

# “纯洁 alpha” 动量下的行业轮动策略——资产定价 + 统计学习

分析师：唐军

执业证书编号：S0740517030003

Email: tangjun@r.qlzq.com.cn

分析师：包赞

执业证书编号：S0740520100001

Email: baozan@zts.com.cn

FF5-Latent-Alpha 因子下的分组回溯



“纯洁 alpha” 下的多空对冲组合



风险提示：模型根据历史数据建立，不排除未来一段时间内失效的风险。

## 报告摘要

### ◆ 研究背景

尽管有充分的学术文献记录表明行业回报率是可预测的，之前有关研究也显示，国内的行业指数收益是可以预测的，但研究显示基于预测的模型稳健性不好。考虑到行业动量效应在A股市场越来越明显，从动量的角度来构建行业轮动策略是最直观的方法。当然，单纯的从收益率角度来构建策略也可以，但从稳健性角度来讲，需要构建更科学、更体系的模型。Sarwar et al., 2017 在论文中论述了运用Fama-French 5因子（后简称为FF5）计算行业组合alpha，并使用该alpha动量在美股市场中实现行业轮动策略，获得了超额收益。我们这个模型就是在参考Golam Sarwar等人利用Fama五因子下alpha动量构建行业轮动策略的基础上，引入统计学上的潜在因素（Latent factor）模型，通过获得更“纯洁”的alpha，来构造更有效、更稳健的行业轮动策略。

### ◆ “纯洁 alpha”

传统FF-5 alpha由于因子间的共线性，是一个有偏估计，更重要的是，我们能够观测到或者挖掘到的因子是有限的、静态的，在衡量组合业绩时，应该考虑所有能够解释组合业绩的因子。但是，这个难度较大。我们参考Yang Song 和Qingyuan Zhao (2017) 提出的利用潜在因子模型来评价组合的业绩表现。和事先指定因子集合不同，这个方法利用组合收益率估计出因子。直觉是，如果因子能够解释组合的横截面收益，那么我们一定能够利用基础资产来复制出这个因子，这个因子带来的超额收益不应该被认为是超额收益，因为他不是单个资产特有的。关于这个方法的技术解决方案是一个成熟的统计算法——Confounder Adjusted Testing and Estimation (CATE)，来估计潜在因子模型。CATE当初被用来处理基因之类的复杂数据。计算分两步：第一步，利用极大似然估计来找出因子，并且估计出组合资产在这些因子上的载荷。第二步，利用横截面回归，加入潜在因子，估计出横截面alpha，此时alpha剔除了更多的共同因子，我们称之为“纯洁alpha”。

### ◆ 行业轮动策略业绩表现

回溯利用 20130101-20200731 中信一级行业指数收益率数据。评价文中不同参数下的策略表现，主要看多空对冲组合的夏普比，夏普比越高，代表策略表现越好、稳健性越好，分析显示采用分三组月度调仓的模式，表现最好。普通FF-5alpha和纯洁alpha策略业绩对比，纯洁alpha对冲组合夏普为1.40，普通alpha下的对冲组合夏普为1.26，纯洁alpha策略夏普比有一定的提升。纯洁alpha下的多头行业组合年化收益为20.17%，普通alpha下的年化收益为19.46%，看似年化收益提升不多，实则潜在因子下的纯洁alpha在行业选择上已经有了大幅提升，因为我们组合选取十个优秀行业，每个行业等权配置，纯洁alpha新选入的行业收益一定是有了比较大幅的提升，才带来组合相对传统alpha的0.84%的年化超额收益。

## 正文目录

1. 引言 .....	4
2. FF-5 下的 alpha 动量行业轮动策略 .....	4
2.1. Fama 五因子 .....	5
2.2. 五因子解释性展示 .....	6
2.3. 基于 FF-5 alpha 的回测分析 .....	8
3. “纯洁 alpha” 动量下的行业轮动 .....	13
3.1. alpha 估计有偏 .....	13
3.2. 考虑潜在因子子的 alpha 估计算法 .....	14
3.2.1. 基金收益率因子模型 .....	14
3.2.2. 估计潜在因子 .....	15
3.3. 潜在因子模型运用 .....	18
3.3.1. 模型解释 .....	18
3.3.2. 潜在因子模型的回归展示 .....	18
3.4. “纯洁” alpha 指标下的回测分析 .....	19
4. 总结 .....	24

## 图表目录

图 1: Fama 五因子 .....	5
图 2: FF5-Alpha 因子分组回测 (分五组, 季度调仓) .....	9
图 3: FF5-Alpha 因子下的多空回测 (分五组, 季度调仓) .....	9
图 4: FF5-Alpha 因子分组回测 (分五组, 月度调仓) .....	10
图 5: FF5-Alpha 因子下的多空回测 (分五组, 月度调仓) .....	10
图 6: FF5-Alpha 因子分组回测 (分三组, 季度调仓) .....	11
图 7: FF5-Alpha 因子下的多空回测 (分三组, 季度调仓) .....	11
图 8: FF5-Alpha 因子分组回测 (分三组, 月度调仓) .....	12
图 9: FF5-Alpha 因子下的多空回测 (分三组, 月度调仓) .....	12
图 10: FF5-Latent-Alpha 因子分组回测 (分五组, 季度调仓) .....	20
图 11: FF5-Latent-Alpha 因子下的多空回测 (分五组, 季度调仓) .....	20
图 12: FF5-Latent-Alpha 因子分组回测 (分五组, 月度调仓) .....	21
图 13: FF5-Latent-Alpha 因子下的多空回测 (分五组, 月度调仓) .....	21

图 14: FF5-Latent-Alpha 因子分组回测 (分三组, 季度调仓)	22
图 15: FF5-Latent-Alpha 因子下的多空回测 (分三组, 季度调仓)	22
图 16: FF5-Latent-Alpha 因子分组回测 (分三组, 月度调仓)	23
图 17: FF5-Latent-Alpha 因子下的多空回测 (分三组, 月度调仓)	23
表 1: 因子间相关性 (2010-2019)	6
表 2: 不同分位下的单个基金业绩表现 (括号类为 T 值, 下面是调整后 R 方)	6
表 3: 不同分位区间基金组合的 alpha 表现 (上面为等权, 下面为市值加权)	6
表 4: 中信一级行业 FF-5 回归 (上面一行为年化 alpha 与 T 值, 下面为 R 方)	7
表 5: FF5-Alpha 因子分组回测及多空表现 (分五组, 季度调仓)	9
表 6: FF5-Alpha 因子分组回测及多空表现 (分五组, 月度调仓)	10
表 7: FF5-Alpha 因子分组回测及多空表现 (分三组, 季度调仓)	11
表 8: FF5-Alpha 因子分组回测及多空表现 (分三组, 月度调仓)	12
表 9: 考虑潜在因子下的 alpha 与原来 alpha 对比 (20190701-20191231)	19
表 10: FF5-Latent-Alpha 因子分组回测及多空表现 (分五组, 季度调仓)	20
表 11: FF5-Latent-Alpha 因子分组回测及多空表现 (分五组, 月度调仓)	21
表 12: FF5-Latent-Alpha 因子分组回测及多空表现 (分三组, 季度调仓)	22
表 13: FF5-Latent-Alpha 因子分组回测及多空表现 (分三组, 月度调仓)	23
表 14: FF-5alpha 和“纯洁”alpha 回测效果对比 (下面行为“纯洁”alpha)	24
表 15: 模型历史结论	25

## Industry Rotation Strategy —— Using Innocent Alpha

### Abstract

It is common to evaluate portfolio returns by linear factor models, such as Fama five factor model. We test if such alpha signifies a true industry out/underperformance by applying simple long-only and long-short industry rotation strategies. Our long-only sector rotation strategy that buys an industry with high alpha generates three times higher Sharpe ratio than the low alpha industry portfolio. However, performance measures from these asset pricing models are misleading if there are some omitted factors that explain cross-sectional variation in returns. To improve the robustness of our strategy. We propose to use a latent-factor approach, Confounder Adjusted Testing and Estimation (CATE), for performance evaluation. After considering latent factors, alpha becomes purer and we call it innocent alpha. We demonstrate that CATE outperforms widely used factor models in identifying common variation in industry index returns and that CATE alpha positively predicts future index performance.

