

浅谈统计套利在我国金融市场的应用

◆贺蓝卿

(长沙理工大学经济与管理学院 湖南长沙 410076)

摘要:统计套利策略是一种基于数理模型的量化投资策略,被对冲基金等金融机构广泛使用。我国的融资融券的启动和股指期货上市,统计套利开始在我国有了发展的空间和平台。本文从统计套利的定义和原理、策略种类以及在我国金融市场的应用基础等方面做了简要的分析。

关键词:统计套利;协整理论;程序化交易

一、统计套利的定义与原理

统计套利模型是依据两个或者一组股票(或其他证券)之间所具有的较高的相关性进行套利的一种策略模型。在用一定的方法验证这两种资产的价格波动在一段时间内保持这种良好的相关性之后,那么两者的价差一旦出现了背离的走势,并且这种价格的背离在未来预计会得到恢复,那么便可以产生套利机会。在统计套利实践中,当两个具有此类相关性的资产的价格之间出现背离,那么可以买进价格被低估的证券,同时卖出价格被高估的证券,那么当未来两者之间的价格背离得到纠正时,进行相反的平仓操作,就可获得套利的利润。统计套利原理要得到实现,必须要以均值回复为前提。均值回复指的是资产的价差存在一个均值区间,通常表现为资产价格之差在时间序列上具有平稳性,如果价格差背离至均值区间之外,那这种背离只是短暂的,随着时间的推移,资产价格将会恢复到它的均值区间内。例如,处于同一行业的两只股票,它们在某种程度上具有和许多共同性,其市场价格之间也存在着较为稳定的相关性,一只股票价格的变化通常伴随另一只股票价格的同方向变化,从而价格之差往往会围绕着某一固定值上下波动。因此,如果经检验发现某资产的价格差时间序列是平稳的,便可以构造统计套利交易的信号发现机制,该信号机制将会显示资产价格是否已经偏离了长期均值从而寻找套利的机会。

二、统计套利策略

统计套利策略如今被对冲基金广泛使用,其具体的操作方法策略种类繁多,如成对交易、多因素模型、均值回归策略、协整策略等。另外在计算机科学发展基础上,越来越多的资产价格预测方法都开始运用模式识别、系统理论、数据挖掘、机器学习等理论。虽然这些理论基础在金融市场的应用还处于试探的起步阶段,但是随着理论的不完善和在其他领域应用拓展,可以想象今后的资产交易模式会广泛使用这些理论,而它们比传统的统计套利方法更加高效。

(一)均值回复策略

均值回复策略是股票市场的一种典型的投资策略,它与趋势追踪策略相反,认为股价经过大幅度的上涨后便比较可能出现下跌的情况,反之股价大幅度下降后很可能会上升。这种策略特点是尽可能多地捕捉小的机会,虽然每一次收益率不高,但是成功地可能性较高。但我国股市发展还处于不成熟的阶段,投资者行为不理性,导致我国股票价格趋势性比较明显,而均值回复并不经常发生,所以这种策略在我国股市上并不容易实现。但是,统计套利策略基于的均值回复策略不单单是利用单个资产的价格的变化,更多地是基于两个或者多个具有较高相关性的资产之间的均值回复现象。即两个或多个相关资产之间的超额收益率如果存在着均值回复现象,那么在一段时间内,当某一资产与另一资产相比具有明显的超额收益时,则在将来的某一段时间内,另一资产的超额收益率往往会超过该资产。

(二)多因子策略

多因子策略的思想是股票的收益与多种选择因素相关,这类方法以套利定价模型为代表。多因子策略首先要选择和定义影响股票收益的因素,用股票收益对这些因素进行多元回归,再基于多元回归结果的相关性选择和确定投资组合。多因子模型通过寻找引起资产价格变动的共同因素,再构建基于这些因素的多元模型,简化了投资组合策略中需输入的资产间的相关系数的要求。在实践操作中,多因子模型效果的优劣主要是取决于影响因子的选择。

