**Fama-French五因子模型**

**在中国股票市场的实证检验**

181098273 王天诚 181098253 屠心怡

181250057 姜春妮 181860018 冯启航

**摘 要:**本文基于Fama and French(2015a)提出的五因子模型，以我国A股市场以 2014年1 月至2018年6月A股上市公司的交易和财务数据为样本，对五因子模型在中国股票市场的适用性及有效性进行探究，并分析了背后潜在的原因。此外，基于线上量化平台——聚宽的数据库，我们搭建了基于五因子模型的量化投资选股策略，在A股市场进行了回测，并对策略的回测表现进行了分析。研究结论表明，五因子模型在我国股票市场上有较强的解释能力，市值效应和价值效应明显而投资风格效应较弱。

**关键词: 因子模型; 投资风格;** **资产定价**

1. **引言**

股票投资管理是资产管理的重要组成部分之一，其目标就是实现效用最大化，即使股票投资组合的风险和收益特征能够给投资者带来最大的满足。因此，构建股票投资组合的原因有二：一是为降低证券投资风险；二是为实现证券投资收益最大化。组合管理是一种区别于个别资产管理的投资管理理念。组合管理理论最早由Markowitz(1952)系统地提出，他开创了对投资进行整体管理的先河。构建投资组合并分析其特性是职业投资组合经理的基本活动。在构建投资组合过程中，就是要通过证券的多样化，使由少量证券造成的不利影响最小化。

股票投资组合的收益率由何种因素决定，这个经典的问题经久不衰，是资产定价领域研究中的重要议题，驱使着一代又一代金融学者投身该领域进行探索,最早的资本资产定价模型——CAPM模型，其实在1961年就有了雏形。股票投资实践经验和对交易数据的分析能够丰富和完善定价理论，某些投资策略长期获得超额收益促使学者们不断完善理论体系，修改定价模型，从而提升资本市场的定价效率并形成交易策略。从CAPM(Sharpe, Linter, and Mossin)，FF-3(Fama and French)，Carhart-4(Fama, French, and Carhart)，到FF-5(Fama and French)，用来检验和被检验的模型层出不穷：从单纯考虑市场风险因素，发展到综合考虑公司的各方面，包括规模、成长性、股票流动性、动量效应和盈利能力等；从简单的单因素模型，发展到了多因子模型。

Fama and French的三因子模型深入人心，无论是在学术界还是投资界，都得到广泛的应用，被奉为因子投资的开山鼻祖。特别是美国市场，三因子模型被广泛地应用于收益率预测、风险管理、基金业绩评价等各个方面。Fama and French随后基于红利贴现模型的讨论框架提出五因子模型，在三因子模型基础上加入盈利能力和投资风格因子，以更好地描述横截面上股票组合的期望收益率。

五因子模型的适用性已经在国际市场得到检验(Fama and French, 2015b)，对五因子模型在中国证券市场的应用进行研究，具有重要的现实意义。资产定价模型的解释能力因资本市场发展水平而异，但目前学术界大多是以美国的股票市场的数据为样本进行分析的。然而对于新兴资本市场，投资者交易理念、交易所成交制度，信息环境等与成熟股票市场存在较大差异，这些因素通过交易过程影响会资产定价机制在该市场的有效性。五因子模型的解释和预测能力是否存在市场差异，公司经营因素是否能解释我国股票市场投资组合的平均回报，HML是否成为冗余变量，这都需要分析和验证。

本文基于Fama and French(2015a)提出的五因子模型，结合我国A股市场的交易和财务数据，对五因子模型在中国股票市场的适用性及有效性进行探究，并分析了背后潜在的原因。此外，基于线上量化平台——聚宽的数据库，我们搭建了基于五因子模型的量化投资选股策略，在A股市场进行了回测，并对策略的回测表现进行了分析。研究结果对于产定价模型的理论分析和实证检验，乃至个人的量化投资策略的选择与构建，均具有一定的借鉴意义。

本文分为七个部分，第二部分是文献回顾，第三部分为数据处理与因子构建方法，第四部分是五因子模型在中国 A 股市场的适用性考察，第五部分为外部数据补充验证，第六部分为结论，第七部分为参考文献。

1. **文献回顾**

自Markowitz(1952)正式提出投资组合理论，Fama and Malkiel(1970)正式提出有效市场假说(EMH)以来，有关资产定价模型的研究得到了长足发展。定价模型的有效性对于投资决策具有重要意义，关于资产定价模型的理论研究和实证研究汗牛充栋，从简单的单因素模型发展到五因子模型，从单纯的考虑市场风险因素发展到综合考虑公司的规模、成长性、股票流动性、动量效应和盈利能力等因素。

Sharpe(1964)、Lintner(1965)和 Mossin(1966)提出了CAPM模型，在一般均衡的基础上，该模型奠定了资产的收益来源于资产所承担的风险这一基本理念，成为现代金融市场价格理论的支柱。CAPM 模型以单一的风险因素描述资产的收益率使其受到多方面的挑战。一方面，CAPM 模型无法解释市值效应(size effect)，即股票收益率和流通市值显著的负相关关系(Banz,1981；Reinganum,1981;Keim,1983)。Banz (1981)发现市场风险溢价的解释能力很弱，而公司的市值对股票收益率的解释能力更强，这种现象在世界上许多主要的交易所屡见不鲜(Hawawini和Keim2000)。另一方面，CAPM 无法解释价值效应(value effect)，即账面市值比与股票收益率显著的正相关关系(Stattman，1980；Chan et al.，1991)。Fama 和 French(1992)、Lakonishok 等(1994)以及 Davis(1994)发现公司当前的股价、所有者权益账面价值和每股收益能够解释公司股票未来的回报率。为了提高模型的解释和预测能力，资产定价模型由单因子模型逐步向多因子模型发展。