## 1. 引言

尽管有充分的学术文献记录表明行业回报率是可预测的，之前有关研究也显示，国内的行业指数收益是可以预测的，但之前研究显示基于预测的模型稳健性不好。考虑到行业动量效应在A股市场越来越明显，消费、医药等行业在最近三年展现出较强的动量效应，业绩表现显著好于其它行业，所以从动量的角度来构建行业轮动策略是最直观的方法。当然，单纯的看收益率涨跌来构建策略也可以，但从稳健性角度来讲，需要构建更科学、更体系的模型。Sarwar et al., 2017在论文中论述了运用Fama-French 5因子模型（后简称为FF5）计算行业组合中的Alpha因子，并使用该Alpha动量在美股市场中实现行业轮动策略，获得了超额收益。我们这个模型就是在参考Golam Sarwar等人利用Fama五因子下alpha动量构建行业轮动策略的基础上，引入统计学上的潜在因素（Latent factor）模型，通过获得更“纯洁”的alpha，来构造更有效、更稳健的行业轮动策略。

金融中普遍利用CAPM alpha或者FF-5控制下的alpha来评价组合业绩，但是这个标准下的alpha似乎并不科学，因为他们只考虑的我们熟悉的因子。事实上，金融学中有很多因子，能够产生超额alpha，我们称之为金融异像，譬如低波异象等。正如Pastor and Stambaugh (2002)所说“alpha常被认为基金经理选择错误定价股票的能力，但是非零alpha并不代表投资能力，因为很多被动投资也能有非零alpha。” Barber, Huang and Odean (2016)的观点是“理论上讲，当评价投资经理投资能力时，投资者应该考虑所有能够解释组合横截面业绩的因子，无论这些因子是合理定价或者错误定价”。

Fama三因子模型或者Carhart四因子模型常被当做标准的业绩基准来衡量投资组合的表现。2013年Fama在三因子的基础上添加了盈利（RMW）和投资（CMA）因子，解释力度有所提升。尽管我们实证分析显示q-factor模型的解释性在国内稍微好于FF-5，出于认可度的原因，我们依然还是采用最经典的FF-5体系，来作为投资组合业绩分析的基准。当然，正如上文所说，最好的业绩基准就是能包含解释组合收益的所有因子，尽管这个难度会很大。首先，很难事先决定相关的因子集，金融文献中，关于解释组合业绩横截面收益的因子多达数百个。另外一个挑战是投资组合不是静态的，即便投资组合是静态的，驱动组合业绩的因子也是在改变，所以，解释组合业绩的因子也是动态变化的。

由于本文借鉴Sarwar的alpha动量思路，为了得到更“纯洁”alpha，需要考虑到尽可能多的因子，所以在Sarwar等人基础上，作者借鉴Wang, J., Q. Zhao, T. Hastie, and A. B. Owen (2017)在《Confounder adjustment in multiple testing》一文介绍的潜在因子构造方法，在Fama-French 5因子模型中加入Latent因子，并提取新的更“纯洁”Alpha因子，后文会展示，“纯洁alpha”动量下的行业轮动策略获得了更好的风险收益表现。

本文的模型有以下应用：第一是可以为主观投资提供行业配置观点；第二是可以帮助指数增强产品产生收益增强的效果。相比于行业收益的预测模型，本文模型无需动态调整参数，稳健性良好。后文会首先介绍FF-5 alpha下的行业轮动策略，并且给出策略回测分析。然后介绍潜在因子模型的具体算法，展示其能够提取出更“纯洁”的alpha，最后会描述“纯洁”alpha动量下的行业轮动模型收益表现。

## 2. FF-5 下的 alpha 动量行业轮动策略

本文利用中国版的 FF5 模型研究行业组合的风险调整后收益。基金研究的相关人员也可能对这一部分的结果感兴趣。建模目标是利用 alpha 动量来构建行业轮动策略。由于 ETF 产品的不断丰富，也使得本文提出的行业轮动策略更有实战意义。即便考虑到交易成本，本策略也是可行的。



首先来看国外现有的有关行业组合业绩分析的数据。大多数组合业绩分析文献都在关注共同基金领域，这一领域的研究数量庞大，例如见 Carhart (1997 年)、Daniel (1997 年)、Wermer (2000 年)、Chen 等人。相比之下，文献中对行业组合业绩或者说行业组合 alpha 的关注要少得多。Dellva, DeMaskey 和 Smith(2001)研究了 35 只 Fidelity 行业型基金从成立到 1998 年 12 月的发展情况。Faff (2004) 对澳大利亚 24 个产业组合的业绩进行了测试，发现矿产资源行业倾向于产生负的 FF3 alpha，而工业则倾向于产生正的 FF3 alpha。在美国，Kacperczyk、Sialm 和 Zheng (2005) 调查了 1984 年 1 月至 1999 年 12 月期间，重仓工业行业基金的业绩。他们认为，基金管理人可能会因其投资组合具有特定行业集中度，而背离被动的市场组合，并证明了，**专注于特定行业的基金会有更好的表现**。Dou 等人 (2014 年)研究发达国家(北美、英国、日本和欧洲)各行业在不同经济体制下的资产分配情况，他们发现，能源、高科技、医疗卫生行业的 alpha 为正值；不论牛市还是熊市，电信和制造业行业的阿尔法为负值。

Sorensen 和 Burke(1986 年)以及 Grauer、Hakansson 和 Shen(1990 年)等人做行业轮动策略较早，他们的行业轮动策略基于均值方差框架，在 Beller, Kling 和 Levinson(1998)的研究中已被证明是成功的。上世纪 80 年代，富达投资(Fidelity Investments)推出了一系列所谓“精选”系列的行业主题基金，从而将行业投资推入主流。1998 年 12 月，第一个 ETF 被成功推选给了投资者，交易基金的时代来临。以“富达行业精选”(Fidelity Select Sector picking)作为标准，Sassetti 和 Tani(2006)使用了三种简单的行业轮动策略，根据涨跌幅、Alpha 和相对强度指标对行业进行排名。他们在 1998 年 1 月至 2003 年 9 月期间，将策略应用于富达精选的 41 个基金。**他们的研究结果显示，根据过往 alpha 值来做行业轮动策略，似乎比单单依靠涨跌幅的策略更加稳健**。他们的行业轮动策略持续优于买入持有战略。Conover 等 (2008) 根据宏观经济条件对行业进行调整，他们发现，在 33 年的研究期间，他们的行业轮动策略很少能有持续的显著回报，这也显示利用宏观数据来构建行业轮动策略，似乎效果并不好。

## 2.1. Fama 五因子

Fama-French(1993)三因素模型(以下简称 FF3)和 Carhart(1997)四因素模型被学术界和投资者用作标准定价模型和基准模型。Fama 和 French(2015)五因素模型(以下简称 FF5)是在研究中经常使用的多因素模型(资产定价和绩效评价)中最新的一种。FF3 模型虽然被频繁使用，但它解释力度还是不够好，尤其是涉及到盈利能力和投资相关的变化(见 Chen, 2011; Aharoni, 2013; Titman 等, 2004)。为此,Fama 和 French(2013) 提出了五因素模型，在 FF3 模型基础上增加了两个额外因素，即盈利能力因子和投资因子。他们发现，对于规模、账面市值比 (B/M)、盈利能力、投资这些因子形成的组合，五因素模型比原 FF3 模型提供了更好的解释能力。

图 1: Fama五因子



资料来源：中泰证券研究所

表 1：因子间相关性（2010-2019）

Factor Portfolio	Std Dev	t-test for Mean=0	Cross-Correlations					
			MKT	SMB	HML	RMW	CMA	MOM
MKT	22.79%	0.13	1.00					
SMB	6.26%	6.24	-0.20	1.00				
HML	7.08%	1.52	0.31	-0.44	1.00			
RMW	4.21%	2.61	0.17	-0.72	0.35	1.00		
CMA	2.92%	2.54	0.16	-0.24	-0.23	0.39	1.00	
MOM	10.80%	-2.86	-0.09	0.01	-0.47	0.00	0.21	1.00

资料来源：中泰证券研究所

表 1 可以看出，因子间相关系数除了 SMB 和 RMW 较高外，其余均较小，表示多重共线性对因子载荷的影响不大。SMB 和 RMW 的相关系数虽较高，考虑 Fama 论文中也出现过大约相同的相关系数，并未特殊处理，本文在第一部分也不做处理。后文会展示出，多重共线性会带来截距项估计的有偏，而且本文研究 alpha 数值大小，所以，引入潜在因子模型来进行更优的估计。

## 2.2. 五因子解释性展示

为了展示五因子的解释能力，我们利用股票基金收益率序列来展示，因为股票基金组合涵盖各种可能的情况，与特定因子下的组合有规律不同，股票基金组合随机性更大。

表 2：不同分位下的单个基金业绩表现（括号类为 T 值，下面是调整后 R 方）

	Return	1-Factor	3-Factor	4-Factor	5-Factor
<b>10<sup>th</sup> pcnt</b>	198.67%	12.23% (1.79) 62.12%	6.47% (1.20) 77.25%	6.53% (1.22) 77.58%	-2.34% (-0.44) 78.49%
<b>25<sup>th</sup> pcnt</b>	152.96%	9.92% (1.28) 53.66%	11.21% (1.96) 75.54%	11.30% (1.99) 76.06%	-0.18% (-0.03) 77.41%
<b>Median</b>	110.43%	7.75% (0.77) 44.22%	4.64% (0.66) 73.38%	4.70% (0.66) 73.50%	-2.23% (-0.31) 73.86%
<b>75<sup>th</sup> pcnt</b>	75.23%	3.64% (0.48) 53.60%	3.20% (0.52) 70.26%	3.23% (0.52) 70.30%	-4.25% (-0.68) 70.91%
<b>90<sup>th</sup> pcnt</b>	42.02%	-0.59% (-0.08) 63.41%	-14.77% (-2.86) 83.20%	-14.73% (-2.87) 83.33%	-15.51% (-2.94) 83.18%

