

# 沪深 300ETF 期现套利机会差异性实证分析

丁挺立

(江西财经大学 信息管理学院 江西 南昌 330013)

**摘要:**文章分析了持有成本套利方法与协整套利方法的优劣性,并基于沪深 300ETF 对两期现套利方法做了实证分析。结果发现:协整套利方法比持有成本套利法的适用范围更广,更能发掘套利机会。这与现有的基于仿真数据和其它 ETF 数据所得的结论一致。同时,基于协整套利方法发现,华泰柏瑞沪深 300ETF 的套利机会较小,而嘉实沪深 300ETF 的套利机会较大。

**关键词:**沪深 300ETF; 套利机会; 差异性; 实证分析

**中图分类号:** F830      **文献标识码:** A      **文章编号:** 2095 - 0098(2012) 05 - 0044 - 06

## 引言

2012 年 4 月我国首次推出了两只沪深 300ETF,在近一个月的募集期之后,华泰柏瑞沪深 300ETF 和嘉实沪深 300ETF 于 5 月 28 日分别在上海证券交易所和深圳证券交易所上市交易,从而改变了沪深 300 股指期货套利时缺少对应的现货工具的现状。国内外对于套利方法研究大量集中于持有成本理论,近年来也有一些学者开始了协整套利方法的研究。

陈守东、韩广哲、荆伟<sup>[1]</sup>利用协整模型,分析了股票指数间的长期平稳关系。仇中群、程希骏<sup>[2]</sup>采用股指期货仿真交易数据,比较研究了持有成本套利与协整套利的效果,然而由于数据源于仿真交易,实证结果并不理想。葛翔宇、吴洋、周艳丽<sup>[3]</sup>主要从理论上探讨了门限协整套利的方法,并对英国富时 100 股指期货与德国法兰克福 DAX 股指期货的协整关系做了验证,提出了跨市场套利的设想。黄晓坤、侯金鸣<sup>[4]</sup>研究了 ETF 类基金的持有成本套利法,并发现套利收益率服从广义极值分布。刘伟、陈敏、梁斌<sup>[5]</sup>采用一分钟高频数据,基于持有成本理论研究了华夏上证 50ETF 与华安上证 180ETF 的一、二级市场之间的套利交易。Burgess, N.<sup>[6]</sup>对 FTSE 100 指数进行套利时,就是采用了协整统计模型,并且取得了良好的效果。Board J 与 Sutcliffe C.<sup>[7]</sup>利用协整方法对 Singapore, Chicago and Osaka 的 Nikkei 225 指数合约进行了价差套利研究,发掘出套利交易的空间。Park, T. H. and Switzer, L. N.<sup>[8]</sup>采用日数据,应用 GARCH(1,1) 模型研究了 TIPS 推出之后,市场交易量的改变状况。Engle<sup>[9]</sup>第一次使用 ETF 每笔交易的高频数据,研究了 ETF 二级市场交易价格和一级市场参考净值的溢价与折价问题,得到结论:以美国国内指数为标的的基金跟踪误差要比以国外指数为标的在美国发行的 ETF 的跟踪误差小得多。Switzer L. N., Varson P. L. and Zghidi S.<sup>[10]</sup>采用小时数据,研究了股指期货与交易型开放指数基金之间的关系,发现股指期货有显著的正价格误差,而 ETF 的上市可以改善股指期货的定价效率。

传统的套利研究大部分是基于持有成本定价理论的无套利价差空间的分析,近年来,虽然有一些文章研究协整统计套利的方法。但是由于国内股指期货与相应的 ETF 刚刚推出,导致国内市场的研究数据几乎都来自仿真交易或组合 ETF。本文于沪深 300ETF 上市之际,采用一分钟高频数据,试图基于上述两种方法对分别在上交所和深交所上市的两只沪深 300ETF 期现套利机会的差异性做出实证分析。

收稿日期:2012 - 08 - 16

作者简介:丁挺立(1986 -),男,安徽黄山市人,硕士研究生,研究方向:金融计量与投资分析。

### 一、持有成本与协整期现套利、沪深两市300ETF要素的差异性

在国外市场上,指数现货和期货的套利策略是目前ETF研究的一个重要的方面。当股指期货与现货的价差上升至无套利区间的上界时,开仓卖出期货并同时买入相同价值的现货,持有至价差回到无套利区间后,进行平仓了结,获取套利收益,称为正向套利;反之,当股指期货与现货的价差下降到无套利区间的下界时,开仓买入期货并卖出现货,持有至价差回到无套利区间后,进行平仓了结,获取套利收益,称为反向套利。

