

# 基于基本面分析与多因子建模分析的量化投资策略

李佳妍 胡旖雯 冯洁宜 刘烙轩

(广东理工学院 广东 肇庆 526070)

**【摘要】**本文基于近年的国内标的资产数据指标以及因子分析模型,分析并设计出统计套利之配对量化分析金融交易策略。分析并得出,新能源行业未来发展趋势较好,存在套利机会。结合多因子模型建模分析并根据因子得分综测得出沪深300中该行业10支龙头股票。经过此次统计套利之配对量化分析金融交易策略,实现量化分析策略最终的目的。

**【关键词】**基本面分析;量化投资策略;多因子模型;统计套利

## 一、引言

中国股票市场正在逐步完善,如何在市场上众多上市公司中寻找投资机会,提高投资收益、降低投资风险成为广大投资者日益关注的问题。目前市场上主要的投资方法可分为3种:技术分析、基本面分析,量化投资。技术分析以股票价格作为主要研究对象,从历史K线图入手,对价格波动规律进行分析总结基本面分析认为证券的价格是由其内在价值所决定,短期内受外界干扰会导致价格频繁波动,但长期会趋于内在价值;量化投资对大量的历史数据进行分析,运用现代信息技术搭建量化选股模型。运用量化投资的手段,将基本面分析和技术分析相结合,试图构建出一个能持续战胜市场的量化选股模型。<sup>[1]</sup>

选股思路如下:首先进行基本面因素分析,即试图通过对与股票价格存在逻辑联系的各种因素的分析,探索股票价格决定及其变动的内在原因,并在此基础上对所选股票价格的走势进行判断。其次是评估证券的内在价值,通过对多因子选股进行建模,选出显著性相关的因子,其作用在于为判断股票市场价格的高低确立一个参照标准,组合所选股票进行套利投资。

## 二、选股思路

### (一) 运用量化投资手段

通过对大量的历史数据进行分析,运用现代信息技术搭建量化选股模型。运用量化投资手段,将基本面分析和技术分析相结合,构建出能持续战胜市场的量化选股模型。

### (二) 运用基本面分析<sup>[4]</sup>

#### 1. 内在价值分析

通过价值型、成长型指标来分析股票市场价格的高低,确立选取股票的参照指标。

#### 2. 因素分析

国家“十四五”规划分析。关于基础材料及新能源行业纲要中提出的新能源是我国能源转型的重要力量加上两会中提出的碳中和成为市场焦点,更主要的是新能源企业自身展现出来的发展前景一片光明。新能源行业的国际合作态势非常明显,尤其是“一带一路”的倡议提出以后,二三线城市国家对基础设施需求量加大,在此基础上,我国的新能源技术将得到进一步的发展。新能源不仅具有良好的经济前景,而且,随其产业化发展,将提供越来越多的就业机会。对经济发展的潜在影响和作用巨大的。加之国家发展政策对新能源的倾斜以及逐步成熟的新能源技术将进一步推动我国新能源产业化发展。

依据新能源行业、市场发展情况。在市场环境、政策环境以及国际气候环境驱动下,我国新能源产业进入高速发展阶段。常规能源向新能源转变、电力企业向新能源企业融合,即将成为新的发展趋势。这一阶段上,形成了支持新能源快速发展的政策体系;新能源装备制造能力位居世界前列,关键技术取得了突破。

随着时间轴的递增,新能源企业日益增加。新能源不仅

拥有良好的经济前景,而且,随其产业化发展,将提供越来越多的就业机会。对经济发展的潜在影响和作用巨大的。加之国家发展政策对新能源的倾斜以及逐步成熟的新能源技术将进一步推动我国新能源产业化发展。

经过多年发展,我国的新能源产业资产规模逐年增加,新能源领域科技创新能力、关键技术和技术装备自主化水平显著提升。在国家的推动下,行业加速发展,2020年新能源行业的营收规模已超过1.1万亿元。经统计,中石化、国家能源集团、国家电投、华能集团、中核集团、三峡集团、国家电网等多家央企均发布了专项用于碳中和或新能源项目的债券,合计发行规模高达182亿元。随着经济发展,市场化改革进程加快,能源利用的市场化逐渐成为必然。

### (三) 运用多因子量化选股策略

其核心思想是基于数学和统计的方法,检验与股票价格相关的一系列因子的有效性,并且通过组合多个有效因子来建立一个定量模型对股票进行分析,依据相应原则选取财务指标优秀的股票进行投资组合获取超额收益。