Fama-French(1993，1996)的三因子模型结合了此前的研究，在 CAPM 的基础上加入了解释市值效应的因子 SMB 和解释价值效应的因子 HML。该模型的重要特点是，虽然 SMB、HML 代表了公司市值和公司账面市值比因素，但实际上二者代表的是不同资产组合的收益之差而非公司市值或账面市值比本身，这种因子的构建方式成为后续研究构建因子的标准方式。随后Fama and French(1998)将三因子模型扩展到全球视野，提出了全球市场风险溢价、全球市场的 SMB 和全球市场的 HML，并与使用全球市场的市场风险溢价的 CAPM 模型进行比较并发现三因子模型的解释能力更强。三因子模型引发了学界的广泛关注。Griffin(2002)用日本、英国和加拿大的数据检验了三因子模型，发现模型能够有效地解释股票组合回报的差异，但使用本国市场数据计算得到的三因子模型要比用全球市场的三因子模型解释能力更好。

Cao 等(2005)利用我国的股票交易数据对比了 CAPM 模型和 Fama-French 三因子模型，发现后者的解释能力更好。朱宝宪和何治国(2002)验证了三因子模型在我国股票市场的适用性。吴世农和许年行(2004)采用早期的数据进行研究,发现中国股市存在显著的账面市值比效应和规模效应。潘莉和徐建国(2011)发现对我国市场来说用市盈率代替市净率的三因子模型具有较好的解释能力。田利辉等(2014)认为A股市场存在显著的规模效应,但账面市值比效应并不显著。

股价走势常常表现出趋势和惯性。Jegadeesh和Titman(1993)发现前期股票的收益率有助于解释当期股票的收益率,使用动量策略,即构建一个投资组合,买入过去股价表现较好的公司,卖出过去股价表现较差的公司能够获得显著的回报,这种现象被称为动量效应。Carhart(1997)将过去股价走势较好的“赢家”组合和过去股票走势较差的“输家”组合的收益率之差作为动量因子WML (winners minus losers)并将其加入到三因子模型中构建了四因子模型。Jegadeesh 和 Titman(2001)验证了动量策略的收益率并发现动量因子有助于解释股票组合收益的差异。这种现象在各国交易所屡见不鲜。Griffin(2003)发现全球多个交易所都存在动量效应,而且动量策略与宏观经济形势无关,在经济周期的各个环节下动量策略都能取得一定收益。L'Her 等(2004)研究了加拿大股市,Lam(2010)研究了香港股市,Gregory等(2013)研究了英国股市,都发现了动量效应,证实了Carhart(1997)四因子模型的有效性。

动量策略也存在诸多争议。Novy-Marx (2012)提出了一个与直觉相悖的观点,认为依据过去7~12个月的表现构造投资组合比依据近期的表现构造投资组合收益率更高。Barroso和Santa-Clara(2015)研究发现动量效应非常不稳定,随时间变化剧烈波动,动量策略由于收益不稳定使投资者敬而远之。此外, Cohen等(2002)认为,动量效应的产生是由机构投资者在市场上持续从其他投资者手中购买被低估的股票引起的,本质上是投资者对企业价值判断。

就我国来说,动量效应是否存在,是否稳定,是否与其他因子相关存在诸多争议。郑方镳等(2007)发现,无论是牛市、熊市还是平衡市,我国股票收益率在随后的交易中多表现为“反转”现象,原因在于中国投资者的资产配置交易和过度投机交易行为。舒建平等(2012)认为,我国动量效应与反转效应存在类似的周期性相互演化规律，但演化的周期长度并不固定,即动量效应是时变的。无独有偶,高秋明等(2014)发现我国A股市场不存在显著的月度频率上的动量效应,而当形成期为2~4周,持有期为1 ~3周时则存在稳定的动量收益,规模和账面市值比和行业等因素可以解释50%的动量效应。因此,对于我国来说对动量效应是否作为一个定价因子存在争议。

流动性衡量了股票交易的速度和成本,Amihud (2002)研究发现流动性可以解释股票的收益。Pastor和Stambaugh(2003)在四因子模型的基础上加入了衡量市场流动性风险的因素作为第五个因子并用美国股票市场进行了验证,发现市场的流动性风险影响了股票组合的回报。Chan和 Faff(2005)在三因子模型的基础上引入了换手率作为流动性因子研究了澳大利亚股票市场,发现流动性风险能够带来价值。Nguyen和Tribhuvan(2009)研究发现流动性能够解释股票回报率,并认为三因子模型中的市值因子实际上代表了流动性风险。陈青和李子白(2008)在CAPM的基础上采用Amihud(2002)的方法构建了流动性因子调整的CAPM模型。

我国学者结合我国市场的特点,对流动性因素对资产组合收益率的解释能力存在一些争议。郑振龙和汤文玉(2011)认为,波动率能够衡量股票的流动性,是影响股票收益率的重要风险因素,且在一定程度上能够解释市值效应和价值效应。苏冬蔚和麦元勋(2004)用换手率代表流动性对我国股票市场股票组合的预期收益率进行了实证分析。田利辉和王冠英(2014)将换手率和成交量作为两个因子与三因子模型结合提出了研究我国股市的五因子模型。实际上换手率和成交量反映了投资者的异质信念,即对股票价值判断的分歧,并不严格等同于股价流动性。张峥和刘力(2006)研究了我国股票市场股票换手率与横截面股票收益间的负相关关系,认为流动性溢价理论并不能完全解释这种异象,投机性交易造成的投机性泡沫是更为合理的解释。流动性是否适合作为我国资产定价因素有待商榷。

