

文章编号:1005-2542(2020)03-0417-08

中美股市配对因子实证分析

周志中, 徐 杰

(上海交通大学 安泰经济与管理学院, 上海 200030)

【摘要】配对交易是基于统计套利,利用两个资产的短暂价格偏离,进行风险对冲以获取两个资产的 Alpha 收益,其核心假设是配对资产的价差具有均值回复性。多因子选股则将多个具有逻辑背景的因子策略相结合,选取在各个因子上综合得分较高的股票构建投资组合,其核心是如何挖掘具有逻辑背景的因子。由于中美经济关联度很大,两国股市间存在着联动性。根据配对交易的思想构建美股-A 股配对因子,利用该因子对 A 股进行选股,通过实证分析,发现配对因子对 A 股股票有较好的区分度,且与传统的选股因子存在较低的相关性。因此,根据配对交易思想构建的美股-A 股配对因子可用于多因子模型进行选股。贡献在于提出了以往研究未提出的因子,通过实证研究证明了该因子的有效性。

关键词:量化投资; 配对交易; 算法交易; 中美金融市场联动

中图分类号:F 830.9

文献标志码:A

DOI:10.3969/j.issn.1005-2542.2020.03.001

An Empirical Analysis of Pairing Factors in Sino-US Stock Markets

ZHOU Zhizhong, XU Jie

(Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai 200030, China)

【Abstract】 Paired trading and multi-factor stock selection are typical quantitative trading strategies. The former is based on statistical arbitrage, which takes advantage of the short-term price deviation of two assets and then hedges the risk to obtain the Alpha return of the two assets. The core assumption is that the spread of paired assets has the property of mean reversion. The latter selects stocks with high comprehensive scores on factors with logical meanings. Due to the high degree of economic connection between China and the United States, there exists price co-movement in the stock markets of the two countries. We construct US equities-A share paired factor according to the ideas of the paired trading. Then the factor is used to select stocks. The empirical analysis reveals that the constructed factor has a relatively good performance in differentiating A share. And it has a low correlation with traditional stock selection factors. Therefore, the factor can be used in the multi-factor model to select stocks. The contribution of this paper lies in the fact that we propose a stock selection factor not examined in the previous research, and verify the validity of this factor in the empirical research.

Key words: quantitative investment; paired trading; algorithmic trading; correlation of the Sino-US financial market link

在经济全球化的今天,跨国商品与服务贸易及

资本流动规模和形式不断增加,技术的广泛迅速传播使世界各国经济的相互依赖性增强。经济全球化有 4 大表现:贸易自由化、生产国际化、资本全球化和科技全球化。2008 年全球金融危机后,美国和中国在经济和文化的联系和合作上不断深化。表现在 2008 年中国成为美国国债的第一大持有者。两国还扩展经济与战略对话机制、进行中美投资协定

收稿日期:2019-02-24 修订日期:2019-09-08

基金项目:国家自然科学基金资助项目(71771148,71371121,71531010,71421002)

作者简介:周志中(1975-),男,博士,副教授。研究方向为金融科技、量化投资和算法交易以及信息系统经济学。

E-mail:zhouzhzh@outlook.com

(BIT)谈判。美国还推出“十万强”“百万强”计划,并对中国护照实行首次 10 年免签。中国资本市场也逐步走向开放和自由化,从发行外资股(B 股)到单向开放(QFII、QDII 和 RQFII)到现在的双向开放阶段(沪港通、深港通、沪伦通、中日 ETF 互联互通),中国金融市场和国际金融市场尤其是和美国金融市场的联动性也在不断增强,这在以往实证研究中也得到证实^[1]。

以往量化选股模型仅利用国内股票市场的信息进行选股,而中美金融市场的联动性为构建新的量化选股模型提供了新的思路。本文旨在利用中美股市内在联系,通过对标中国 A 股股票的美股股票价格差,为中国 A 股股票建立选股因子。这个选股因子构建的思想在以往论著中甚少进行深入研究,本文使用配对交易的思想找到和美股股票对标的 A 股股票,并利用两者价差构建配对因子,然后利用配对因子对 A 股进行选股,实证研究证明了该因子的有效性。

量化投资策略大致可以分为股票策略、宏观策略和统计套利策略 3 类。股票策略主要是指单一的应用于股票市场的交易策略。中国的股票市场目前主要存在的股票策略是多因子选股模型和基于流动性的高频交易策略。宏观策略投资范围不局限于单一类型的市场,而是进行全类型市场的投资,其中包括动量策略的代表策略 CTA 策略。统计套利策略是一种模型驱动的投资过程,当资产价格偏离其理论价格或模型预测价格时,则通过同时构建多头和空头组合而获取利润。配对交易策略是统计套利的一种。本文的思路包括了股票策略中的多因子选股和统计套利策略中的配对交易。