资料来源：中泰证券研究所

表 2 是用不同分位数下的单个基金和不同因子回归得出的 alpha，时间段是 2010-2017。表 3 是对不同分位下的基金的组合进行分析，从表 3 可以看出，随着因子的增多，R 方呈现逐渐增大趋势。但是，四因子和五因子给 R 方提高的程度并不明显，表示分析区间内的这八年，国内基金在动量因子、盈利因子、投资因子上的暴露并不多。主要还是暴露在 Fama 经典的三因子，即市场因子，市值因子，账面市值比因子上。从 R 方大小来看，五因子下的 R 方在 90% 左右，显示出五因子在国内良好的解释性。

表 3：不同分位区间基金组合的 alpha 表现（上面为等权，下面为市值加权）

	return	1-Factor	3-Factor	4-Factor	5-Factor
<b>0-20<sup>th</sup> pcnt</b>	215.95%	12.56% (2.69) 76.76%	8.79% (2.74) 89.41%	8.83% (2.78) 89.60%	1.68% (0.54) 90.35%
	211.70%	12.31% (2.61) 76.88%	8.12% (2.56) 89.80%	8.16% (2.60) 89.98%	1.50% (0.48) 90.61%
	150.41%	8.56% (1.76) 76.03%	4.21% (1.28) 89.53%	4.25% (1.30) 89.68%	-2.82% (-0.88) 90.35%
<b>20-40<sup>th</sup> pcnt</b>	148.98%	8.42% (1.74) 76.65%	3.57% (1.08) 89.56%	3.61% (1.10) 89.70%	-3.48% (-1.08) 90.36%

40-60 <sup>th</sup> pcnt	116.12%	6.08% (1.20) 74.79%	2.73% (0.83) 89.65%	2.78% (0.85) 89.89%	-3.78% (-1.16) 90.34%
	117.16%	6.19% (1.20) 73.96%	2.89% (0.85) 89.23%	2.93% (0.88) 89.41%	-3.94% (-1.18) 89.99%
60-80 <sup>th</sup> pcnt	86.22%	3.50% (0.69) 74.99%	-0.82% (-0.25) 89.47%	-0.78% (-0.23) 89.67%	-6.94% (-2.08) 90.08%
	85.93%	3.45% (0.69) 75.58%	-0.48% (-0.15) 89.87%	-0.43% (-0.13) 90.11%	-6.16% (-1.89) 90.38%
80-100 <sup>th</sup> pcnt	40.67%	-1.31% (-0.25) 74.65%	-5.10% (-1.51) 89.55%	-5.06% (-1.51) 89.72%	-10.64% (-3.15) 90.02%
	36.86%	-1.69% (-0.30) 72.20%	-5.30% (-1.43) 88.07%	-5.26% (-1.43) 88.17%	-11.67% (-3.17) 88.66%

资料来源：中泰证券研究所

用我们此次建模用到的**中信一级行业指数组合**来对**FF-5**进行回归分析。见表4，表中我们能看出，同一个行业在不同的年份，五因子对其解释力度差异很大，石油化工行业在2016年能被五因子很好的解释，R方高达89%，但是在2017年，R方只有52%。不同的行业在相同的年份，R方差异也很大，在2017年，银行、地产、有色、钢铁等行业的R方较小，不到60%，但是机械、纺织服装这两个行业的R方在2017年较高。总的来看，五因子对行业的解释力度还是较大的，平均有80%。但是，不同的时间区间内，解释力度差异很大，所以，这也是我们下文用到潜在因子模型的原因，我们需要经典定价因子之外的共同因子，这样才能获得更纯洁的alpha。

表 4：中信一级行业FF-5回归（上面一行为年化alpha与T值，下面为R方）

	2015	2016	2017	2018	2019	2015-2019
石油石化	-27.80% (-1.73)	10.10% (1.03)	-0.89% (-0.10)	2.10% (0.13)	-9.94% (-1.06)	-4.92% (-0.88)
	88.79%	89.44%	51.85%	63.99%	73.96%	81.05%
煤炭	-20.23% (-0.96)	1.01% (0.05)	-10.11% (-0.56)	-17.02% (-0.94)	-4.97% (-0.45)	-1.87% (-0.22)
	87.15%	71.44%	43.15%	65.07%	77.54%	73.17%
有色金属	-38.26% (-1.53)	-7.09% (-0.33)	-5.86% (-0.32)	-12.26% (-0.84)	18.27% (1.11)	-2.84% (-0.32)
	83.15%	71.53%	42.40%	71.83%	62.07%	71.34%
电力及公用事业	-42.77% (-2.07)	-13.97% (-1.94)	-1.28% (-0.17)	-14.73% (-1.63)	-6.59% (-0.88)	-16.24% (-2.77)
	87.81%	94.52%	65.00%	80.24%	82.77%	82.84%
钢铁	-34.79% (-1.51)	-1.42% (-0.09)	-0.02% (0.00)	-18.78% (-1.14)	-11.67% (-0.97)	-5.67% (-0.68)
	84.95%	82.81%	44.50%	64.94%	74.15%	73.51%
基础化工	-35.85% (-1.73)	-6.76% (-0.72)	-5.96% (-0.69)	-18.95% (-1.74)	-10.21% (-0.93)	-10.30% (-1.66)
	87.57%	94.04%	73.22%	82.62%	81.18%	84.36%
建筑	1.75% (0.07)	10.44% (0.82)	-7.84% (-0.71)	-16.99% (-1.40)	-20.68% (-2.43)	-6.69% (-0.98)
	82.14%	88.34%	56.96%	75.21%	85.34%	80.33%
建材	-42.17% (-2.19)	-7.77% (-0.81)	9.01% (0.72)	-20.66% (-1.46)	11.75% (1.00)	-4.11% (-0.62)
	89.23%	93.89%	60.58%	79.02%	78.05%	82.84%
轻工制造	-34.29% (-1.55)	-5.67% (-0.59)	-8.68% (-1.02)	-12.89% (-1.21)	-1.94% (-0.20)	-7.02% (-1.10)
	85.89%	93.40%	74.91%	84.49%	83.97%	83.42%
机械	-27.94% (-1.28)	-6.34% (-0.85)	3.63% (0.51)	-10.68% (-1.15)	0.46% (0.06)	-6.23% (-1.08)
	87.18%	96.25%	82.27%	87.68%	91.19%	86.90%
电力设备及新能源	-36.31% (-1.53)	-6.90% (-0.80)	-0.52% (-0.06)	-11.79% (-1.08)	-5.84% (-0.48)	-6.97% (-1.02)
	86.03%	95.09%	71.19%	80.79%	78.41%	82.49%
国防军工	-33.11% (-0.80)	15.86% (0.86)	22.76% (1.37)	24.17% (1.31)	13.21% (0.77)	4.14% (0.37)
	68.46%	81.04%	49.60%	66.58%	63.72%	65.76%
汽车	-23.72% (-1.30)	-2.04% (-0.28)	2.34% (0.30)	-7.70% (-0.84)	-6.87% (-0.65)	-5.15% (-0.97)
	0.89	0.95	0.70	0.85	0.81	0.86

商贸零售	-48.18% (-1.94)	-16.29% (-1.58)	-2.75% (-0.34)	-12.81% (-1.15)	-11.65% (-1.40)	-15.57% (-2.31)
	84.56%	92.25%	70.83%	80.40%	85.68%	81.91%
消费者服务	-2.09% (-0.08)	-32.92% (-2.69)	0.39% (0.03)	37.82% (1.92)	-5.08% (-0.35)	1.74% (0.21)
	80.50%	87.77%	50.77%	64.25%	60.62%	72.42%
家电	0.60% (0.03)	8.57% (0.67)	15.90% (1.06)	-4.17% (-0.23)	10.07% (0.70)	10.93% (1.46)
	81.74%	86.29%	47.81%	66.57%	71.06%	75.38%
纺织服装	-30.27% (-1.35)	-5.58% (-0.59)	-14.41% (-2.47)	-15.48% (-1.42)	-11.66% (-1.54)	-13.54% (-2.11)
	86.68%	93.38%	82.05%	78.27%	86.41%	82.44%
医药	-33.72% (-1.57)	-6.09% (-0.74)	2.39% (0.33)	3.28% (0.20)	-5.95% (-0.50)	-6.35% (-0.98)
	84.91%	93.65%	70.73%	71.37%	75.43%	80.80%
食品饮料	-7.50% (-0.46)	10.07% (0.71)	27.02% (1.80)	20.96% (1.08)	7.17% (0.48)	12.99% (1.82)
	86.47%	77.01%	42.63%	67.58%	68.14%	73.53%
农林牧渔	-41.77% (-1.60)	-17.02% (-1.38)	-7.13% (-0.71)	4.28% (0.28)	23.80% (0.85)	-2.33% (-0.25)
	81.79%	88.47%	61.36%	56.54%	29.65%	66.61%
银行	54.67% (2.22)	16.86% (1.59)	3.80% (0.36)	12.14% (0.97)	6.54% (0.70)	13.38% (2.01)
	67.54%	66.06%	42.96%	70.57%	75.44%	64.25%
非银行金融	-5.91% (-0.24)	27.64% (2.61)	6.90% (0.66)	16.17% (1.46)	22.19% (2.06)	11.11% (1.70)
	80.77%	89.69%	62.80%	86.10%	87.12%	82.16%
房地产	-2.27% (-0.11)	-15.82% (-1.09)	-1.69% (-0.14)	-9.70% (-0.66)	2.92% (0.25)	-3.69% (-0.55)
	87.19%	84.65%	45.70%	75.44%	77.07%	80.21%
交通运输	-22.07% (-1.00)	-17.68% (-1.86)	5.05% (0.62)	-6.28% (-0.73)	-8.28% (-1.03)	-9.69% (-1.65)
	85.96%	91.61%	65.86%	85.21%	85.63%	83.89%
电子	-35.31% (-1.44)	1.62% (0.15)	15.29% (1.37)	-5.14% (-0.37)	25.73% (1.85)	9.29% (1.26)
	83.89%	93.05%	69.60%	82.31%	81.63%	81.73%
通信	-14.85% (-0.63)	6.96% (0.71)	16.55% (1.47)	-4.18% (-0.26)	7.62% (0.51)	9.16% (1.23)
	86.00%	94.19%	67.25%	77.40%	79.81%	82.13%
计算机	-24.82% (-0.77)	12.48% (1.07)	9.48% (0.78)	39.25% (2.16)	27.53% (1.88)	15.97% (1.71)
	81.48%	94.37%	71.74%	76.83%	82.06%	79.16%
传媒	-21.98% (-0.75)	-3.40% (-0.31)	-6.07% (-0.75)	-9.90% (-0.76)	10.29% (0.86)	-3.63% (-0.46)
	79.85%	93.38%	73.92%	79.02%	83.46%	79.03%
综合	-36.75% (-1.52)	-9.55% (-0.72)	-4.69% (-0.37)	-26.25% (-2.01)	7.10% (0.72)	-6.60% (-0.88)
	86.53%	90.12%	60.83%	80.08%	87.07%	81.78%