在传统的持有成本定价理论中,股指期货理论价格为:

$$F_{t,T} = S_t e^{\frac{(r-d)(T-t)}{365}} \quad (1)$$

其中  $F_{t,T}$  为股指期货在  $t$  时刻的理论价格;  $S_t$  为  $t$  时刻现货指数的价格;  $r$  为无风险利率;  $d$  为股息收益率;  $T$  为合约到期时间;  $t$  表示当前的时间。

$$\text{套利活动的交易成本和跟踪误差标准为: } 2(F_t + S_t)C_R + S_t T_E \quad (2)$$

其中  $C_R$  为单向交易成本;  $T_E$  为跟踪误差。

当期货的实际价格与理论价格之差大于(2)式时,就可以开始正向套利。即:

$$F_t - F_{t,T} > 2(F_t + S_t)C_R + S_t T_E \quad (3)$$

定义期现价格比  $P_R$  为下式

$$P_R = \frac{F_t}{S_t} \quad (4)$$

于是,将(1)代入(3)式,再代入(4)式可得

$$P_R = \frac{F_t}{S_t} > \frac{e^{\frac{(r-d)(T-t)}{365}} + 2C_R + T_E}{1 - 2C_R} \quad (5)$$

此式表明,用期现价格比可进行正向套利是有下限的。反之,当期货的理论价格与实际价格之差大于(2)式时,就可以进行反向套利。即:

$$F_{t,T} - F_t > 2(F_t + S_t)C_R + S_t T_E \quad (6)$$

同理可解得

$$P_R = \frac{F_t}{S_t} < \frac{e^{\frac{(r-d)(T-t)}{365}} - 2C_R - T_E}{1 + 2C_R} \quad (7)$$

此式表明,用期现价格比可进行反向套利是有上限的。由(5)和(7)可得期现价格比的无套利区间:

$$\frac{e^{\frac{(r-d)(T-t)}{365}} - 2C_R - T_E}{1 + 2C_R} < P_R < \frac{e^{\frac{(r-d)(T-t)}{365}} + 2C_R + T_E}{1 - 2C_R} \quad (8)$$

当期现价格之比  $P_R$  上升超过(8)式所示区间上界时,就存在正向套利机会;当期现价格之比  $P_R$  下降超过(8)式所示区间下界时,就存在反向套利机会。

然而,这种传统的持有成本套利方法有以下两个缺陷:一是股息收益率  $d$  不易确定,尤其是在我国资本市场中,上市公司一般不配发现金红利,并且权益分配方式和数量都不确定,只有凭借经验设定股息收益率  $d$  的值,从而使得(8)式中的无套利区间不易确定。二是单向交易成本对于采用实物申赎的ETF尚且可以确定,而对于采用部分现金替代方式的ETF则无法确定。

基于协整的统计套利的方法可以避免以上所述的两个缺陷,但需要充分的市场交易数据,来拟合出期货价格与现货价格的长期均衡关系,而金融时间序列大量的数据完全满足统计分析的需要。本文试图设定两个门限值  $\pm \theta$  ( $\theta > 0$ ),一旦价差偏离均衡,高于  $+\theta$  或低于  $-\theta$  时,则认为存在套利机会。也就是说,期、现价格间的协整关系并非是线性的。在门限区域  $(-\theta, +\theta)$  内部,二者价格服从随机游走过程;而在门限区域  $(-\theta, +\theta)$  外,二者价格迅速收敛于均衡;这种非线性的协整关系被称为门限协整,而这里  $(-\theta, +\theta)$  的区间正体现了期现套利的交易成本。