多因子模型是量化投资领域应用最广泛也是最成熟的量化选股模型之一,建立在投资组合、资本资产定价(CAPM)以及套利定价理论(APT)等现代金融投资理论基础上。此处的因子是股票价格影响因素的简称,或简单理解为指标。候选因子的选取大致分为基本面指标因子与技术面指标因子两大类。前者主要是基于对股票基本面的分析,如 ROA、ROE 以及资产负债率等,后者主要是基于股票技术层面的指标,如换手率、波动率以及流通市值等。例如,无论国内外,小市值效应一直存在,即小盘股比大盘股的收益率高。其内在逻辑在于,市值越小的股票,被操纵的可能性就越大,存在较大市场风险,需要用较高收益率来进行风险补偿,因此,规模因子(包括总市值、流通市值、自由流通市值等)是一个有效的选股因子。多因子模型的缺陷在于,因

子来源基本是基本面指标和技术指标,同质化很高,也没有反映外部因素的影响,这使得因子表现很不稳定甚至经常失效,影响策略表现,因此需要挖掘新的因子和可以体现外部因素影响的因子。

配对交易是常用的统计套利投资策略,主要是基于两只股票的价差来盈利。配对交易的基本原理是:两个具有高度协整关系的股票,股价价差长期来看会比较稳定,但稳定性有时会因投资者的非理性交易而打破,导致价差偏离长期均衡值。但这种价格偏差可能只是暂时现象,未来价差还是会回到长期均衡值。配对交易就是利用价格偏差来进行交易:一旦两只股票价差扩大,则卖空相对高估的股票,买入相对低估的股票,然后等待价差回复长期均衡值;一旦价差回归长期均衡值,则同时将多头头寸和空头头寸平仓,获得套利收益。例如,中国平安同时在上海和香港上市,分别为中国平安 A 股和中国平安 H 股,可以作为配对股票。如果中国平安 H 股短时间内大幅上涨,而中国平安 A 股涨幅不大,使得两者的价差偏离了均值,则可以卖空中国平安 H 股而买入中国平安 A 股。当价差回归长期均衡值时同时平仓获取套利收益。经过市场验证,配对交易不仅可以在成熟的资本市场获利,也是 A 股市场中一种可获利的市场中性策略^[2]。

随着中国金融市场对外开发程度越来越大,尤其是 QDII 实施后,中美股市联动性也越来越大。从投资者的角度,可以通过研究中美股市之间的内在联系,挖掘新的选股因子——配对因子,将配对交易思想应用到 A 股市场投资中,提高选股策略的表现和可行性。配对因子策略的核心思想是,寻找中美股市间具有高度协整关系的股票进行配对,计算配对股票的价差与历史均值的偏离度作为配对因子。配对因子反映了 A 股股票的低估程度,低估程度越高,后续反弹的几率则越大。因此,在构建投资组合时,应选取低估程度最大的 N 只股票。具体操作步骤为,通过使调仓日前 1 个月内的标准化的中美股票价格序列间的平方距离最小化来选取相应的配对,并将平方距离最小的 300 对配对股票定义为“相似”股票;然后,将“相似”股票中的美股价格对数减去 A 股价格对数,得到“A 股-美股配对因子”。从“相似”股票对中选取“A 股-美股配对因子”值最大的 50 只 A 股,作为当期的投资组合进入交易期,一直持有到下个调仓日,重复上述过程。以构建 2018 年 2 月的投资组合为例,首先计算 2018 年 2 月每只 A 股与同行业所有美股的价格序列间的平方距离,作为该 A 股的配对距离。经过计算,万里

股份、现代制药等300只股票的配对距离最小,将它们作为2018年2月的配对股票库。在2018年2月的第1个调仓日,计算所有配对股票的股价对数价差与配对期内股价对数价差均值的差值,作为配对因子。选取配对因子最大的50只股票,等权配置,持有到2018年3月的第1个调仓日。

选取A股与美股市场的股票为样本,通过构建配对因子的方法来构建投资组合,检验其在2016~2018年的市场表现,探讨该策略的投资风险和可行性,从而提出有价值的投资策略。经过实证分析,配对因子与传统因子的相关性很低,通过配对因子构建的配对交易策略在近3年都取得了稳定的超额收益,说明了配对因子和配对策略的有效性和稳健性。

1 文献综述

随着中国商品市场和金融市场的不断开放和国际化,中美之间的经济联系越来越密切,这也体现在中美股票市场的联动性上。张金萍等^[3]在DCC-GARCH模型基础上,发现上证指数收益率和美国道琼斯指数收益率在2003年后一直保持正相关。莫悠等^[4]使用Johansen协整检验和Granger因果检验发现,中美股市具有长期均衡的正相关关系。刘继明^[5]在GARCH(1,1)-M模型基础上,发现中美两国股市联动性有不断增强的趋势。唐勇等^[6]基于小波视角发现,中美股市在短期尺度上联动效应较弱,而在长期尺度上联动效应较强。李合龙等^[7]使用集成经验模态分解(Ensemble Empirical Mode Decomposition, EEMD)发现,中美股市中长期周期下联动效应明显。赵若瑜^[8]通过研究中、美、英、日四国以及香港股市之间的关系,发现上证综合指数与道·琼斯指数、伦敦金融时报指数、日经225指数和香港恒生指数之间存在联动性。这些研究加深了本文对中美股市间协整关系的理解,也是尝试利用美国股市标的,在中国股市实现配对交易的理论基础。