资料来源：中泰证券研究所

### 2.3. 基于 FF-5 alpha 的回测分析

本次研究使用了 2013-2020 年期间，中信一级行业的 30 个行业指数日频率收益率序列数据，研究显示，FF-5 比 FF-3 更适合描述行业收益，具有较高的调整后的 R 方。FF-5 也是统计上更好的拟合模型，额外的两个因子(RMW 和 CMA) 显著增加了模型的解释性。此行业轮动策略在思想上与 Sassett 和 Tani (2006 年) 最为相似，遵循了他们的一项主要研究结果：即**基于 alpha 的行业轮动策略比基于其他指标的行业轮动策略表现更好**。而 Sassetti 和 Tani (2006) alpha 估计是基于 30、60 和 90 天的窗宽的，为使行业组合稳健性更好，我们利用过去 125 天的历史数据。我们根究每个行业回归出来的 alpha 大小，选择 alpha 大的一组行业作为未来的多头组合。

每次回归计算alpha用到的是过去125天的收益率数据，关于未来持有期限，我们下面分别展示了月度调仓和季度调仓的结果。为了获得较好的参数，我们对分三组和分五组，都进行了试验。下面分别展示了分五组按季度调仓、分五组按月度调仓、分三组按季度调仓、分三组按月度调仓的回测结果。分组数字越大的是alpha越大的组。



(1) 分五组按季度调仓

图 2: FF5-Alpha 因子分组回测 (分五组, 季度调仓)



资料来源: 中泰证券研究所

图 3: FF5-Alpha 因子下的多空回测 (分五组, 季度调仓)



资料来源: 中泰证券研究所

表 5: FF5-Alpha 因子分组回测及多空表现 (分五组, 季度调仓)

	万得全 A	Group5 - 1	Group5	Group4	Group3	Group2	Group1
累计收益率	134.77%	95.84%	255.99%	149.11%	114.16%	118.55%	77.34%
年化收益率	12.38%	9.63%	18.97%	13.30%	10.98%	11.29%	8.15%
年化夏普比	0.46	0.78	0.65	0.48	0.38	0.40	0.29
年化波动率	26.65%	12.41%	29.30%	27.89%	28.69%	28.46%	27.79%
最大回撤	55.99%	28.63%	57.10%	61.04%	61.43%	58.29%	59.82%
月度胜率		46.15%	59.34%	56.04%	52.75%	46.15%	41.76%

资料来源: 中泰证券研究所

(2) 分五组按月度调仓

图 4: FF5-Alpha 因子分组回测（分五组，月度调仓）



资料来源：中泰证券研究所

图 5: FF5-Alpha 因子下的多空回测（分五组，月度调仓）



资料来源：中泰证券研究所

表 6: FF5-Alpha 因子分组回测及多空表现（分五组，月度调仓）

	万得全 A	Group5 - 1	Group5	Group4	Group3	Group2	Group1
累计收益率	134.77%	168.62%	275.73%	241.83%	74.22%	129.61%	34.86%
年化收益率	12.38%	14.47%	19.85%	18.31%	7.89%	12.04%	4.18%
年化夏普比	0.46	1.12	0.68	0.65	0.28	0.42	0.15
年化波动率	26.65%	12.89%	29.03%	28.13%	28.40%	28.52%	27.85%
最大回撤	55.99%	22.23%	53.42%	54.50%	67.11%	56.80%	63.46%
月度胜率		47.25%	63.74%	61.54%	47.25%	48.35%	37.36%

资料来源：中泰证券研究所

(3) 分三组按季度调仓

图 6: FF5-Alpha 因子分组回测（分三组，季度调仓）



资料来源：中泰证券研究所

图 7: FF5-Alpha 因子下的多空回测（分三组，季度调仓）



资料来源：中泰证券研究所

表 7: FF5-Alpha 因子分组回测及多空表现（分三组，季度调仓）

	万得全 A	Group3 - 1	Group3	Group2	Group1
累计收益率	134.77%	65.00%	209.51%	144.26%	81.62%
年化收益率	12.38%	7.09%	16.72%	13.00%	8.51%
年化夏普比	0.46	0.76	0.60	0.46	0.30
年化波动率	26.65%	9.30%	27.94%	28.33%	27.97%
最大回撤	55.99%	22.40%	56.72%	62.23%	59.01%
月度胜率		45.05%	56.04%	43.96%	46.15%

资料来源：中泰证券研究所

(4) 分三组按月度调仓

图 8: FF5-Alpha 因子分组回测 (分三组, 月度调仓)



资料来源: 中泰证券研究所

图 9: FF5-Alpha 因子下的多空回测 (分三组, 月度调仓)



资料来源: 中泰证券研究所

表 8: FF5-Alpha 因子分组回测及多空表现 (分三组, 月度调仓)

	万得全 A	Group3 - 1	Group3	Group2	Group1
累计收益率	134.77%	131.04%	266.80%	138.90%	54.05%
年化收益率	12.38%	12.14%	19.46%	12.65%	6.09%
年化夏普比	0.46	1.26	0.69	0.45	0.21
年化波动率	26.65%	9.59%	28.21%	27.93%	28.04%
最大回撤	55.99%	15.84%	53.50%	61.13%	61.79%
月度胜率		45.05%	65.93%	49.45%	39.56%



### 3. “纯洁 alpha” 动量下的行业轮动

尽管金融中普遍利用CAPM alpha或者FF-5控制下的alpha来评价组合业绩,但是这个标准下的alpha似乎并不科学,因为他们只考虑的我们熟悉的因子。事实上,金融学中有很多因子,能够产生超额alpha,我们称之为金融异像,譬如低波异象等。正如Pastor and Stambaugh (2002)所说“alpha常被认为基金经理选择错误定价股票的能力,但是非零alpha并不代表投资能力,因为很多被动投资也能有非零alpha。” Barber, Huang and Odean (2016)的观点是“理论上讲,当评价投资经理投资能力时,投资者应该考虑所有能够解释组合横截面业绩的因子,无论这些因子是合理定价或者错误定价”。

所以,在这样的逻辑下,我们在衡量组合业绩时,应该考虑所有能够解释组合业绩的因子。但是,这个难度较大。首先,很难事先决定相关的因子集。金融文献中,关于解释组合业绩横截面收益的因子多达数百个。另外一个挑战是投资组合不是静态的,尽管投资组合是静态的,驱动组合业绩的因子也是在改变。所以,解释组合业绩的因子也是动态变化的。

我们根据Yang Song 和Qingyuan Zhao (2017)提出的利用潜在因子模型来评价组合的业绩表现。和事先指定因子集合不同,这个方法利用组合收益率估计出因子。他的直觉是,如果因子能够解释组合的横截面收益,那么我们一定能够利用基础资产来复制出这个因子。这个因子带来的超额收益不应该被认为是超额收益,因为他不是单个资产特有的。关于这个方法的技术解决方案是一个成熟的统计算法——干扰因子检验与估计Confounder Adjusted Testing and Estimation (CATE),来估计潜在因子模型。CATE当初被用来处理基因之类的复杂数据。计算分两步:第一步,利用极大似然估计来找出因子,并且估计出组合资产在这些因子上的载荷。第二步,利用横截面回归,加入潜在因子,估计出横截面alpha,此时alpha剔除了更多的共同因子,我们称之为“纯洁” alpha。