现将样本分为两段,前段样本长,后段样本短。首先,用第一段数据建模。如果期货价格  $F_{t,T}$  与现货价

格  $S_t$  的对数序列都是一阶单整的,即为  $I(1)$  过程,则对两对数序列建立回归方程:

$$\ln F_{t,T} = \alpha + \beta \ln S_t + \mu_{1t} \quad (9)$$

若残差项  $\mu_{1t}$  不含单位根,则期货价格与现货价格的对数序列存在协整关系。

然后根据该模型的运算结果,用第二段数据计算残差  $\mu_{2t}$ 。

$$\mu_{2t} = \ln F_{t,T} - \beta \ln S_t - \alpha \quad (10)$$

若满足假设  $\mu_{1t}, \mu_{2t} \sim W(0, h_{it})$ , 令

$$\mu_{3t} = \frac{\mu_{2t}}{h_{1t}} \quad (11)$$

其中  $h_{1t}$  为  $\mu_{1t}$  样本方差。这样开平仓的标准设定如下:当  $|\mu_{3t}|$  超过 1 且又回落到 1 时,则开仓;若  $|\mu_{3t}|$  超过 1 且不再回落到 1,则认为两序列的协整关系可能不再存在。开仓后,当  $|\mu_{3t}|$  回落到 0 时平仓。开仓后,若  $|\mu_{3t}|$  超过 2,则认亏平仓。这样的标准可以确保价差在可控范围内,从而不像持有成本套利理论那样出现爆仓。

从理论上讲,协整统计套利的方法要比基于持有成本理论的套利方法所需的约束条件更少。两交易所的规则对两只沪深 300ETF 的约束也不相同,如表 1 所示。

表 1 两只沪深 300ETF 要素比较

基金名称	华泰柏瑞沪深 300ETF	嘉实沪深 300ETF
上市交易所	上海证券交易所	深圳证券交易所
1 个申赎单位	90 万份	200 万份
1 个申赎单位对应的合约数	3 张沪深 300 股指期货合约	约 6.67 张沪深 300 股指期货合约
申赎模式	沪市证券申赎 + 深市证券现金替代;	场外实物申赎;
流动性	T 日买入股票, T 日可申购 ETF; T 日申购 ETF, T 日可卖出; T 日买入 ETF, T 日可赎回; T 日赎回 ETF, T 日(沪) T+2 日(深) 卖出	T 日买入股票, T+1 日可申购 ETF; T 日申购 ETF, T+2 日可卖出; T 日买入 ETF, T+2 日可赎回; T 日赎回 ETF, T+2 日得股票卖出

深交所上市的沪深 300ETF 在申赎时,采用国际惯用的场外实物申赎方式,使得单向交易成本可以确定,但 T+2 的交收模式减小了其流动性,可能适合基于持有成本的套利方法。而上交所上市的沪深 300ETF 在申赎时,其深市的证券部分不可以直接用证券进行申赎,而要用现金进行替代,则投资者需要承受现金代买深市证券时的冲击成本,这使得其单向交易成本无法确定,但其采用 T 日交收模式使其流动性很高,更加适用于协整统计套利的方法。

## 二、实证分析

### (一) 数据选取与处理

根据中金所的规定,同一时点上有四种沪深 300 股指期货合约可以进行交易,分别是当月合约、下月合约和随后两个季月合约,每种合约在合约交割月份的第三个周五进行交割。因此,沪深 300 股指期货的数据是间断的。可以对数据进行连续性处理:先选取交易最为活跃的一种合约作为基准数据,通常为当月合约,再依据合约交割日进行连接,当月合约交割完毕,以下月合约进行连接。然而这样的处理会使得交割日当天的连接点上出现一个价格的跳空点。为了保证数据的连续性和可靠性,本文仅选用当月合约中的数据进行研究,即 2012 年 6 月 18 日到 7 月 20 日间共 25 个交易日。而 IF1207 合约刚开始交易不活跃,最后几个临近交割的交易日又过分活跃,因此剔除掉前 6 个交易日和后 5 个交易日的数据。考虑到套利机会的瞬时性,若采用日数据,则无法迅速捕捉短暂的套利机会。本文选用的是两只沪深 300ETF 与沪深 300 股指期货的 1 分钟高频数据。14 个交易日共有 3360 组数据,将 14 个交易日的 3360 组数据分为长、短两段,前一段为前 13 个交易日 3120 组数据,后一段为后 1 个交易日的 240 组数据。然后采用第一段数据来建立模型,而第二段