以往对配对交易的研究主要聚焦于对传统配对交易策略的改进或检验配对交易在不同市场的有效性。在改进传统配对交易策略方面,麦永冠等^[9]构建了WM-FTBD建仓改进策略;黄晓薇等^[10]使用了基于OU过程的配对交易策略;胡伦超等^[11]使用了基于协整和距离的两阶段方法;刘永辉等^[12]使用了基于协整-OU过程的配对交易策略;胡文伟等^[13]使用了基于强化学习的自适应配对交易模型;陈晓芬等^[14]对配对交易进行了配对阶段和策略阶段的双重改进;毕秀春等^[15]使用动态优化方法求解带止损条件的配对交易最优阈值。在检验配对交易有效

性方面,赵胜民等^[16]检验了配对交易在A股市场上的有效性;于孝建等^[17]检验了配对交易在商品期货市场上的有效性;徐杰等^[18]检验了配对交易在中美跨市场交易中的有效性。与上述文献的研究内容不同,本文没有直接研究配对交易的策略和有效性,而是使用配对交易的思想构建配对因子,然后利用配对因子进行选股并构建投资组合。

多因子模型是一类重要的选股模型,核心思想是根据股票在各个因子下的得分给股票进行打分。各种多因子模型的核心区别为因子的选取和构造以及打分的方法。如何选取和构造因子,是多因子模型的关键步骤,也是中外学者研究的热点。Piotroski^[19]选取了9个财务指标对股票进行打分,涉及盈利能力、运营效率、杠杆率以及流动性和资金来源4个方面。各个指标设定一个临界值,如果达到或超过临界值,则得1分,否则不得分。最后,根据9项指标的总分进行股票挑选,构建投资组合。Mohanram^[20]在文献[19]的基础上进行了改进,选取盈利能力、财务的稳健性以及增长的稳定性3个层面的指标,进行投资组合构建。国内对因子的研究主要集中在券商的研究报告中,其中比较经典的是安信证券的金融工程分析师潘凡的研究报告,他在模型中选取了30个常见的选股因子,并验证了它们的有效性。上述研究的因子基本上可以分为基本面和技术分析两大类,主要是根据股票的财务指标和价格走势进行因子挖掘,并未涉及中美股市间的个股联系。本文贡献则是通过中美股市间的个股联系进行股票配对,构造配对因子,加入到多因子模型中。本文首次提出了配对因子,并通过实证研究证实其有效性。

2 实证分析

A股市场对融券卖空进行限制,是导致配对交易作用受限的主要原因。2015年的股灾,使得限制甚至禁止融券卖空的声音甚嚣尘上。此后,很多大中型券商暂停了融券券源供给,上海证券交易所和深圳证券交易所将融券交易规则从 $T+0$ 改为 $T+1$,很大程度上对融券卖空进行了限制。对卖空进行限制,是否有利于维护市场稳定,推动股市的繁荣发展,还存在很多争论,未来是否会放开限制也存在很大变数。因此,本文需要研究在这些限制客观存在的情况下,如何使用配对交易的思想在A股市场进行量化投资。

由于中美股市具有联动性,很多机构投资者会根据美股的表现来调整自己在A股的投资策略,因

而也使得部分 A 股与美股走势存在较强的相关性。本文通过挖掘这种相关性来进行统计套利。由于 A 股市场中融券标的稀缺,很多情况下难以进行传统意义上的配对交易,因而希望使用配对交易的思想,将 A 股-美股价差构建为一个因子,放入多因子选股模型,进而提高 A 股买入标的股票的选股精准性。具体交易策略:首先,将 A 股与美股进行配对;然后,使用美股相似股票价格对数减去 A 股个股价格对数,得到一个“A 股-美股配对因子”,如果该因子的值减去均值是正数而且变大,则意味着 A 股被低估,应该买入 A 股头寸;如果该因子减去均值是负数并且变小,则意味着 A 股被高估,应该卖出 A 股头寸。

配对因子的经济意义,是用来刻画某只 A 股相对于对标美股的低估或高估程度。由于指数报价在时间序列上具有一致性,可以解决数据不对称问题,因而本文用当日股价对数价差与形成期内的股价对数价差期望的偏离度刻画低估或高估程度。配对因子本质上是一个反转投资策略因子,核心是股票价格均值回归。文献[21]中对美国纽约股票市场进行了实证研究,得出股票收益长期呈均值回归的结论。文献[22]中的研究同样证明了股票价格均值回归的存在。

2.1 股票配对

配对交易在资本市场的作用已经得到了市场验证,本文采用最小化偏差平方和非参数化的方法进行股票配对。最小化偏差平方和方法,是用与给定股票价格时间序列的价差平方和最小的股票构成股票对,相应的阈值是配对股票价差序列的样本标准差的 2 倍,超过阈值则进行交易^[23]。为了提高策略的可解释性,本文在同行业内对中美股票进行配对。股票行业划分的依据是 Wind 一级行业,具体如表 1 所示。