理论上，公司的价值取决于未来股权现金流的现值，从实践经验来看，盈利能力强的公司的股票往往回报率较高,而高投资水平的公司股票回报率较低。许多学者从公司经营角度考虑资产定价问题。Fama和 French(2006,2008)用股利贴现模型说明了预期收益、预期投资和账面市值比对股票未来收益具有解释和预测能力。Cooper 等(2008)发现公司的总资产增长率可以预测美国股票收益率并在统计意义上和经济意义上都显著。Novy-Marx (2013)认为,总资产利润率(ROA)对截面数据的公司回报的解释能力与账面市值比不分伯仲,在使用价值投资策略时考虑ROA将提高策略收益。除盈利能力之外,投资水平也是影响资产回报率的重要因素。Arharoni 等(2013)认为,高投资水平和高投资回报率增加公司的未来总资产,但高投资水平会降低当前的总资产,因而公司的投资决策影响公司的价值。Titman等(2004)发现投资对公司股票回报有明显的负影响,尤其是现金流充裕和负债比率低的公司。Liu等(2009)发现公司股票的回报率与杠杆投资的回报率相等,公司的回报率与意外盈余、账面市值比和资本投资有关。Cooper和Priestley(2011)认为,投资水平高和投资水平低的公司组合回报率的差异源于系统风险,系统风险在投资水平高时较低,对投资组合回报具有较好的解释能力。综合前述研究,投资和盈利因素成为后续资产定价模型的因子。Chen等(2011)基于MM公式提出包含市场超额收益、投资和净资产收益率三个因素的模型,与 Fama和 French(2015)的区别在于因子计算方法不同。Chen等(2011)用计算特殊项目前的净收益占总资产的比率衡量盈利能力,用固定资产和存货占总资产的比重衡量投资,而 Fama-French五因子模型分别用营业利润占总资产比和总资产上一年度增长率衡量利润和投资。Nichol和Dowling(2014)用英国股票对比了二者的解释和预测能力,发现 Fama 和 French(2015)模型更胜一筹。Hou等(2015)提出了新的四因子模型,包含了市场超额收益、规模、投资和权益净利率四个因子。与 Fama 和 French五因子的差异在于Hou等(2015)模型不包含价值因素 HML,投资因素用当年的总资产增长率衡量,盈利能力用权益净利率衡量。

当前对公司盈利能力和投资水平的研究成为资产定价模型的主要发展方向,但结合盈利因素和投资因素对新兴市场的研究尚不多见。目前国内学术界的研究大多停留在对三因子模型的实证检验，且近年来我国经济和投资环境变化较快,股市规律也在不断演变，早期研究的数据也亟待更新。

因此，本文基于五因子定价模型，结合A股的交易和财务数据，研究我国市场的盈利因素和投资因素是否能解释和预测股票的回报率，因子定价模型的作用是否因市场而异，对这一领域的研究形成补充。并且，本文基于聚宽平台，搭建了基于五因子模型的量化投资选股策略，在A股市场进行了回测分析。为我国的资产定价模型提供经验证据。

1. **数据处理与因子构建方法**

**(一)样本数据选取与处理**

本文选取的数据由俞红海老师提供，实际来源为CSMAR国泰安数据库，其数据源为对交易所授权数据和上市公司权威公开资料的整合。

考虑到我国境内机构、组织、个人一般主要参与 A 股交易，本文的股票样本只选取 A 股股票。中国股市早期股票数量少、波动大、信息反映程度较差(田利辉等，2014)，多数国内学者进行相关研究时都选择剔除掉早期的数据。1994 年之前的中国股市数据不足以完成 Fama and French(2015a)中对因子的分组方法，因此本文选取 A 股上市公司2014年1月-2018年6月共54个月的月度数据作为样本。与大多数学术文献一致，本文的无风险利率采用中国人民银行公布的人民币三个月整存整取利率，而月度个股收益率采用数据库中考虑现金红利再投资的月度收益率。财务数据均选择合并报表后数据。

文章对数据进行了如下处理:

1. 更改时间周期t为月

按照字段说明.txt文件中的要求，我们初期尝试确定周期t的单位为年，选用年度数据进行实证检验。但由于只有2014-2018年的数据，在计算各因子后发现仅有2年数据可用，理论上无法应用五因子模型进行回归。在广泛查阅市场上实际用于投资策略搭建的Fama-French五因子模型后，我们发现业界使用的时间周期最大就是月，实际多以天、小时、分钟作为周期。因此，在完成年度数据的处理后，我们决定对其进行数据加工，更改时间周期t为月，并以此计算各个因子。

这种周期选取，以及因子计算的方法与字段说明.txt文件中的要求相悖，然而是我们在撰写本文中为应对数据不足所能采取最佳策略。

2. 月数据的补充

在老师提供的原始数据中，balance sheet与income statement只有1、3、6、9、12月的数据，mktmnth、rf、stockmnth中有所有月份的数据。因此，为了得到所有月份的数据，我们需要对balance sheet 和income statement中的数据进行处理，补全缺失月的数据。

我们尝试弃用老师提供的原数据，去国泰安数据库下载更为完整的数据来完成实证检验。但由于数据源不同，会导致老师和助教难以给出公平正当的分数。

另外，我们曾考虑对缺失数据的月份进行线性插值，以补为完整的月度连续时间序列。但由于缺失数据量过大，线性插值法较为复杂，花费时间较长，并且线性插值与市场的实际情况偏离较大，因此弃用该方法。