数据则根据第一段数据建好的模型来进行交易。

由于股指期货每天比ETF多交易30分钟,为了保持数据一致,剔除掉每天上午9点15分到9点30分与下午3点到3点15分的数据;最终本文采用华泰柏瑞和嘉实沪深300ETF以及IF1207合约的6月27日至7月13日共13个交易日的3120组一分钟数据,进行建模;用7月16日的200组一分钟数据进行期现套利交易。本文的数据来自于WIND咨询和同花顺行情软件。

## (二) 持有成本套利机会实证分析

先需要确定股息收益率 $d$ 和无风险利率 $r$ ,考虑到我国很少发放现金股利的现状,可设 $d$ 为0; $r$ 取一年期定期存款利率3.25%; $T_E$ 为模拟误差,由于沪深300ETF与沪深300股指期货具有相同的标的资产,可以把模拟误差视为0; $C_R$ 为单向交易成本,如表2所示。

表2 单向交易成本 $C_R$

标的指数	期货交易成本	ETF现货交易成本	总交易成本
沪深300	0.15%	0.2%	0.35%

根据(8)式 
$$\frac{e^{\frac{(r-d)(T-t)}{365}} - 2C_R - T_E}{1 + 2C_R} < P_R < \frac{e^{\frac{(r-d)(T-t)}{365}} + 2C_R + T_E}{1 - 2C_R}$$
 可以得到: 分别在两个交易所上市的沪深

300ETF的期现价格比 $P_R$ 以及无套利区间的上下界,如图1与图2所示。

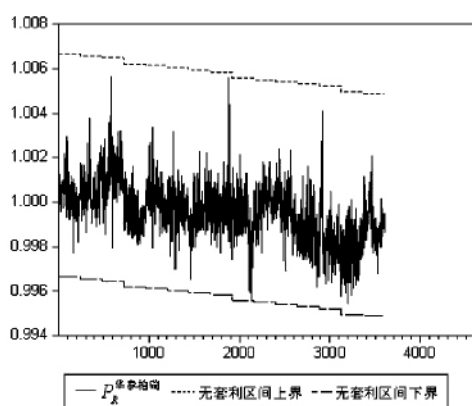


图1 华泰柏瑞沪深300ETF的 $P_R$ 与无套利区间

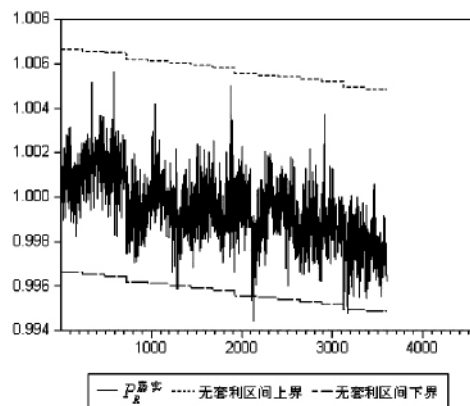


图2 嘉实沪深300ETF的 $P_R$ 与无套利区间

综上所述,若按照持有成本套利理论计算,对于上交所上市的华泰柏瑞沪深300ETF在样本数据范围内并未出现套利机会;而深交所上市的嘉实沪深300ETF则出现了5次套利机会,但是每次套利时间都很短暂,很难把握套利时机。这与上文理论分析的结论是一致的。

## (三) 协整套利机会实证分析

本文尝试采用协整统计套利的方法,选取同样的样本数据,来进行实证检验,并与传统的持有成本套利理论的结果做比较。

首先,对 $\ln F_{t,T}$ 和 $\ln S_t$ 做单位根检验,结果见表3。

表3 价格序列单位根检验

序列	t-统计量(ADF)	P值	t-统计量(1%显著性水平)	t-统计量(5%显著性水平)
$\ln F_{t,T}$	-2.122835	0.5322	-3.961854	-3.411673
$\ln S_t^{510300}$	-1.736640	0.7349	-3.960430	-3.410976
$\ln S_t^{159919}$	-2.880224	0.1691	-3.960591	-3.411055
$D(\ln F_{t,T})$	-65.75644	0.0001	-2.565565	-1.940907
$D(\ln S_t^{510300})$	-50.80640	0.0001	-2.565923	-1.940955
$D(\ln S_t^{159919})$	-49.38638	0.0001	-2.565605	-1.940912