#### 3.1. alpha 估计有偏

潜在变量,指与自变量 X 和因变量 Y 都相关的变量,该变量能使 X 和 Y 之间产生虚假的相关关系。假设有三个变量 X、Y、Z,在研究 X、Y 之间的相关关系时,如果 Z 能够同时影响 X、Y,那么对于 X、Y 之间的回归分析很可能会出现虚假的相关关系。例如,研究 X 表示药物效果、Y 表示治疗效果、Z 表示性别年龄。显然 Z 对 X 和对 Y 都有影响,如果不控制 Z,那么在对 X、Y 回归分析中总就有可能得到虚假相关关系。为什么要消除 latent factor 的影响? 在应用 FF 因子模型计算组合 alpha 时,如果不消除 latent factor 的影响,在某些情况下,我们无法得到真正剥离系统因子的 alpha,也就是说 alpha 不“纯”,因为我们需要 alpha 动量,所以,“纯洁”的 alpha 模型效果会更好。

假设组合收益满足如下回归方程:

$$\mathbf{Y}_{n \times p} = \mathbf{1}_{n \times 1} \boldsymbol{\alpha}_{p \times 1}^T + \mathbf{Z}_{n \times d} \boldsymbol{\beta}_{p \times d}^T + \boldsymbol{\epsilon}_{n \times p}$$

$\mathbf{Y}_{n \times p}$  表示 n 个组合在 p 个交易日里的收益率矩阵;  $\mathbf{Z}_{n \times d}$  表示驱动该组合的因子数据。

不失一般性,我们可以写成:

$$\mathbf{Y}_{n \times p} = \mathbf{X}_{n \times 1} \boldsymbol{\alpha}_{p \times 1}^T + \mathbf{Z}_{n \times d} \boldsymbol{\beta}_{p \times d}^T + \boldsymbol{\epsilon}$$

在传统方法中，我们可以采用最小二乘法估计  $\boldsymbol{\alpha}_{p \times 1}^T$ 。在  $\mathbf{X}_{n \times 1}$  与  $\mathbf{Z}_{n \times d}$  正交的情况下，这种估计是无偏的。如果  $\mathbf{X}_{n \times 1}$  与  $\mathbf{Z}_{n \times d}$  不正交，即  $\mathbf{X}_{n \times 1}$  在  $\mathbf{Z}_{n \times d}$  上的投影不为 0，则该估计是有偏的，换言之估计出的  $\boldsymbol{\alpha}_{p \times 1}^T$  没有完全排除  $\mathbf{Z}_{n \times d}$  的影响。

若  $\mathbf{X}_{n \times 1}$  与  $\mathbf{Z}_{n \times d}$  不正交，设

$$\mathbf{Z}_{n \times d} = \mathbf{X}_{n \times 1} \boldsymbol{\gamma}_{d \times 1}^T + \mathbf{W}, \text{ 其中 } \mathbf{W} \perp \mathbf{X}$$

$$\text{故 } \mathbf{Y}_{n \times p} = \mathbf{X}_{n \times 1} \boldsymbol{\tau}_{p \times 1}^T + \mathbf{W}_{n \times d} \boldsymbol{\beta}_{p \times d}^T + \boldsymbol{\epsilon},$$

其中  $\boldsymbol{\tau} = \boldsymbol{\alpha} + \boldsymbol{\beta} \boldsymbol{\gamma}$ ， $\tilde{\boldsymbol{\tau}}$  是上式的最小二乘法无偏估计。

由此得到， $\boldsymbol{\alpha} = \boldsymbol{\tau} - \boldsymbol{\beta} \boldsymbol{\gamma}$ ，在  $\boldsymbol{\gamma}$  不为 0 时有偏。

## 3.2. 考虑潜在因素子的 alpha 估计算法

### 3.2.1. 基金收益率因子模型

我们首先定义基金收益率的线性因子模型。令  $R_{it}$  为基金  $i$  在  $t$  时的回报率， $R_{0t}$  为  $t$  时的无风险利率。那么基金的超额收益可以用下式表示：

$$R_{it} - R_{0t} = \alpha_i + \boldsymbol{\tau}_i^T \mathbf{V}_t + \epsilon_{it}, \quad i = 1, 2, \dots, N \quad (1)$$

其中  $\alpha_i$  和  $\tau_i$  是需要估计的参数， $\mathbf{V}_t$  是可以影响基金截面收益的一组系统性风险因子， $\epsilon_{it}$  是基金收益率里的特异性噪音。多因子模型 (1) 把基金的回报率拆解成了三部分：最重要的截距项  $\alpha_i$  表示基金能够在风险因子之外带来超额收益的能力。第二项  $\boldsymbol{\tau}_i^T \mathbf{V}_t$  是基金暴露在风险因子下所能得到的收益。最后一项  $\epsilon_{it}$  是特异性风险（噪音），我们认为其均值为 0，且和风险因子不相关。这里我们假设  $\epsilon_{it} \sim N(0, \sigma_i^2)$ ，在任意基金、任意时刻都是相互独立的，并且我们承认它有异方差性 (heteroskedasticity)。

我们常用的衡量对冲基金表现的模型有：

- (1) CAPM，用市场回报作为因子；
- (2) FFC 4 因子模型 (Fama, French, 1993; Carhart, 1997)，在 CAPM 上加入了规模、价值、动量；
- (3) FF 6 因子模型在 FFC 模型 (Fama, French, 2015) 的基础上加入了盈利和投资；
- (4) 9 因子模型 (Pastor, Stambaugh, 2002) 在 FF6 因子的基础上加入了三个行业因子。

Pastor and Stambaugh (2002) 和 Barber, Huang and Odean (2016) 提到, 当评估基金收益的时候投资者们应当考虑所有的能解释截面基金收益的因子。但是想要找出所有因子客观上是非常困难的。因此, 我们提议与其强行寻找所有可能的因子, 不如从基金回报率的截面中估算这些因子。这么做背后的逻辑是, 假如某个因子能解释很大一部分收益率, 且能在众多基金的收益表现里被发现, 那么该因子带来的这部分 alpha 就不能被认为是基金表现出色的指标。

令  $\mathbf{V}_t \equiv (\mathbf{F}_t, \mathbf{Z}_t)$  其中  $\mathbf{F}_t$  是我们观察到从而给定的因子 (observed factors), 而  $\mathbf{Z}_t$  是隐藏因子 (latent factors), 因此我们可以把 (1) 改写成如下形式:

$$R_{it} - R_{0t} = \alpha_i + \beta_i^T \mathbf{F}_t + \gamma_i^T \mathbf{Z}_t + \epsilon_{it} \quad (2)$$

同时我们令  $\delta = E[\mathbf{Z}_t]$  代表潜在因子风险溢价的期望。

在讨论如何用 (2) 模型做回归之前, 我们先简要说明下扰动偏差 (confounding bias) 或者缺失变量偏差 (omitted variable bias)。假设  $\gamma_i^T \delta \neq 0$  且模型里只含有  $\mathbf{F}_t$ , 那么 OLS 估算出的是  $\alpha_i$  的有偏估计。令  $\mathbf{R}_i$  为超额收益  $(R_{i1} - R_{01}, \dots, R_{iT} - R_{0T})$ ,  $\tilde{\mathbf{F}} = (1, \mathbf{F})$  其中  $\mathbf{F}$  是因子回报率矩阵。那么 OLS 估计如下:

$$\begin{pmatrix} \hat{\alpha}_{\text{OLS},i} \\ \hat{\beta}_{\text{OLS},i} \end{pmatrix} = (\tilde{\mathbf{F}}^T \tilde{\mathbf{F}})^{-1} \tilde{\mathbf{F}}^T \mathbf{R}_i$$

在标准的正则化条件下, 我们可以直接得到:

$$\hat{\alpha}_{\text{OLS},i} \xrightarrow{p} \alpha_i + \gamma_i^T \delta \quad (3)$$

从上式可以看出, 如果我们只考虑我们能观测到的因子 (observed factors), 我们会错误估计 “真实 alpha”。

假如在  $\delta = 0$  (即所有隐藏因子的回报均值为 0) 的特殊情况下, 此时 OLS 估计的  $\alpha_i$  和真实值是一致的。在这种情况下我们还是应该考虑在模型中加入潜在因子 (latent factor), 因为能减少估算的方差。

### 3.2.2. 估计潜在因子

我们按照 Wang et al.(2017) 中 Confounder Adjusted Testing and Estimation (CATE) 的步骤来估计潜在因子模型 (latent-factor model)。CATE 原本是用来剔除基因数据中的批次效应的, 在这里也可以用来消除经济面板数据中的扰动偏差。

#### (1) 估计潜在因子的因子荷载

我们首先通过因子分析来估计  $\gamma_i$ 。令  $p_F = \mathbf{F}(\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T$  为投向  $\mathbf{F}$  列空间的投影矩阵。在用线性模型只对已观测到的因子 (observed factors) 回归时, 残差矩阵可以表示为:

$$(I - P_{\tilde{F}})R = \gamma^T \left[ (I - P_{\tilde{F}})Z \right] + (I - P_{\tilde{F}})\hat{\epsilon} \quad (4)$$

其中  $R, Z, \hat{\epsilon}$  分别是超额收益，潜在 (latent factor) 和特异性误差的矩阵形式。(4) 的右边， $(I - P_{\tilde{F}})R$  是低秩矩阵  $\gamma^T \left[ (I - P_{\tilde{F}})Z \right]$  和误差  $(I - P_{\tilde{F}})\hat{\epsilon}$  之和。因此，潜在因子的荷载  $\gamma$  可以通过对残差矩阵  $(I - P_{\tilde{F}})R$  做因子分析获得。