因此，我们补充的方法是，对于缺失月份t, 我们取过去时间里最接近t的已知月份数据填入，也就是t/3向下取整的月份数据填入，类似于“向下取整”。代码逻辑层面上，选取月份完整的表格（如size），与待补充的表格（如equity）做merge操作，通过在两个表格上对交易日期除以周期向下取整打tag，标记“共享”同一数据的多个交易日期，进行表格的合并，则可得到日期完整的表格。

代码段1 merge合并表格以实现月度数据补充

1. **size ['tag'] = (size['ym'].astype(int) / 3).astype(int)**
2. **equity['tag'] = (equity['ym'].astype(int) / 3).astype(int)**
3. **bm\_merge = pd.merge(size, equity, how = 'inner', left\_on=['Stkcd', 'tag'], right\_on=['Stkcd', 'tag'])**

同时，文章对数据格式有如下约定:

表1 数据格式要求

|  |  |
| --- | --- |
| 数据名称 | 数据格式 |
| 股票代码(Stkcd) | 6位str类型，不足用0 补全 |
| 日期(ym) | 6位int类型，即年份乘以100加上月份 |

**(二)收益率及因子计算方法**

Fama and French(2015a)提出的五因子模型如下:

其中是投资组合i在第t期的收益率，是无风险利率，是市值加权平均的市场组合收益率，反映市场风险溢价。是市值因子，小市值与大市值股票组合的收益率之差。是账面市值比因子，高账面市值比与低账面市值比股票组合的收益率之差。是盈利能力因子，营运利润率高与营运利润率低的股票组合收益率之差。是投资风格因子，投资风格保守与投资风格激进的股票组合收益率之差。

表2 各因子含义

|  |  |
| --- | --- |
| 因子名称 | 因子含义 |
| 市值因子(SMB) | 小市值与大市值股票组合的收益率之差 |
| 账面市值比(HML) | 高账面市值比与低账面市值比股票组合的收益率之差 |
| 盈利能力因子(RMW) | 营运利润率高与营运利润率低的股票组合收益率之差 |
| 投资风格因子(CMA) | 投资风格保守与投资风格激进的股票组合收益率之差 |

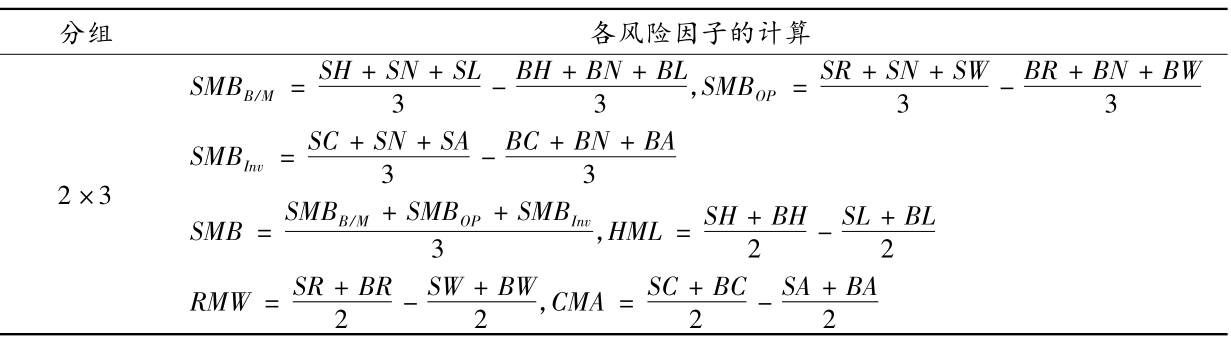
与 Fama and French的计算类似，本文在第 t 月底根据各指标的分位点来分组，用于确定第t月1日至 t +1 年30日的各股票组合。对于股票 i，以其在第 t 月底的流通市值作为指标；“账面市值比”()是用第t -1月末的账面价值，除以第 t －1 月底股票 i 的流通市值;鉴于中美会计准则存在的差异，我们直接使用“营业利润/股东权益合计”来反映 A 股市场的“营运利润率”指标；而“投资风格”( )是用第 t-1月末相对于第 t-2月末的总资产增加额，除以第 t-2月末的总资产。

表3 各指标含义

|  |  |
| --- | --- |
| 指标名称 | 指标含义 |
| 市值(Size) | 股票 i在第 t 年 6月底的流通市值 |
| 账面市值比(BM) | 第t-1年末的“账面价值/股票i的流通市值” |
| 营运利润率(OP) | 第t-1年末的“营业利润/股东权益合计” |
| 投资风格(INV) | 第t-1年末相对于第t-2年末的总资产增加额，除以第t-2年末的总资产 |

本文采用 Fama and French (2015a)提出的 3 种分组方法里的第一种来构建风险因子，以检验文章结论是否稳健。该方法是 2×3 方法，首先，按股票市值的中位数把全体股票分成小市值(S)和大市值(B)两组；按账面市值比的 30% 和 70% 分位点把全体股票分成高(H)、中(N)、低(L)三组；其次，将市值和账面市值比两个指标交叉，可把全体股票分成SH、SN、SL、BH、BN、BL 六个组合；再次，分别以营运利润率和投资风格代替账面市值比，重复上述步骤，可把全体股票分成 SR、SN、SW、BR、BN、BW、SC、SN、SA、BC、BN、BA 这 12个组合，其中 Ｒ 代表盈利稳健，W 代表盈利较弱，C 代表投资风格保守，A 代表投资风格激进，N 代表居中的盈利能力或投资风格;接下来计算上述各组合每一期的市值加权平均收益率；最后，利用不同组合收益率之差构造四个因子。