### (四) 运用量化择时策略与回测

利用数量化的方法,通过对各种宏观微观指标的量化分析,试图找到影响大盘走势的关键信息,并对未来走势进行预测。本文基于的多因子选股模型主要应用于经初步挑选的较优的96支沪深300股,根据投资标的相关特性,筛选出可能有效的因子。本文所建立的待测因子库包括了成长型因子和价值型因子。经过对单因子检测后的有效因子,建立多因子模型并且在聚宽平台上进行回测。<sup>[2]</sup>

综合分析,新能源行业得到极大的关注,这为我们构建股票组合投资提供了利好的消息,我们选择了新能源行业作为我们选股策略的研究行业,并根据基本面分析以及相关的多因子策略分析选出了其中的10支比较有代表性的股票来进行模拟套利策略的投资。根据相关的选股指标,例如:夏普比率,贝塔系数,ROE,ROA,PE,PB等等指标对整个新能源行业的股票进行基本面分析的一种行业筛选,最后在总计96支股票的新能源行业中选取10支股票,进行根据价值型因子与成长型因子的技术分析,进行多因子分析的建模然后再进行多因子(GARP)模型的打分并且对这10支股票进行一个排名与分析。

综上所述,本文选取新能源行业的股票作为统计套利操作的标的资产。通过对选股行业的确定,我们在聚宽平台筛选出以下十家基本指标良好的公司企业来作为备选能够组合价值投资的股票。分别是:南网能源、新天绿能、比亚迪、动力源、中材节能、美锦能源、银星能源、江苏新能、宝新能源、豫能控股。

## 三、指标选取

### (一) 候选因子

本节基于新能源上市公司财务信息构建多因子打分模型,对影响股价的因素展开全面分析,目的是选出10只长期优势股。并且根据各个公司指标数据,运用GARP选股策

略挑选出十支具有高投资价值的股票。为了使投资研究更贴近现实情况,本文提取的财务数据均为该公司对应的季度报表信息。多因子打分模型的候选因子包括4大类16个因子,本文选取了7个因子进行分析,涵盖估值、盈利、质量、营运多个维度。

### (二) GARP 策略指标选取及构建模型

GARP 选股策略是一种结合了价值投资和成长投资策略,试图寻找价值被市场低估,同时又有较强持续稳定增长潜力的股票,具有分散化投资、数量化选股以及收益较稳定的点。GARP 选股策略将指标分为价值型因子和成长型因子,综合考虑指标的意义各方面因素本文选择的价值型因子包含5个子项目,成长型因子包含6个子项目。这些指标可以综合反映上市公司的价值和成长状况,保证了通过 GARP 策略选出的股票具有好的统计套利的特性。根据该模型,选股时同时考虑价值型因子(市净率和市现率)和成长型因子(净利润、ROE、ROA、净利润率、利润同比增长率、收入同比增长率)<sup>[3]</sup>。

本文将选择股票的基本面分析与量化手段结合,通过反映一只股票价格波动和内在价值指标进行相似度的分类,用显著性靠前的相关因子进行系统打分,选取出得分靠前的前十名的新能源公司股票。

### (三) 新能源候选因子<sup>[1]</sup>

不同行业的公司针对同一财务指标的大小相差甚远,比如信息行业上市公司的市盈率普遍高于材料行业。此外,每个行业的行业特征不同,比如对消费品行业,存货周转天数很好地衡量了企业销售消费品的效率,但是对于金融企业,存货周转天数不再是合适的衡量指标。通过运用分行业因子测试结果统计得出,新能源行业中因子 IC 绝对值最大的五个有效因子为 NP\_GRT 扣非归母净利增长率、PC 市现率、ROE 净资产收益率、PB 市净率、ROA 总资产收益率。

另外再运用公共指标因子 SGP\_MARGIN 毛利率、ASS\_LIAB 资产负债比、PE 市盈率等进行多因子打分。此外并对量化策略的年化收益率、夏普比率、最大回撤率、收益波动率、贝塔系数等主要指标进行综合评比。

在聚宽平台上获得新能源行业100多家企业的所需要的指标,包括了GARP所涉及的价值型因子和成长型因子,对每一项指标数据进行标准化处理,选出了新能源行业中十支有价值的股票。

## 四、实证分析

### (一) KMO 以及 Bartlett

构建模型前,需对所选指标进行 KMO 样本测度和 Bartlett 球体检验,看样本间是否具有相关性,是否适合进行因子分析。KMO 值越大大于 0.5,说明变量间的相关性越强,越适合做因子分析。检验结果显示,KMO 值为 0.570, Bartlett 球形检验值为 23.814, P 值为 0.069,经过检验得出,所选指标适合作为因子分析。