我们使用 (Bai and Li, 2012) 中的极大似然法 (maximum likelihood method) 来估计  $\gamma$  和  $\Sigma \equiv \text{diag}(\sigma_1^2, \dots, \sigma_N^2)$ ：

$$(\hat{\gamma}, \hat{\Sigma}) = \arg \min_{\tilde{\gamma}, \tilde{\Sigma}} \left\{ -\log |\tilde{\gamma}(\tilde{\gamma})^T + \tilde{\Sigma}| - \text{tr} \left\{ S \left[ \tilde{\gamma}(\tilde{\gamma})^T + \tilde{\Sigma} \right]^{-1} \right\} \right\}$$

其中  $S$  是  $(I - P_{\tilde{F}})R$  基于样本数据的协方差矩阵， $\det(\cdot)$  和  $\text{tr}(\cdot)$  分别是矩阵的行列式和迹。当  $\hat{\epsilon}$  是同方差时 (homoskedastic)，以上最大似然法和主成分分析 (PCA) 等价。

## (2) 估算潜在因子的收益率和基金的 alpha

接下来我们要估算潜在因子 (latent factors) 的期望。基于以上因子分析中我们估算出的  $\gamma_i$ ，(3) 式可以表示为：

$$\hat{\alpha}_{\text{OLS},i} \approx \alpha_i + \hat{\gamma}_i^T \delta, \quad i=1, \dots, N \quad (5)$$

显然，我们可以从 (5) 中估算出  $\alpha$  和  $\delta$ 。这里我们需要考虑上式中有  $N + \dim(Z)$  个未知参数，比等式的数量略多。通过假设  $\alpha$  的近似稀疏性可以解决这个问题，我们假设：

$$\|\alpha\|_1 / N \rightarrow 0 \quad \text{当} \quad N \rightarrow \infty$$

此假设意味着尽管基金的数量非常多，但是好的投资机会并不多。基于此假设我们可以通过以下方法来估计  $\alpha$  和  $\delta$ 。我们认为在  $\hat{\alpha}_{\text{OLS}}$  对  $\hat{\gamma}$  回归时所得的  $\alpha$  值很大的话说明这是一个离群值 (outlier)。因此为了保证回归的稳定性在这里我们用稳健回归 (robust regression) 来估计  $\delta$ ：

$$\hat{\delta}_{\text{CATE}} = \arg \min_{\delta} \sum_{i=1}^N \rho \left( \frac{\hat{\alpha}_{\text{OLS},i} - \hat{\gamma}_i^T \delta}{\hat{\sigma}_i} \right) \quad (6)$$

其中  $\rho$  是损失函数， $\hat{\sigma}_i$  是上一步因子回归中估算出的特异性风险用来表示 (5) 中的异方差性 (heteroskedasticity)。同时对损失函数，我们也有以下假设：



- (i).  $\rho(0) = 0$
- (ii).  $x < 0$  时  $\rho(x)$  非增,  $x > 0$  时  $\rho(x)$  非减
- (iii).  $\rho(x)$  一阶导数存在且有界
- (iiii).  $\rho(x)$  二阶导 0 附近恒正

常用的  $\rho$  有 Huber's loss function, 我们在之前的指数增强模型中也用到过。

$$\rho_H(t) = \begin{cases} t^2 / 2, & \text{if } |t| \leq k \\ k^2 / 2 + k |t|, & \text{otherwise} \end{cases}$$

这里当我们令损失函数的参数  $k=1.345$  时, 可以达到正态分布样本 95% 的渐进有效性。而 OLS 损失函数不满足假设, 因为一阶导数不满足有界条件。过大或过小的  $\alpha$  都会对目标函数造成难以控制的影响, 从而导致估计偏差。

最后通过对 naïve OLS 估计除偏就可以得到  $\alpha_i$  的 CATE 估计:

$$\hat{\alpha}_{\text{CATE},i} = \hat{\alpha}_{\text{OLS},i} - \hat{\gamma}_i^T \hat{\delta}_{\text{CATE}} \quad (7)$$

上述的步骤是 CATE (Wang et al, 2017) 的一个特例, 这里我们只关注截距项。但是广义的 CATE 理论对于上述设定也是成立的。详细的技术细节请参考 (Wang et al, 2017)。

作为对比, 假如我们在 (6) 使用标准的二次损失函数, 即  $\rho(t) = t^2 / 2$ , 那估算的结果是有偏的。假设噪音项是同方差的 (homoskedastic),  $\sigma_i \equiv \sigma$ , 此时隐藏因子只有一个。那么基于二次损失函数的  $\delta$  估计就变成了  $\delta + \overline{\text{Cov}}(\alpha_i, \gamma_i) / \overline{\text{Var}}(\gamma_i)$ 。 $\overline{\text{Cov}}$  和  $\overline{\text{Var}}$  分别是截面的协方差和方差。 $\overline{\text{Cov}}(\alpha_i, \gamma_i) / \overline{\text{Var}}(\gamma_i)$  就是  $\alpha_i$  对  $\gamma_i$  的截面回归系数。因此对  $\alpha$  的估计就会有有很大的偏差。

综上所述, 我们用两步法来剔除 alpha 里的扰动偏差 (confounding bias):

1. 我们首先对 (4) 中的残差矩阵做因子分析并获得  $\hat{\gamma}$  和  $\hat{\sigma}$ 。这一步的关键假设是隐藏因子足够强使因子分析能够保持一致。
2. 然后我们用  $\hat{\alpha}_{\text{OLS}}$  对  $\hat{\gamma}$  回归并获得  $\hat{\delta}$  和  $\hat{\alpha}$ 。这一步的关键假设是  $\alpha$  的近似稀疏性使得它在 (5) 中可识别。

上述方法与 Fama and Macbeth(1973), Bai(2009), Giglio and Xiu(2016)所提及的方法紧密相关。其中最接近的是

Giglio and Xiu (2016)。Giglio and Xiu (2016) 的前两步分别使用主成分分析法提取隐藏因子以及 OLS 做截面回归，而上述的方法在第一步使用最大似然法来处理异方差性，在第二步用稳健回归做截面回归。

### 3.3. 潜在因子模型运用

#### 3.3.1. 模型解释

我们使用中信一级行业指数，自 2012 到 2020 的历史收益率序列数据。在实际操作中，我们使用市场、规模、价值、盈利、投资这五个因子作为我们的可观测到的因子（observed factors），即  $F_t \equiv (MKT_t, SMB_t, HML_t, RMW_t, CMA_t)$ ，因为这五个因子被认为是最稳健的因子，且我们之前研究认为能有效解释投资组合的收益。同时在实证分析中我们假定三个潜在因子（Bada and Liebl, 2014），即  $Z_t \equiv (Z_t^1, Z_t^2, Z_t^3)$ 。综上，我们的潜在因子模型表达形式如下：

$$R_{it} - R_{0t} = \alpha_i + b_i MKT_t + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + \sum_{k=1}^3 \gamma_i^k Z_t^k + \epsilon_{it} \quad (8)$$

RHS 的第一项  $\alpha_i$  是因子调整之后的平均收益，这里我们称它为  $CATE_{-\alpha}$ 。  $b_i, s_i, h_i, r_i, c_i$  and  $\gamma_i^k$  分别是行业指数在市场、规模、价值、动量以及隐藏因子上的暴露。

我们在以上提及的 CATE alpha、CAPM alpha 以及 FFC alpha 之间的关系如下所示：

$$\begin{aligned} & E(R_{it} - R_{0t}) - b_i E(MKT_t) \\ &= \underbrace{\alpha_i + \sum_{k=1}^3 \gamma_i^k E(Z_t^k)}_{\text{CATE}_{-\alpha} \quad \text{Latent factor}} + \underbrace{s_i E(SMB_t) + h_i E(HML_t) + r_i E(RMW_t) + c_i E(CMA_t)}_{\text{Observed factor (non-market)}} \\ & \quad \text{CAPM}_{-\alpha} \\ & \quad \text{FFC}_{-\alpha} \end{aligned}$$

FFC alpha 是 CATE alpha 和隐藏因子  $\gamma_i^T E(Z_t)$  的平均回报均值之和。 $CAPM_{-\alpha}$  在 FFC alpha 的基础上还包含了规模、价值、盈利、投资的平均收益。

#### 3.3.2. 潜在因子模型的回归展示

下表展示了添加潜在因子后，因子模型对行业指数的解释度大幅上升。alpha 大幅下降，表示 FF-5 的 alpha 不是很纯，包含了共同因子对收益贡献的部分，考虑潜在因子后，这部分被剔除，alpha 变得更“纯洁”，数值大幅下降。下表是用 19 年下半年的数据做展示，我们能看出，煤炭、基础化工、国防军工等行业的 R 方有大幅

上升，上升幅度都是10%左右。Alpha 数值除了家电、非银、银行外，基本上都有所下降，但是下降的幅度不是一致的，所以，提纯后的“纯洁”alpha 并不是原先 alpha 的简单映射，他们的排序是不相同的，利用 FF-5 alpha 和“纯洁”alpha 得出排名靠前的行业会有所不同。