表 1 Wind 一级行业列表

行业代码	行业名称	行业代码	行业名称
10	能源	35	医疗保健
15	材料	40	金融
20	工业	45	信息技术
25	可选消费	50	电信服务
30	日常消费	55	公用事业

在同一行业内,计算每一只 A 股的配对美股。首先要规定一个配对的形成期,即选取多久的时间长度进行股票配对。确定形成期后,需要将股票价格序列进行数据标准化。设股票 i 的价格序列为 $\{P_{it}\}$,标准化的方法为

$$P_{it}^* = \frac{P_{it} - E(P_{it})}{\sigma_i} \quad (1)$$

式中, $E(P_{it})$ 和 σ_i 分别为 P_{it} 的期望值与标准差。其中:

$$E(P_{it}) = \frac{1}{N} \sum_{t=1}^N P_{it} \quad (2)$$

$$\sigma_i = \sqrt{\frac{1}{N} \sum_{t=1}^N [P_{it} - E(P_{it})]^2} \quad (3)$$

数据标准化后要确定形成期和交易期。形成期指选取多久的时间长度进行配对距离计算,交易期指选取多久的时间长度进行交易。以每个交易日前 1 个月为形成期,按月调仓,采用最小化偏差平方和的方法进行配对。对每只 A 股 i ,计算它与美股 j 的配对距离 d 。形成期内 A 股的股价序列为 $\{C_{it}\}$,美股的股价序列为 $\{U_{jt}\}$,则具体计算公式为

$$d = \sum_{t=1}^N (C_{it} - U_{jt})^2 \quad (4)$$

选取配对距离最小的美股,作为给定 A 股的配对股票,并记录配对距离,部分配对数据如表 2 所示。配对距离的定义涉及两种货币,但是汇率因素对配对因子对 A 股股票的区分度影响不大。原因是已有实证研究发现,中国 A 股股价和汇率存在长期均衡的协整关系^[24-25]。这就意味着,如果人民币汇率突然发生变化,则所有 A 股股价会一起远离(或接近)美股股价,不会有的接近、有的远离。这样只会造成所有配对距离一起变大或一起变小,不会有的变大、有的变小。因此,配对距离的相对位置不会因为汇率突然发生变动而改变。本文选股时只使用配对距离的相对位置检查哪些配对距离最小,而不关心配对距离的绝对值,因此,汇率因素对配对因子对 A 股股票的区分度影响不大。

2.2 股票选择

在每个调仓日期,按照股票配对的结果,通过配对距离,对 A 股市场的 3 000 余只股票进行排序,选取配对距离最小的 300 只股票作为股票库。

2.3 配对因子计算

在每个调仓日期,计算每只股票的配对因子 m 。当日 A 股 i 标准化后的股价为 C_i ,配对美股 j 标准化后的股价为 U_j ,形成期内 A 股的股价序列为 $\{C_{it}\}$,美股的股价序列为 $\{U_{jt}\}$,则配对因子 m 的计算方法为

$$m = (\ln U_j - \ln C_i) - E(\ln U_{jt} - \ln C_{it}) \quad (5)$$

式中: $\ln U_j - \ln C_i$ 为调仓日配对美股 j 与 A 股 i 的股价对数的价差; $E(\ln U_{jt} - \ln C_{it})$ 为形成期内配对美股 j 与 A 股 i 的股价对数价差的期望值,计算公式为

表 2 中美股市股票配对表

配对日期	A 股代码	美股代码	配对日期	A 股代码	美股代码
2018-02-01	600847.SH	WPRT.O	2018-02-01	600074.SH	RENN.N
2018-02-01	600420.SH	BASI.O	2018-02-01	002456.SZ	RENN.N
2018-02-01	600289.SH	WSTG.O	2018-02-01	600679.SH	REVG.N
2018-02-01	300497.SZ	ADVM.O	2018-02-01	002618.SZ	RENN.N
2018-02-01	002139.SZ	RENN.N	2018-02-01	300381.SZ	VVUS.O
2018-02-01	002583.SZ	LKM.N	2018-02-01	300684.SZ	CTSH.O
2018-02-01	600091.SH	FRTA.O	2018-02-01	002090.SZ	PLUG.O
2018-02-01	603938.SH	FRTA.O	2018-02-01	002076.SZ	CTRN.O
2018-02-01	002478.SZ	FRTA.O	2018-02-01	002418.SZ	PVH.N
2018-02-01	600388.SH	MG.N	2018-02-01	002758.SZ	CELC.O
2018-02-01	002571.SZ	PIR.N	2018-02-01	601668.SH	IIIN.O
2018-02-01	300256.SZ	VERI.O	2018-02-01	000663.SZ	GHC.N
2018-02-01	000913.SZ	HOV.N	2018-02-01	601009.SH	HDB.N
2018-02-01	002609.SZ	GRVY.O	2018-02-01	601998.SH	CPTA.O
2018-02-01	000952.SZ	ODT.O	2018-02-01	600654.SH	PDFS.O

$$E(\ln U_{jt} - \ln C_{it}) = \frac{1}{N} \sum_{t=1}^N (\ln U_{jt} - \ln C_{it}) \quad (6)$$

