-sample predictability tests by means of the feasible generalized least squares method (FQGLS). Out-of-sample predictability tests are conducted in different economic cycles. It is found that: 1) The portfolios can be successfully predicted and the predictability is dependent on the cycle and the portfolio type; 2) The predictability of most portfolios can be explained by the conditional CAPM model; and 3) The return predictability in industries is significantly negatively correlated with industry concentration.

**Key words:** feasible generalized least squares method; in-sample test; out-of-sample test; conditional CAPM; industry concentration