表4 因子构建方法



1. **五因子模型在中国 A 股市场适用性检验**

本文对 2014年1月至2018年6月共54个月中国 A 股市场的月度数据进行考察，以更全面地检验五因子模型的适用性。限于篇幅，仅报告基于月度数据的检验结果，而采用日度数据检验时本文主要结论并未改变。下文将报告五因子模型在 A 股市场的各项检验结果。

**（一）适用性分析：**

**1.1 五因子模型和无风险利率的累积收益率比较**

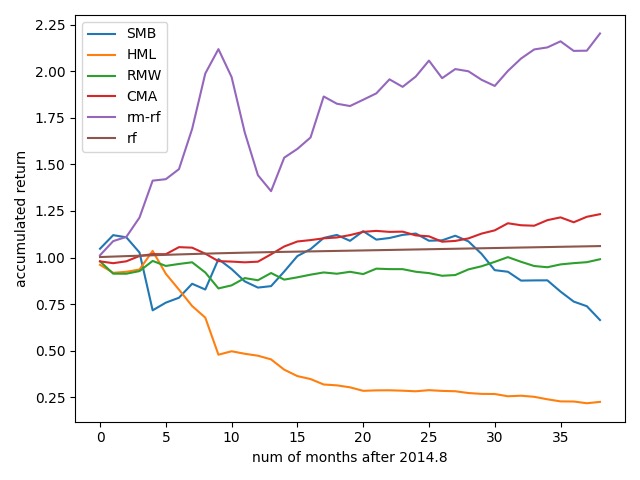


图1 五因子模型和无风险利率的累积收益率比较图

由图可知，市场超额收益和HML对组合收益的贡献最大，其他三个因子与无风险收益的差别不大，显著性检验见后文。

这一现象产生的原因可能在于：对于SMB因子，中国股市小市值股票存在借壳上市的情况，这一部分股票的市值评估存在误差，实际操作时，应排除这部分小市值股票；对于RMW因子，我国股票市场以换手率高著称，大部分股票投资者以短期投资获取股票价格波动收益为目标，而非“买入并持有”，投资者较少关注财务信息和公司经营决策，因此对该因子不敏感；对于CMA因子，我国机构投资者以基金为主，投资者结构比较单一，同质化严重。市场存在明显的“羊群效应”，因此投资风格差距不大。

从投资组合收益率的变化规律来看：

第一，我国股票市场市值效应非常明显，流通市值小的公司组合的回报高于流通市值大的公司组合的回报，而且在各组都十分明显。

第二，除了流通市值最小的20%,的公司组合以外，股票组合收益随着公司盈利能力的提高而提高，尤其是流通市值较大的公司。对于流通市值最小的20%,的公司组合来说，股票组合收益随公司盈利能力的提高而下降，这说明对于这类公司投资者不关注业绩，反而更关注小市值公司业绩不佳导致资产重组的可能性，投机性更强。

**1.2 因子相关性分析**



图2 因子相关性系数图

观察相关性系数表发现以下特点：

特点一：SMB与HML相关性高：

由于二者有相同的计算因子，因此高相关性符合常理。

特点二：RMW与SMB、HML相关性高：

由于公司的运营利润率很大程度上取决于公司的市值和账面市值，因此高相关性符合常理。

特点三：CMA与RMW的相关性高：

运营利润的高低与该公司的运营风险相关，而投资风格的激进与保守使得投资者对不同运营风险的公司存在偏好，因此投资风格与运营利润的高相关性符合常理。

**1.3 多元回归显著性分析**

**1.3.1 R2**

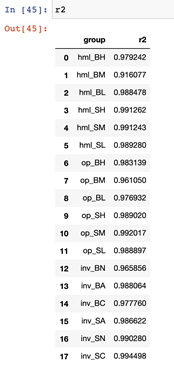
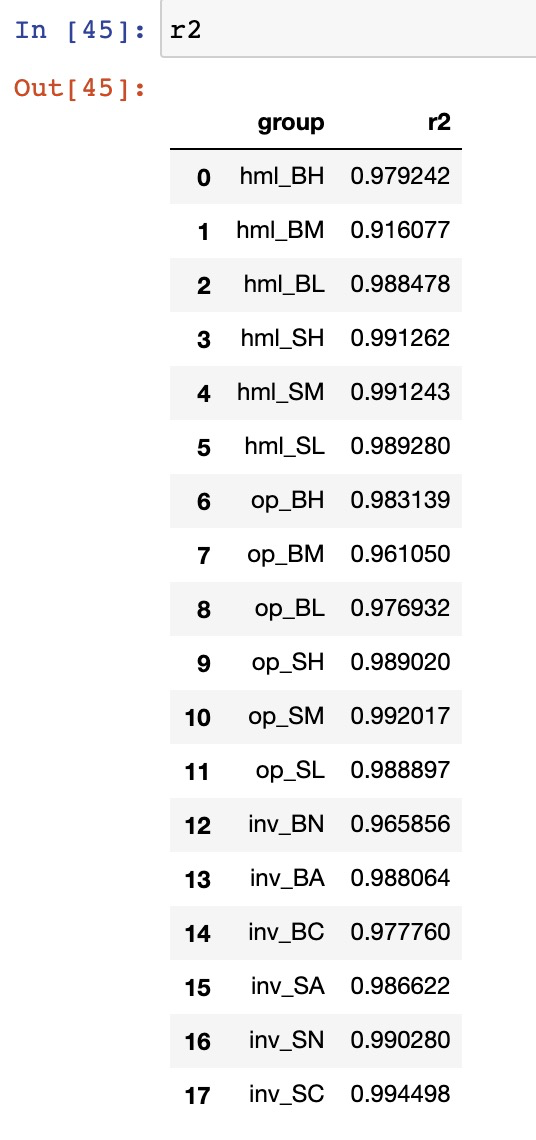