如果配对因子的数值越大,说明 A 股被低估的程度越高,后续反弹的概率也越高。

2.4 构建投资组合

每次选取配对因子最高且为正值的 50 只股票,等权构建组合,如果遇到涨跌停或停牌等情况的个股,则顺延调仓。

下面展示几组具体的操作。根据模型的计算结果,2018 年 3 月 1 日的部分股票配对如表 3 所示。

表 3 2018-03-01 股票配对表

A 股代码	美股代码	配对距离	距离排名
600733.SH	HHS.N	0.0226	1
300167.SZ	CVON.O	0.0548	2
002494.SH	LAUR.O	0.0803	3
002058.SZ	CFX.N	0.0804	4
002760.SZ	MNGA.O	0.0086	5
000070.SZ	SCOR.O	0.1882	296
600284.SH	MLI.N	0.1883	297
601555.SH	JP.N	0.1884	298
600750.SH	SMMT.O	0.1884	299
002527.SZ	WLDN.O	0.1884	300

根据股票配对表,进行配对因子计算,并按照配对因子值倒序排序,具体数据如表 4 所示。

选取配对因子最大且为正的前 50 只股票,作为本月的股票组合,持有到下个调仓日 2018 年 4 月 2 日。

根据模型的计算结果,2018 年 4 月 2 日的部分股票配对如表 5 所示。

根据股票配对表,进行配对因子计算,并按照配对因子值倒序排序,具体数据如表 6 所示。

表 4 2018-03-01 配对因子表

A 股代码	配对因子	因子排名
600733.SH	0.435 027 733	1
002054.SZ	0.178 550 223	2
300233.SZ	0.173 215 068	3
000663.SZ	0.148 553 248	4
002856.SZ	0.111 473 98	5
002773.SZ	0.094 185 942	6
002760.SZ	0.093 151 378	7
300104.SZ	0.065 126 804	8
600355.SH	0.050 597 891	9
002157.SZ	0.047 763 754	10
600664.SH	-0.424 370 10	299
600829.SH	-0.439 738 54	300

表 5 2018-04-02 股票配对表

A 股代码	美股代码	配对距离	距离排名
002663.SZ	PBI.N	0.0665	1
002035.SZ	TOUR.O	0.0775	2
600308.SH	ATI.N	0.0780	3
300536.SZ	LDL.N	0.0845	4
603889.SH	DISCA.O	0.1023	5
002124.SZ	AGRO.N	0.2448	296
000933.SZ	MTX.N	0.2448	297
000727.SZ	DWCH.O	0.2449	298
300376.SZ	BGG.N	0.2453	299
002036.SZ	CLGX.N	0.2455	300

表 6 2018-04-02 配对因子表

A 股代码	配对因子	因子排名
002566.SZ	0.174 447 595	1
600634.SH	0.166 811 360	2
300626.SZ	0.064 220 226	3
600258.SH	0.058 979 964	4
002001.SZ	0.056 747 091	5
600066.SH	0.051 193 910	6
002112.SZ	0.050 522 802	7
603355.SH	0.050 156 592	8
002327.SZ	0.049 750 251	9
000663.SZ	0.048 958 786	10
002387.SZ	-0.001 864 60	299
600599.SH	-0.004 547 80	300

选取配对因子最大且为正值的前 50 只股票,作为本月的股票组合,持有到下个调仓日。

2.5 因子有效性检测

对因子有效性的检验基于两个维度:① 计算在同一时刻下,个股的配对因子值和未来一段时间收益的相关性,即量化研究中常用的信息系数(Information Coefficient, IC)。② 按照配对因子值的大小对股票进行分组,从时间序列的角度观察各分组的累计收益、信息比率以及胜率等。

信息系数的经济意义,是用来刻画因子预测股票收益的能力。信息系数的绝对值越大,表示因子预测股票收益的能力越强。信息系数通常分为 Normal IC——皮尔逊相关系数,以及 Rank IC——斯皮尔曼相关系数,其中最常用的是 Rank IC,即某时刻某因子在全部股票因子暴露值排名与其下期回报排名的截面相关系数,计算公式为

$$\text{Rank IC} = \frac{\text{cov}(\text{order}_{t-1}^f, \text{order}_t^r)}{\sqrt{D(\text{order}_{t-1}^f)D(\text{order}_t^r)}} \quad (7)$$

式中: order_{t-1}^f 为 $t-1$ 期各股票的因子排名值; order_t^r 为 t 期各股票的收益率排名值。

首先计算配对因子的因子 IC,如图 1 所示。

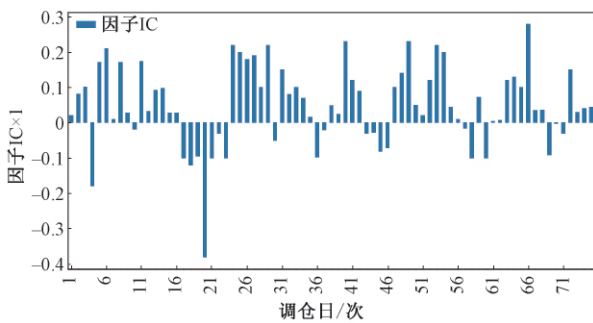


图 1 各期因子 IC 值

一般而言,因子 IC 绝对值大于 0.05,就认为是有效的因子。在统计的调仓日内,配对因子的有效率达到了 72.2%,并且大部分情况下是正相关。

然后,根据配对因子的大小对股票进行分组,选取因子值最大的 50 只股票作为低估组合,因子值最低的 50 只股票作为高估组合,计算区间累计收益率,并选取万得全 A 指数作为基准(见表 7)。模型表现如图 2 所示。