### (二) R 特征值以及贡献率

将各项指标的原始数据标准化后建立变量的相关系数矩阵 R 的特征值及贡献率。结果显示,可提取 3 个因子,变量相关系数矩阵有特征值分别为:2.547、1.457、1.191,基于过程内特征根大于 1 的原则主因子分析过程相应提取表中成分前三个作为主成分量,并称为 F1、F2、F3。并且这三个因子共解释 X 标准方差的 64.944% (累计贡献率)。即前 3 个成分反映了原始数据所提供的足够信息,满足研究的绝大部分要求。

### (三) 因子旋转

对于上述所提取的 3 个主因子分量 F1、F2、F3,建立起原始因子载荷矩阵。为了使得所提取的指标对于标的资产设计有一定的实际代表意义,我们选取了利用方差最大正交

旋转法来处理因子载荷矩阵。由此可得,F1 在市盈率、市净率以及市现率上有比较大的载荷,这些指标通常是衡量企业盈利能力和成长能力的重要指标,能呈现公司的基本情况、潜力与否,也能反映公司的估值高低,根据上述所选用的 GARP 模型,在此称 F1 为价值绩效因子(估值因子)。F2 在换手率、夏普比率和资产负债率上有较大的载荷,这几个指标是衡量公司风险控制能力的重要指标,同时也在一定程度上可以反映企业公司偿债能力,因此在此称 F2 为风险性绩效因子(营运因子)。F3 则是在市盈率、净资产收益率等有较大的载荷,主要衡量了公司企业的盈利能力,称 F3 为盈利绩效因子(质量因子)。在这三个主因子中,F1、F2、F3 因子的方差分别是 31.84%、18.22%、14.89%,即对于新能源企业公司影响其绩效最大的因素是公司的价值估值,其次是偿债能力和盈利能力。

综上所述,以各因子的信息贡献率作为加权数计算新能源企业公司的综合测评得分,根据主因子各贡献率/累计贡献率,得出公式如下:

$$Z=0.4902Y1+0.2804Y2+0.2292Y3.$$

其中,Z 为各新能源公司财务绩效状况得分,Y 的系数则为各因子的信息贡献率,是各因子的方差贡献率与三个主成分的累计贡献率(31.84%+18.22%+14.89%=64.95%)的比值。根据公式得出,Y1, Y2, Y3 的系数分别为 0.4902, 0.2804, 0.2292。因此,价值绩效因子的信息贡献率是 0.4902,风险性绩效因子的信息贡献率是 0.2804,盈利绩效因子的信息贡献率是 0.2292。为了避免偶然性,我们增加了样本数量对因子综合进行评分,选取了 17 家上市新能源企业公司,通过计算得出了以下 17 家公司绩效状况的因子综合测评评分以及前十个得分高的综合排名:

003035 南网能源 1.41, 600956 新天绿能 1.11, 002594 比亚迪 0.67, 601997 动力源 0.29,

603126 中节材能 0.24, 000723 美锦能源 0.09, 000862 银星能源 0.06, 603693 江苏新能 -0.08, 000690 宝新能源 -0.15, 001896 豫能控股 -0.19。

通过上述分析,采用多因子分析法对新能源股票进行分析,由于选取指标具有普遍代表性,预测效果有限。因此本文将多因子分析与基本面分析相结合,本文选股策略通过基于基本面分析的量化策略进行选股,并且模拟盘收益率,最终达到了统计套利的目的。

### 五、动态模拟跟踪

采取上述的选股策略,在东方财富平台上进行模拟选股组合投资,并根据价差趋势的变动判断其市场短期走势,合理地确定开平仓的时机,顺势交易。根据量化策略选择的 10 支股票分散投资,在近半个月来,新天绿能、比亚迪等股票曾出现大幅涨停,总体股票投资收益率达到 9% 以上。

夏普比率大于 1,表明这个策略组合承担单位风险可以获得的超额风险高,绩效表现比较好。贝塔系数小于 1,风险度可以得到控制。胜率 0.65,可以获利的概率比较高。年化收益率 36%,可以获得可观的收益。

### 参考文献:

- [1] 吕凯晨, 闫宏飞, 陈翀. 基于沪深 300 成分股的量化投资策略研究 [J]. 广西师范大学学报, 2019, 37 (01): 1-12.
- [2] 王春丽, 刘光, 王齐. 多因子量化选股模型与择时策略 [J]. 东北财经大学学报, 2018 (05): 81-87.
- [3] 高兴波, 郭甲蕾, 胡智磊. 统计套利策略在对冲基金投资中的应用研究 [J]. 中央财经大学学报, 2014 (06): 31-37.
- [4] 侯晓辉, 王博. 基于基本面分析的量化投资: 研究述评与展望 [J]. 东北师大学报 (哲学社会科学版), 2021 (01): 124-131+141.