图3 R2图

上图展示了五因子模型在18个分组中多元线性回归的拟合优度。由图可知：所有组别回归拟合优度R2均大于0.9，绝大多数分组的回归拟合优度大于0.95，约一半左右的组拟合优度达到0.99。这充分说明了Fama-French五因子对单个分组内的投资组合收益率具有很好的解释度。

**1.3.2 p值**

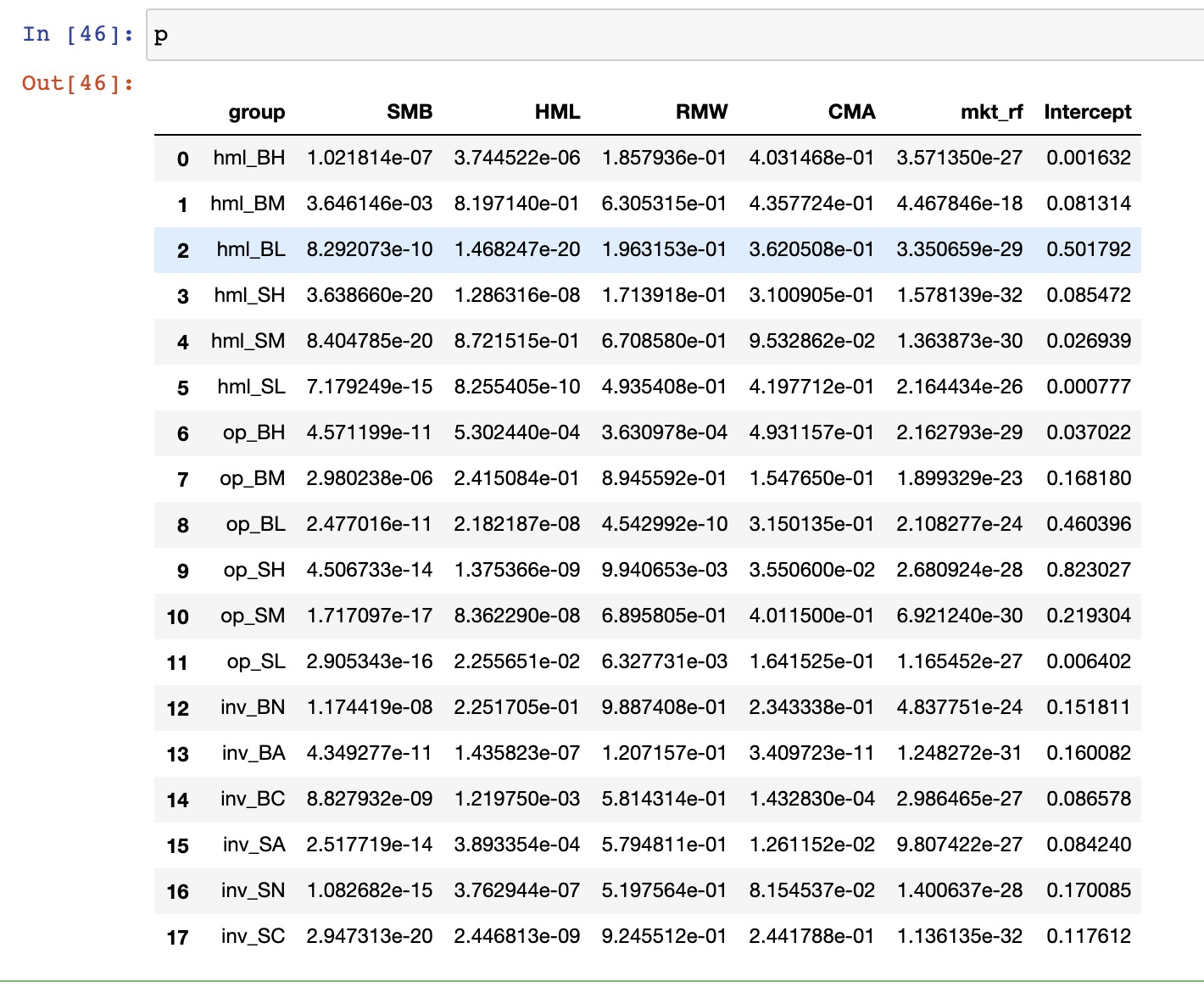


图4 p值图

上图展示了五因子模型在18个分组中多元线性回归的p值与t值，由图可知，除截距项以外的所有p值均远小于0.05，说明Fama-French五因子的多元线性回归在统计意义上具有显著性。

**1.4 冗余分析**

按照 Fama and French(2015)所采取的方式，进行冗余因子的鉴别。Fama and French(2015)在构造五因子模型前，逐个将每一个因子作为被解释变量，其他因子作为解释变量，进行因子之间的回归，以此来识别冗余因子，并发现在北美市场 HML 是冗余因子，在欧洲市场、日本市场和亚太市场，SMB 和 CMA 是冗余因子。冗余因子对于解释股票收益的作用微乎其微，其对于解释动量效应的帮助也不大，因此，在对动量效应进行回归之前，及时鉴别冗余因子，可以更好的优化模型，保证模型的解释力和简洁性。