表 7 风险收益指标对比

组合	区间累计收益率/%	最大回撤/%
低估组合	-24.82	-31.34
高估组合	-35.36	-39.62
万得全 A	-30.12	-31.35

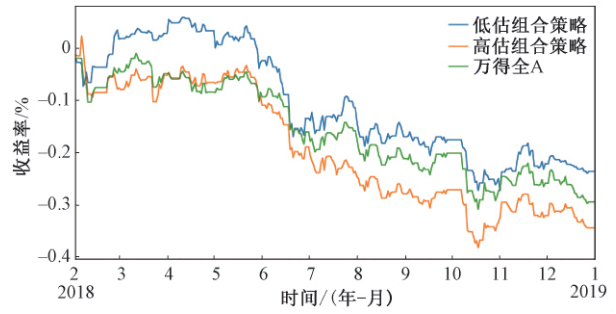


图 2 区间累计收益率

由图 2 可以看出,低估组合(配对因子选股组合)的模型表现远好于高估组合,区间累计收益率更高,最大回撤更小。低估组合也跑赢了万得全 A 指数。这说明了配对因子对 A 股股票有较好的区分能力。

2.6 稳健性检验

本文在 2016 和 2017 年用同样的策略进行投资组合构建,并统计低估组合及万得全 A 在 2016~2017 年的半年以及全年累计收益情况(见图 3~8)。