**表 9：考虑潜在因子下的 alpha 与原来 alpha 对比（20190701-20191231）**

中信一级行业	FF-5			FF-5 & Latent factor		
	Alpha annualized	P value	R.Square	Alpha annualized	P value	R.Square
石油石化	0.48%	0.97	63.16%	-7.37%	0.49	73.40%
煤炭	7.46%	0.60	71.14%	-3.22%	0.77	83.09%
有色金属	36.79%	0.12	34.87%	25.97%	0.24	44.12%
电力及公用事业	-8.07%	0.38	74.93%	-14.26%	0.06	83.23%
钢铁	8.18%	0.57	66.88%	-3.42%	0.75	82.12%
基础化工	6.63%	0.58	79.74%	-4.97%	0.45	93.78%
建筑	-6.18%	0.58	79.04%	-15.24%	0.06	89.07%
建材	15.60%	0.34	67.61%	7.11%	0.63	73.78%
轻工制造	11.09%	0.35	79.08%	2.09%	0.82	87.69%
机械	4.26%	0.65	87.00%	-4.67%	0.39	95.64%
电力设备及新能源	4.71%	0.73	75.41%	-6.47%	0.52	87.01%
国防军工	11.81%	0.58	56.35%	-0.41%	0.98	66.21%
汽车	21.21%	0.11	72.99%	12.81%	0.26	80.58%
商贸零售	-8.35%	0.38	80.67%	-15.12%	0.05	87.72%
消费者服务	-18.12%	0.39	42.39%	-12.54%	0.55	44.86%
家电	10.34%	0.61	57.87%	14.03%	0.48	58.56%
纺织服装	-1.84%	0.85	78.78%	-8.88%	0.28	86.22%
医药	-14.46%	0.35	67.35%	-14.91%	0.34	67.10%
食品饮料	-17.41%	0.39	53.19%	-5.13%	0.77	65.29%
农林牧渔	-8.60%	0.80	20.32%	-9.88%	0.78	19.73%
银行	16.16%	0.20	63.03%	23.79%	0.03	72.56%
非银行金融	-0.15%	0.99	83.27%	1.47%	0.90	83.34%
房地产	14.28%	0.36	67.95%	10.31%	0.50	69.20%
交通运输	-13.22%	0.14	83.70%	-16.64%	0.06	85.29%
电子	28.60%	0.16	78.05%	17.99%	0.32	82.36%
通信	-4.40%	0.81	78.74%	-19.70%	0.15	88.85%
计算机	15.42%	0.43	79.69%	-2.09%	0.87	91.69%
传媒	43.31%	0.01	75.22%	30.55%	0.03	84.56%
综合	3.04%	0.79	83.30%	-6.61%	0.43	91.59%
综合金融	20.07%	0.44	67.64%	9.58%	0.70	71.16%

资料来源：中泰证券研究所

### 3.4. “纯洁”alpha 指标下的回测分析

和上文一样，每次回归计算alpha用到的是过去125天的收益率数据，回归加入了潜在因子。关于未来持有期限，我们下面分别展示了月度调仓和季度调仓。为了获得较好的参数，我们对分三组和分五组，都进行了试验。

下面分别展示了分五组按季度调仓、分五组按月度调仓、分三组按季度调仓、分三组按月度调仓的回测结果。分组数字越大的是alpha越大的组。

### (1) 分五组按季度调仓

图 10: FF5-Latent-Alpha 因子分组回测 (分五组, 季度调仓)



资料来源: 中泰证券研究所

图 11: FF5-Latent-Alpha 因子下的多空回测 (分五组, 季度调仓)



资料来源: 中泰证券研究所

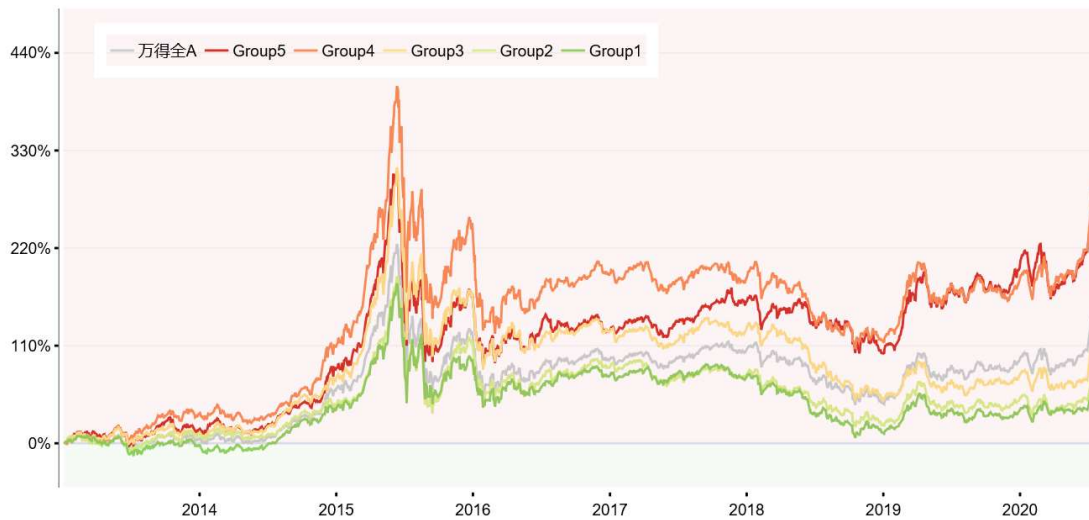
表 10: FF5-Latent-Alpha 因子分组回测及多空表现 (分五组, 季度调仓)

	万得全 A	Group5 - 1	Group5	Group4	Group3	Group2	Group1
累计收益率	134.77%	132.64%	287.68%	172.40%	110.88%	104.96%	60.14%
年化收益率	12.38%	12.24%	20.37%	14.69%	10.75%	10.32%	6.65%
年化夏普比	0.46	1.03	0.72	0.51	0.38	0.37	0.24
年化波动率	26.65%	11.94%	28.34%	28.93%	28.55%	28.23%	27.77%
最大回撤	55.99%	26.79%	54.90%	64.93%	60.13%	58.07%	59.17%
月度胜率		47.25%	62.64%	52.75%	47.25%	45.05%	39.56%



(2) 分五组按月度调仓

图 12: FF5-Latent-Alpha 因子分组回测 (分五组, 月度调仓)



资料来源: 中泰证券研究所

图 13: FF5-Latent-Alpha 因子下的多空回测 (分五组, 月度调仓)



资料来源: 中泰证券研究所

表 11: FF5-Latent-Alpha 因子分组回测及多空表现 (分五组, 月度调仓)

	万得全 A	Group5 - 1	Group5	Group4	Group3	Group2	Group1
累计收益率	134.77%	144.02%	294.41%	261.81%	91.06%	66.55%	56.84%
年化收益率	12.38%	12.98%	20.65%	19.24%	9.26%	7.23%	6.35%
年化夏普比	0.46	1.11	0.72	0.67	0.33	0.25	0.23
年化波动率	26.65%	11.70%	28.68%	28.72%	27.99%	28.43%	27.73%
最大回撤	55.99%	20.29%	54.01%	59.00%	64.15%	59.03%	62.24%
月度胜率		46.15%	61.54%	59.34%	40.66%	37.36%	41.76%

资料来源: 中泰证券研究所

(3) 分三组按季度调仓

图 14: FF5-Latent-Alpha 因子分组回测（分三组，季度调仓）



资料来源：中泰证券研究所

图 15: FF5-Latent-Alpha 因子下的多空回测（分三组，季度调仓）



资料来源：中泰证券研究所

表 12: FF5-Latent-Alpha 因子分组回测及多空表现（分三组，季度调仓）

	万得全 A	Group3 - 1	Group3	Group2	Group1
累计收益率	134.77%	69.39%	232.01%	110.52%	92.96%
年化收益率	12.38%	7.48%	17.84%	10.72%	9.41%
年化夏普比	0.46	0.84	0.63	0.38	0.34
年化波动率	26.65%	8.89%	28.23%	28.33%	27.60%
最大回撤	55.99%	26.43%	58.30%	61.71%	58.01%
月度胜率		46.15%	58.24%	50.55%	42.86%

(4) 分三组按月度调仓

图 16: FF5-Latent-Alpha 因子分组回测（分三组，月度调仓）



资料来源：中泰证券研究所

图 17: FF5-Latent-Alpha 因子下的多空回测（分三组，月度调仓）



资料来源：中泰证券研究所

表 13: FF5-Latent-Alpha 因子分组回测及多空表现（分三组，月度调仓）

	万得全 A	Group3 - 1	Group3	Group2	Group1
累计收益率	134.77%	135.60%	283.10%	117.69%	60.43%
年化收益率	12.38%	12.44%	20.17%	11.23%	6.68%
年化夏普比	0.46	1.40	0.71%	0.40	0.24
年化波动率	26.65%	8.88%	28.47%	27.82%	27.73%
最大回撤	55.99%	20.65%	54.94%	61.74%	60.04%
月度胜率		47.25%	65.93%	46.15%	39.56%