(三)协整策略

协整套利策略是所有统计套利方法中理论基础最坚实,应用最广泛的一种,同时也是实证证明效果非常好的一种策略。在传

统的金融建模理论中,需要保证金融数据具有动态稳定性,也就是说,在分析时间序列数据时要假设时间序列具有平稳性。即时间序列数据的统计特性不会随时间变化而变化,不会因时间节点的选取而变化。但实际情况是,金融时间序列通常不具有平稳性,如果使用这种不平稳时间数据建立一个稳定的模型,会出现“伪回归”的问题。协整理论的出现是这种不稳定的问题得到了解决。协整理论描述的是经济变量之间存在的一种均衡关系,对于两个或者多个自身不具有平稳性的经济变量来说,如果它们之间存在协整关系,也就是说它们之间存在一个长期的均衡关系,而协整套利模型可以利用这种价差的平稳性寻求套利机会。协整套利策略主要有三个步骤:第一,构建资产组合,确定资产比例;第二,建立协整模型,确定交易信号;第三,按照交易信号实施策略。

三、统计套利在我国金融市场的应用基础

(一)理论基础

韩广哲,陈守东(2007)对上证50指数成分股进行统计套利研究,运用方差比分析法发现统计套利策略在样本外的绩效对交易成本非常敏感,计算出针对机构投资者的样本外数据年化夏普比率符合市场中的投资要求,说明统计套利策略是可行的。颜涵(2012)分别运用OLS恒定波动模型、GARCH时变波动模型和O-U随机波动模型对沪深300股指期货的5分钟高频数据进行了统计套利分析。何树红,张月秋等(2013)结合了协整模型和GARCH模型,发现样本数据从结果上存在套利获利的机会。张金雷(2014)使用EGARCH模型进行统计套利的构建,并使用时变方差作为交易阈值的开平仓操作方式,结果表明使用时变方差进行交易阈值的预测比以往使用常数方差更为精准和高效。李乐,张淳奕等(2014)使用了从0.5秒到30分钟的一系列数据对跨期套利进行对比研究分析,实证结果发现,使用1分钟高频数据取得的跨期套利效果最佳。

总体来说,我国金融市场对统计套利的应用尚处于起步阶段,协整模型是使用最广泛的统计套利模型。随着高频数据运用的兴起,计算机科学的发展,投资者越发地开始关注日内高频数据的使用和数据挖掘、机器学习等理论的应用。

(二)市场应用基础

1.做空机制为统计套利提供市场基础

2010年03月30日,上交所、深交所分别发布公告,表示将于2010年3月31日起正式开通融资融券交易系统,融资融券业务正式启动;4月16日,沪深300股指期货在我国正式上市交易,标志着我国资本市场改革发展又迈出了一大步,也标志我国资本市场单边市就此终结。这对于发展和完善我国资本市场体系具有重要而深远的意义。融资融券和股指期货推出前后,各大券商纷纷对融资融券和股指期货加大研究力度,推出了一系列高质量的研究报告,其中出现最多的当属统计套利,中信、海通、华泰联合、光大、招商、齐鲁、浙商等券商都曾出过关于统计套利的研究报告,并据此构建各自的统计套利策略组合。可见,随着“做空时代”的到来,统计套利策略在量化投资策略中占据着越来越重要的地位。

2.程序化交易发展现状

我国程序化交易系统发展得相对较晚,无论是在交易系统研究还是系统应用方面的成果都较为缺乏。我国的程序化交易仍处于起步阶段,在期货市场上,主要是采用短线和超短线的趋势交易策略对商品期货进行投机,证券市场上,则是主流券商运用程序化交易对交易所交易基金进行套利。不过,由于我国的信息技术应用起步较高,近几年随着市场化程度的提高,新的投资理念被源源不断地引进,各交易所也适时推出了套利交易指令,各种各样的程序化交易系统被开发出来,尽管在市面上还看不到商业化的交易系统推出,但底层的程序化交易开发平台、应用平台等均已较为完善。程序化交易发展为统计套利提供了应用的平台,相信将来统计套利会在我国广泛应用和发展。

作者简介:贺蓝卿,女,1993.06.20,湖南怀化,汉族,职称:无,长沙理工大学硕士研究生,研究方向:统计学。