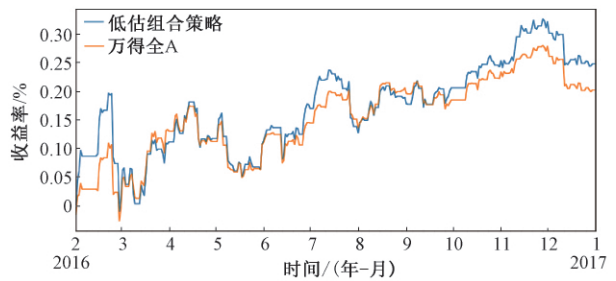


图 3 2016 年全年累计收益率

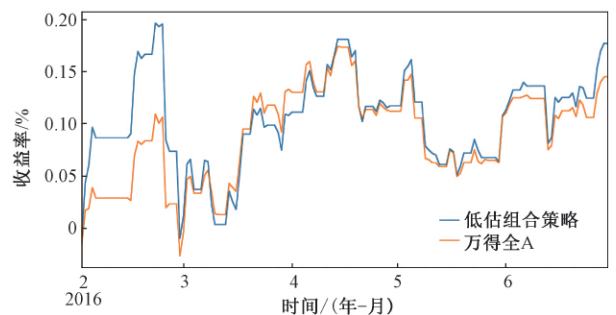


图 4 2016 年上半年累计收益率

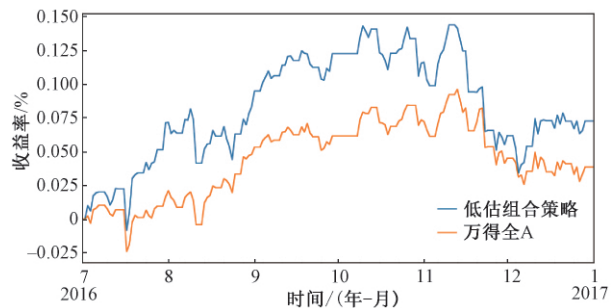


图 5 2016 年下半年累计收益率

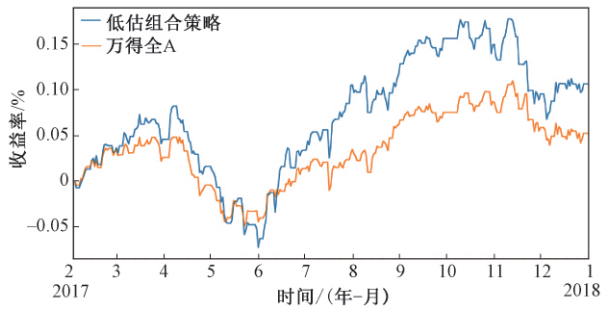


图 6 2017 年全年累计收益率

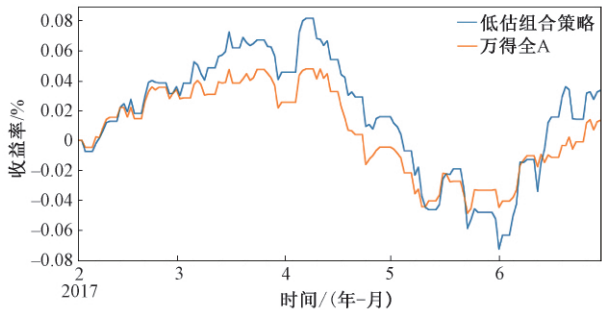


图 7 2017 年上半年累计收益率

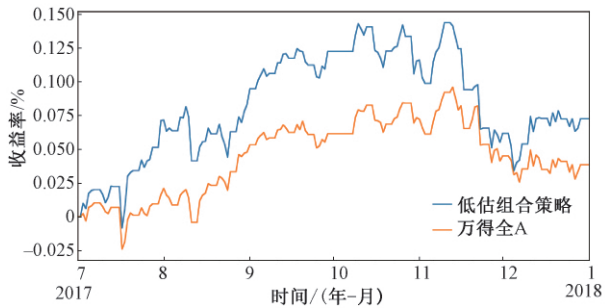


图 8 2017 年下半年累计收益率

由模型表现来看, 低估组合(配对因子选股组合)在 2016~2017 年的全年表现和半年表现都好于万得全 A 指数。具体的超额收益如表 8 所示。

表 8 低估组合超额收益

持有区间	低估组合策略/%	万得全 A/%	超额收益/%
2016 全年	24.78	20.22	4.56
2016 上半年	17.67	14.52	3.15
2016 下半年	7.11	5.7	1.41
2017 全年	10.61	5.20	5.41
2017 上半年	3.36	1.36	2.00
2017 下半年	7.25	3.84	3.41

2.7 与其他因子相关性检验

在多因子选股模型中, 因子间的相关性检验是很关键的一步。因为不同的选股因子间, 可能由于内在的驱动因素大致相同, 而导致所构建的投资组合和收益情况有较高的一致性。多重共线性会造成多元线性回归的结果偏差, 并且会使得投资组合将在同种因子上暴露过多的风险。因此需要检验配对

因子与常用的 Barra China Equity Model (CNE5) 的 10 个因子之间的相关性(见表 9)。

表 9 CNE5 因子说明

因子名称	
Beta, 贝塔因子	Residual Volatility, 残差波动率因子
Momentum, 动量因子	Growth, 增长因子
Size, 规模因子	Book-to-Price, 市净率因子
Non-linear Size, 非线性规模因子	Leverage, 杠杆因子
Earnings Yield, 收益因子	Liquidity, 流动性因子

相关性检验的具体步骤:

- (1) 对不同因子下的 n 个组合进行打分, 分值与该组合在整个模型形成期的收益相关, 收益越大, 分值越高。
- (2) 计算个股的不同因子得分间的相关性矩阵。
- (3) 在计算因子得分相关性矩阵后, 计算整个样本期内相关性矩阵的平均值。

计算结果如图 9 所示。

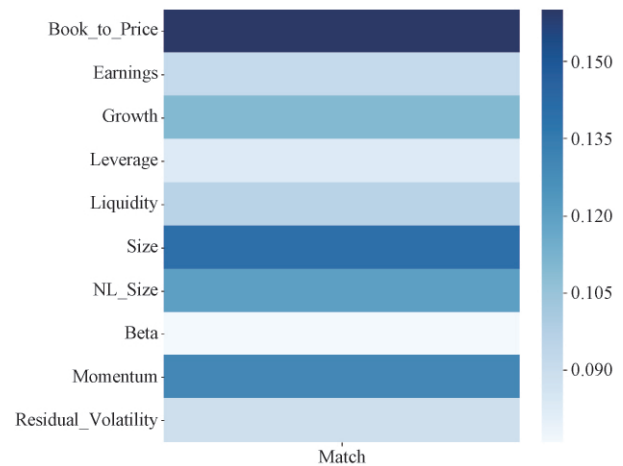


图 9 相关性分析

根据分析结果, 配对因子与 10 个因子的相关系数都低于 0.25, 一致性较低, 可以加入到传统选股因子组合中。并且, 在加入配对因子后, 新的多因子模型比之前的传统模型, 无论是线性还是非线性模型, 效果均有一定的提升。以线性回归模型为例, 加入配对因子后, 模型年化收益率提高了 1.06%。

3 结论

本文通过最小化偏差平方和方法, 在同行业内将美股和 A 股进行配对, 并利用股价对数价差计算美股-A 股配对因子, 该因子反映了 A 股股票被低估的程度。配对因子的数值越大, 说明 A 股被低估的程度越高, 后续反弹的概率也越高。每个调仓日选取配对因子最高的 50 只股票, 等权构建投资组合。通过实证分析, 配对因子对 A 股股票有较好

的区分度。在 2016~2018 年期间, 低估组合(配对因子选股组合)相比万得全 A 指数, 分别取得了 4.56%、5.41% 和 5.30% 的超额收益。同时, 通过与其他因子的相关性检验, 发现配对因子与传统的选股因子存在较低的相关性。