资料来源：中泰证券研究所

## 4. 总结

评价上文不同参数下的策略表现，主要看多空对冲组合的夏普比，夏普比越高，代表策略表现越好、稳健性越好，所以采用分三组月度调仓的模式，模型表现最好。下表展示了该参数下，普通 FF-5alpha 和纯洁 alpha 策略业绩对比，纯洁 alpha 对冲组合夏普为 1.40，普通 alpha 下的对冲组合夏普为 1.26，夏普比有一定的提升。纯洁 alpha 下的多头组合年化收益为 20.17%，普通 alpha 下的年化收益为 19.46%，看似年化收益提升不多，实则潜在因子下的纯洁 alpha 在行业选择上已经有了大幅提升，因为我们组合选取十个优秀行业等权配置，纯洁 alpha 新选入的行业收益一定是有大幅的提升，才给组合带来相对传统 alpha 的 0.84% 的年化超额收益。

表 14: FF-5alpha 和“纯洁”alpha 回测效果对比（下面一行 为“纯洁”alpha）

	万得全 A	Group3 - 1	Group3	Group2	Group1
累计收益率	134.77%	131.04%	266.80%	138.90%	54.05%
		135.60%	283.10%	117.69%	60.43%
年化收益率	12.38%	12.14%	19.46%	12.65%	6.09%
		12.44%	20.17%	11.23%	6.68%
年化夏普比	0.46	1.26	0.69	0.45	0.21
		1.40	0.71	0.40	0.24
年化波动率	26.65%	9.59%	28.21%	27.93%	28.04%
		8.88%	28.47%	27.82%	27.73%
最大回撤	55.99%	15.84%	53.50%	61.13%	61.79%
		20.65%	54.94%	61.74%	60.04%
月度胜率		45.05%	65.93%	49.45%	39.56%
		47.25%	65.98%	46.15%	39.56%

资料来源：中泰证券研究所

图 18: FF5-Latent-Alpha 因子下的多空回测（分三组，月度调仓）



资料来源：中泰证券研究所



表 15：纯洁 alpha 动量策略下的历史行业配置(月末权重应用在下个月)

	20191129	20191231	20200123	20200228	20200331	20200430	20200529	20200630	20200731
煤炭									
有色金属	0.1	0.1	0.1			0.1	0.1	0.1	0.1
钢铁									
建筑									
国防军工	0.1		0.1			0.1	0.1	0.1	0.1
汽车	0.1	0.1	0.1	0.1	0.1	0.1			
银行	0.1	0.1	0.1	0.1	0.1				
非银行金融	0.1	0.1		0.1					0.1
电子	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
综合			0.1	0.1					
基础化工									
轻工制造		0.1			0.1				
家电		0.1	0.1	0.1	0.1		0.1		
医药				0.1	0.1	0.1	0.1	0.1	0.1
电力及公用事业									
房地产		0.1							
石油石化			0.1						
通信									
传媒	0.1	0.1	0.1	0.1	0.1	0.1	0.1		
计算机	0.1			0.1	0.1				0.1
建材	0.1	0.1	0.1	0.1					
商贸零售								0.1	0.1
交通运输								0.1	
电力设备及新能源						0.1	0.1	0.1	0.1
农林牧渔					0.1	0.1	0.1	0.1	0.1
消费者服务						0.1	0.1	0.1	0.1
食品饮料	0.1				0.1	0.1	0.1	0.1	0.1
机械									
纺织服装									

资料来源：中泰证券研究所

风险提示：模型根据历史数据建立，不排除未来一段时间内失效的风险。

## 附录

### 1、参考文献

- [1] Bai, J. and Li, K. (2012). Statistical analysis of factor models of high dimension. The Annals of Statistics. 40, 436-465.
- [2] Miralles-Quirós, J.L.; Miralles-Quirós, M.M.; Nogueira, J.M. (2020). Sustainable Development Goals and Investment Strategies: The Profitability of Using Five-Factor Fama-French Alphas. Sustainability 12, 1842.
- [3] Sarwar, G.; Mateus, C.; Todorovic, N. (2018). US sector rotation with five-factor Fama – French alphas. J. Asset Manag. 19, 116 – 132
- [4] Wang, J., Zhao, Q., Hastie, T. and Owen, A. B. (2017). Confounder adjustment in multiple hypothesis testing. Annals of Statistics 45, 1863-1894.
- [5] Fama, E.F.; French, K.R. (2017). International tests of a five-factor asset pricing model. J. Financ. Econ. 123, 441 – 463
- [6] Foye, J. (2018). A comprehensive test of the Fama-French five-factor model in emerging markets. Emerg. Mark. Rev. 37, 199 – 222
- [7] Blitz, D.; Huij, J. (2012). Evaluating the performance of global emerging markets equity exchange-traded funds. Emerg. Mark. Rev. 2012, 13, 149 – 158
- [8] Bai, J. (2009). Panel data models with interactive fixed effects. Econometrica 77, 1229-1279.
- [9] Barber, B. M., Huang, X. and Odean, T. (2016). Which factors matter to investors? Evidence from mutual fund flows. Review of Financial Studies 29, 2600-2642.
- [10] Kacperczyk, M., Sialm, C. and Zheng, L. U. (2005). On the Industry Concentration of Actively Managed Equity Mutual Funds, Journal of Finance, 60(4), pp. 1983 – 2012.
- [11] Pástor, L. and Stambaugh, R. F. (2002). Mutual Fund Performance and Seemingly Unrelated Assets, Journal of Financial Economics, 63(3), pp. 315 – 349.
- [12] Dou, P. Y., Gallagher, D. R., Schneider, D. and Walter, T. S. (2014). Cross-region and Cross-sector Asset Allocation with Regimes, Accounting and Finance, 54, pp. 809 – 846.
- [13] Frazzini, A. and Pedersen, L. H. (2014). Betting Against Beta. Journal of Financial Economics 111, 1-25.
- [14] Beller, K.R., J.L. Kling, and M.J. Levinson. (1998). Are Industry Stock Returns Predictable? Financial Analysts Journal 54 (5): 42 – 57.
- [15] Pástor, L., Stambaugh, R. F. and Taylor, L. A. (2015). Scale and skill in active management. Journal of Financial Economics 116 23 - 45.

- [16] Bada, O. and Liebl, D. (2014). The R Package phtt: Panel Data Analysis with Hetero-geneous Time Trends. Journal of Statistical Software 59 1-34.
- [17] Giglio, S. and Xiu, D. (2016). Asset pricing with omitted factors. Technical report.
- [18] Sassetti, P. and Tani, M. (2006). Dynamic Asset Allocation Using Systematic Sector Rotation, The Journal of Wealth Management, Vol.8(4), pp.59 – 70.
- [19] Mohanty, S.S. (2019). Does one model fit all in global equity markets? Some insight into market factor based strategies in enhancing alpha. Int. J. Financ. Econ. 24, 1170 – 1192.
- [20] Angelidis, T.; Giamouridis, D.; Tassaromatis, N.(2013). Revisiting mutual fund performance evaluation. J. Bank. Financ. 37, 1759 – 1776.
- [21] Maronna, R. A., Martin, D. R. and Yohai, V. J. (2006). Robust statistics: Theory and Methods. John Wiley & Sons, Chichester.

### 投资评级说明：

	评级	说明
股票评级	买入	预期未来 6~12 个月内相对同期基准指数涨幅在 15%以上
	增持	预期未来 6~12 个月内相对同期基准指数涨幅在 5%~15%之间
	持有	预期未来 6~12 个月内相对同期基准指数涨幅在-10%~+5%之间
	减持	预期未来 6~12 个月内相对同期基准指数跌幅在 10%以上
行业评级	增持	预期未来 6~12 个月内对同期基准指数涨幅在 10%以上
	中性	预期未来 6~12 个月内对同期基准指数涨幅在-10%~+10%之间
	减持	预期未来 6~12 个月内对同期基准指数跌幅在 10%以上

备注：评级标准为报告发布日后的 6~12 个月内公司股价（或行业指数）相对同期基准指数的相对市场表现。其中 A 股市场以沪深 300 指数为基准；新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准；香港市场以摩根士丹利中国指数为基准，美股市场以标普 500 指数或纳斯达克综合指数为基准（另有说明的除外）。

### 重要声明：

中泰证券股份有限公司（以下简称“本公司”）具有中国证券监督管理委员会许可的证券投资咨询业务资格。本报告仅供本公司的客户使用。本公司不会因接收人收到本报告而视其为客户。

本报告基于本公司及其研究人员认为可信的公开资料或实地调研资料，反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响。但本公司及其研究人员对这些信息的准确性和完整性不作任何保证，且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，可能会随时调整。本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。本报告所载的资料、工具、意见、信息及推测只提供给客户作参考之用，不构成任何投资、法律、会计或税务的最终操作建议，本公司不就报告中的内容对最终操作建议做出任何担保。本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。

市场有风险，投资需谨慎。在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。

投资者应注意，在法律允许的情况下，本公司及其本公司的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能为这些公司正在提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。本公司及其本公司的关联机构或个人可能在本报告公开发布之前已经使用或了解其中的信息。

本报告版权归“中泰证券股份有限公司”所有。未经事先本公司书面授权，任何人不得对本报告进行任何形式的发布、复制。如引用、刊发，需注明出处为“中泰证券研究所”，且不得对本报告进行有悖原意的删节或修改。