当某一因子作为被解释变量，其他因子作为解释变量时，如果回归截距项接近 0，则说明该因子可以被其他因子所覆盖，即该因子为冗余因子。在全样本下对五因子模型进行冗余检验，发现 CMA 因子“冗余”，这与 Fama(2015)对亚太市场的研究结果一致。囿于篇幅本文省略了因子之间的回归统计结果。

我们猜测：CMA 为冗余因子，是由于我国企业自改革开放以来大都奉行大量投资的模式，不顾盈利的大量投资导致中国企业投资模式差异化很低，且中国股票市场处于转轨时期，上市公司相对于投资模式更关注财务报表的盈利性，因此，投资模式对股票收益的影响也就很小。

**（二）对收益率呈现效果的原因分析：**

总体来看我国股票收益率趋同现象比较明显，与美国股票市场相比，月均收益率高于美国股市，收益率波动的方差比美国股票市场要小得多。

主要原因可能有：

第一，中国严苛的IPO条件导致小市值公司的价值反映的实际上是买壳上市中的潜在壳资源价值。因此，对于市值较小的公司，他们的价值包括壳价值，因此为了除去这部分不准确的信息、剔除壳价值，需要刨除市值较小的公司，再进行重分类。这对应解释了适用性分析中的第三点。

第二，相对于公司内部人而言，对公众投资者保护不足，投资者处于信息劣势的地位，无法充分判断公司的价值，从而只能采取从众行为。我国股票市场投资者结构中个人投资者比例显著高于美国市场，而机构投资者也以基金为主，投资者结构比较单一，同质化严重。市场存在明显的“羊群效应”，股票市场往往表现为同涨同跌，系统风险是主要的风险，上市公司的盈利情况、财务状况以及个体风险往往被投资者忽视。我国股票市场由于发展时间较短，市场规范程度较低，不对称信息问题比较严重。信息披露机制不健全以及投资者非理性等因素也造成了信息不对称。此外，上市公司具有信息优势，而监管机构和投资者处于信息劣势，上市公司企业制度不完善，加剧了信息不对称程度和逆向选择风险。监管机制和投资者保护机制不完善导致上市公司信息披露不实，内幕知情者利用虚假信息恶意炒作等现象时有发生，因而投资者对上市公司的财务信息缺乏信任。

第三，我国对私人财产权利保护程度低，政治决策和传言会导致市场股票价格大幅波动，而且产权保护不足使得信息知情人的套利交易遭受较大的噪音交易者的阻碍，风险提高而收益降低，无法通过套利交易提高公司的定价效率。

第四，我国股票市场以换手率高著称，大部分股票投资者以短期投资获取股票价格波动收益为目标，而非“买入并持有”，投资者较少关注财务信息和公司经营决策。

1. **外部数据验证分析**

本部分中，我们使用了JointQuant API，分别验证了数据时间跨度、更新频率对Fama French 五因子模型效果的影响。

使用JointQuant API的原因在于：

第一、由于只提供了2014-2018年的数据，实际可用的数据较少，且由于数据的缺失，不能做到更小时间尺度上的回测，因此使用JointQuant API提供的数据可以弥补这一缺陷。

第二、我们的代码适用于更为理想的环境，没有考虑到手续费等现实因素的影响，而这些误差在JointQuant中可以通过调用API消除。

**（一）时间跨度**

分别选取2006年-2016年、2013年-2017年的数据，做回测后，得到如下结果：



图5 2006-2016年回测图



图6 2013-2017年回测图

研究发现，2006年-2016年的收益水平、夏普比率更高，同时最大回撤也更高，反应了买入后可能的最大亏损较大。

这一现象可以解释为，随着市场理性随时间推移的不断增加，Fama-French五因子模型对的效果逐渐下降，也就是说，2013年-2017年收益率比2006年-2016年收益率更低的原因在于，后者的时间更早，利用市场的非理性可以获得更多收益。

**（二）更新频率**

JointQuant API提供了以天、分钟为单位的回测更新频率。以2014年-2017年的数据为例，收益率效果对比如下：



图7 以“天”为周期进行回测图



图8 以“分钟”为周期进行回测回测图

结果显示以“天”为回测更新频率的收益率、夏普比更高，同时最大回撤也更低，反应了买入后可能的最大亏损较小。猜测原因在于“分钟”更新过快，容易受到市场无意义、非理性波动的影响，出现了类似“过拟合”的判断失误。