对配对因子的构建和研究, 在选股因子同质化日益严重的今天, 为多因子选股开拓了新思路。利用配对因子进行选股, 既集成了配对交易和因子选股的优点, 又考虑了美国股市对 A 股市场的影响, 通过实证分析发现, 确实可以获取稳健可观的超额收益。此外, 由于市场监管和交易规则限制, 学术界提出的很多配对交易模型在 A 股市场无法发挥作用。而通过挖掘 A 股和美股内在联系, 将配对交易思想应用到 A 股市场中, 对配对交易的研究进行了补充。

通过挖掘不同市场股票间的内在联系来指导投资, 还有很多可以继续探索的方向。首先, 利用中美股市的关联性研究了配对因子在选股中的有效性, 而中欧股市是否存在类似的关联效应还未进行深入研究。希望未来在获得足够数据的情况下, 对中欧股市的算法交易策略进行深入研究。其次, 基于配对因子的统计套利的策略具有较高的波动性, 可以考虑在后续引入亚式期权, 通过增强对过去价格的依赖性来降低投资组合的波动性^[26]。

参考文献:

- [1] 刘继明. 中美股市的联动性研究[J]. 北方金融, 2018(8): 26-31.
- [2] 崔方达, 吴亮. 配对交易的投资策略[J]. 统计与决策, 2011, 27(23): 156-159.
- [3] 张金萍, 王准. 中美股市联动性实证研究[J]. 商业时代, 2014(16): 74-76.
- [4] 莫悠, 程锐. 基于向量自回归模型的中美股市联动性分析[J]. 中国集体经济, 2017(25): 40-42.
- [5] 刘继明. 中美股市的联动性研究[J]. 北方金融, 2018(8): 26-31.
- [6] 唐勇, 朱鹏飞, 王雅梅. 中美股市联动效应研究——基于小波视角[J]. 北京化工大学学报(社会科学版), 2018(1): 7-16.
- [7] 李合龙, 刘方舟. 基于 EEMD 的中美股市联动性分析[J]. 广西大学学报(哲学社会科学版), 2019, 41(1): 80-86.
- [8] 赵若瑜. 我国股票市场指数与国际股票市场主要指数的联动性研究——基于协整分析[J]. 时代金融, 2017, 24(14): 156-157.
- [9] 麦永冠, 王苏生. WM-FTBD 配对交易建仓改进策略及沪深港实证检验[J]. 管理评论, 2014, 26(1): 30-

- 40.
- [10] 黄晓薇, 余湄, 皮道羿. 基于 O-U 过程的配对交易与市场效率研究[J]. 管理评论, 2015, 27(1): 3-11.
- [11] 胡伦超, 余乐安, 汤铃. 融资融券背景下证券配对交易策略研究——基于协整和距离的两阶段方法[J]. 中国管理科学, 2016, 24(4): 1-9.
- [12] 刘永辉, 张帝. 基于协整-OU 过程的配对交易策略研究[J]. 管理评论, 2017, 29(9): 28-36.
- [13] 胡文伟, 胡建强, 李湛, 等. 基于强化学习算法的自适应配对交易模型[J]. 管理科学, 2017, 30(2): 148-160.
- [14] 陈晓芬, 杨朝军. 配对交易中配对与交易策略双重改进下的行业实证检验[J]. 上海金融, 2017(2): 67-74.
- [15] 毕秀春, 刘博, 袁吕宁, 等. 带止损条件的配对交易最优阈值[J]. 系统科学与数学, 2019, 39(7): 1117-1141.
- [16] 赵胜民, 闫红蕾. A 股市场统计套利风险实证分析[J]. 管理科学, 2015, 28(5): 93-105.
- [17] 于孝建, 邹倩倩. 基于 OU 过程的商品期货市场配对交易策略[J]. 南方金融, 2018(3): 52-60.
- [18] 徐杰, 周志中. 中美股市跨市场配对交易实证分析[J]. 上海金融, 2019(8): 26-30.
- [19] PIOTROSKI J D. Value investing: The use of historical financial statement information to separate winners from losers [J]. Journal of Accounting Research, 2000, 38(1): 1-41.
- [20] MOHANRAM P S. Separating winners from losers among low book-to-market stocks using financial statement analysis [J]. Review of Accounting Studies, 2005, 10(2): 133-170.
- [21] FAMA E F, FRENCH K R. Permanent and temporary components of stock prices[J]. Journal of Political Economy, 1988, 96(2): 246-273.
- [22] PORTERBA J, SUMMERS L. Mean reversion in stock prices: Evidence and implications[J]. Journal of Financial Economics, 1988, 22(1): 27-59.
- [23] GATEV E, GOETZMANN W N, ROUWENHORST K G. Pairs trading: Performance of a relative-value arbitrage rule[J]. Review of Financial Studies, 2006, 19(3): 797-827.
- [24] 周虎群, 李育林. 国际金融危机下人民币汇率与股价联动关系研究[J]. 国际金融研究, 2010(8): 69-76.
- [25] 张兵, 封思贤, 李心丹, 等. 汇率与股价变动关系: 基于汇改后数据的实证研究[J]. 经济研究, 2008(9): 70-81.
- [26] 姚远, 李光举. 基于几何平均亚式期权的投资组合保险策略[J]. 系统管理学报, 2018, 27(3): 529-537.