观察各因子的影响波动如下：

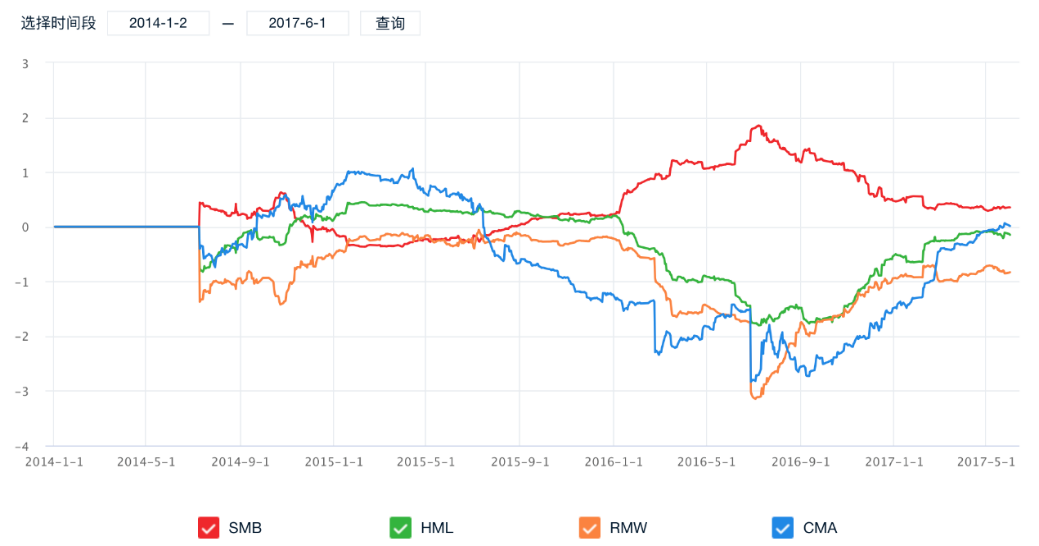


图9 以“天”为周期进行回测，因子波动图

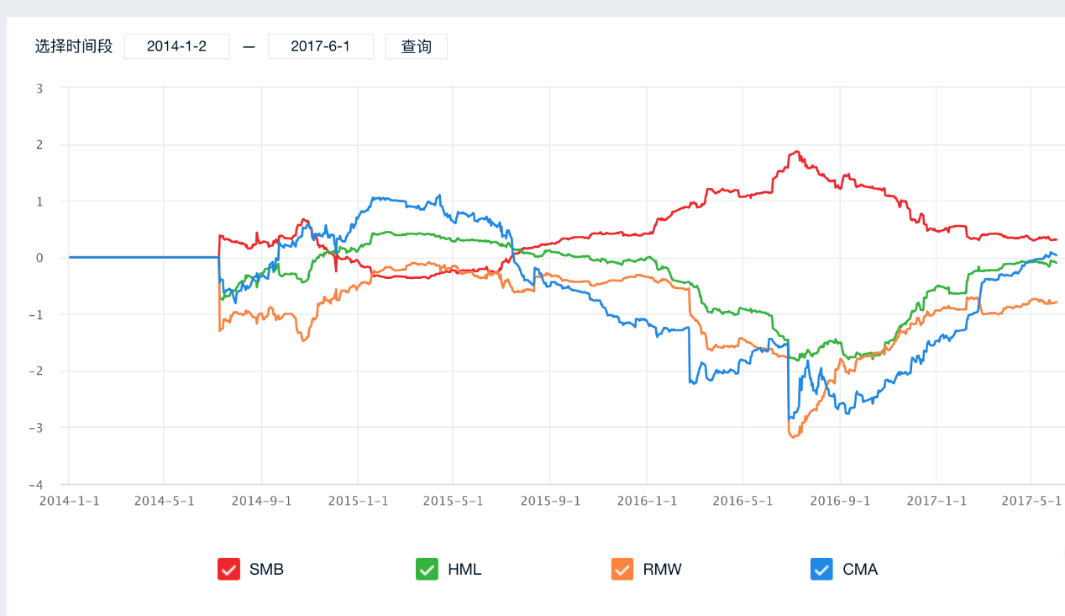


图10 以“分钟”为周期进行回测，因子波动图

结果显示，二者因子影响差别不大，因此过于频繁的更新对结果精确度的提升不大，结合其对收益率和回撤的影响，综合考虑下，适中的回测更新频率，如分钟，更为合适。

1. **结论**

本文选取中国A股上市公司2014年1月-2018年6月共54个月的月度数据,按照Fama and French(2015a)方法，对五因子模型在中国股市的适用性进行检验。本文研究表明：第一、对我国股票市场来说，市值效应和价值效应明显，盈利能力强的公司组合回报率要高于盈利能力差的公司组合的回报率，投资水平低的公司组合的回报率要高于投资水平高的公司组合的回报率，但差异不明显；第二、与国际市场检验结果一致，五因子模型具有非常强的解释能力，比 CAPM、三因子模型、Carhart四因子模型表现更好；第三、因子模型的解释能力因市场而异。Fama和French(2015)对美国股市数据的实证分析结果认为相比于 RMW 和CMA 来说 HML 是冗余变量，而我国恰巧相反CMA为冗余变量。

这也反映了两个市场投资者理念和策略的差异，成熟市场的投资者更看重公司的个体差异和公司未来的发展前景和成长空间，价值投资策略主要关注公司的发展前景和成长机会。对于我国股票市场而言，投资者更关注的是上市公司的估值水平，并不看重公司的发展前景和成长空间。这说明我国投资者对资本市场投资的信息传递机制和投资者的投资理念与美国资本市场还存在一定差异。

因子模型的解释和预测能力因市场而异，模型的构建还需考虑市场的实际情况。探索其他适合中国股市的定价因子和风险因子，以构造更精确、更适用于中国市场的资产定价模型仍是这一领域未来的研究方向，这对于理解中国市场，提供监管依据，以及辅助投资实践都有重要的意义

**参考文献：**

[1]Eugene F. Fama,Kenneth R. French. A five-factor asset pricing model[J].

Journal of Financial Economics,2015,116(1).

[2]宋光辉,董永琦,陈杨炀,许林.中国股票市场流动性与动量效应——基于Fama-French五因子模型的进一步研究[J].金融经济学研究,2017,32(01):36-50.

[3]赵胜民,闫红蕾,张凯.Fama-French五因子模型比三因子模型更胜一筹吗——来自中国A股市场的经验证据[J].南开经济研究,2016(02):41-59.

[4]李志冰,杨光艺,冯永昌,景亮.Fama-French五因子模型在中国股票市场的实证检验[J].金融研究,2017(06):191-206.

[5]王馨悦. 基于Fama-French模型的中国股票市场因子模型研究[D].哈尔滨工业大学,2020.

[6]傅以沫. 投资者情绪对股票收益的影响研究[D].华东师范大学,2020.