表注:510300和159919分别为华泰柏瑞沪深300ETF和嘉实沪深300ETF的交易代码。

由表 3 可知,沪深 300 股指期货与华泰柏瑞沪深 300ETF 以及嘉实沪深 300ETF 的价格对数序列含有单位根,而三个序列的差分序列都通过 ADF 检验,是平稳的,三个序列皆为一阶单整序列  $I(1)$ 。

对  $\ln F_{t,T}$  与  $\ln S_t$  建立 (9) 式所示回归方程,并对  $\mu_{1t}$  做单位根检验,如表 4 所示。

表 4  $\mu_{1t}$  的单位根检验

序列	t-统计量(ADF)	P 值	t-统计量(1% 显著性水平)	t-统计量(5% 显著性水平)
$\mu_{1t}^{510300}$	-5.662937	0.0000	-2.565606	-1.940912
$\mu_{1t}^{159919}$	-4.267180	0.0000	-2.565608	-1.940912

由表 4 可知,  $\mu_{1t}^{510300}$  和  $\mu_{1t}^{159919}$  在 1% 显著性水平下都不含单位根。因此,不论是华泰柏瑞还是嘉实沪深 300ETF,都与沪深 300 股指期货存在协整关系。将上述模型的估计结果带入到第二段数据中,可得:

$$\mu_{2t}^{510300} = \ln F_{t,T} - 1.029408 \ln S_t^{510300} - 6.872906 \quad (12)$$

$$\mu_{2t}^{159919} = \ln F_{t,T} - 1.034620 \ln S_t^{159919} - 7.879525 \quad (13)$$

综合 (11) (12) (13) 式可得  $\mu_{3t}$  的序列分布图如图 3 与图 4 所示:

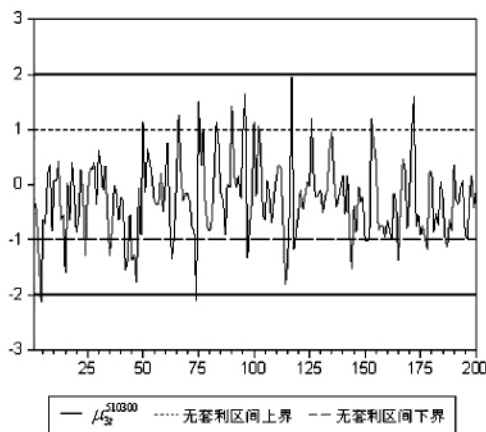


图 3  $\mu_{3t}^{510300}$  序列及无套利区间

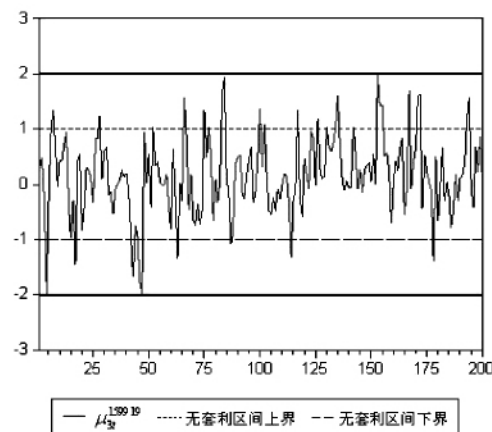


图 4  $\mu_{3t}^{159919}$  序列及无套利区间

依照上文理论部分所述的开平仓标准和  $\mu_{3t}$  序列图来看,两个交易所的沪深 300ETF 都出现了大量套利机会。但在上交所上市的华泰柏瑞沪深 300ETF 有两次套利失败,认亏平仓;而在深交所上市的嘉实沪深 300ETF 的套利,则全部成功获利平仓。

### 三、结论与启示

本文用两种方法实证检验了分别在沪深两个交易所上市的沪深 300ETF 与股指期货的套利效果。结果发现,协整统计套利法的适用范围比基于持有成本理论的套利方法更广泛,这是由于持有成本理论的套利方法面临着股息收益率和单向交易成本不易确定的限制。本文的实证研究表明:在深交所上市的嘉实沪深 300ETF 用两种方法都可以发掘到套利机会,其中协整的方法发掘到的套利机会更多,且失败概率较小;在上交所上市的华泰柏瑞沪深 300ETF 因交易成本不能确定而不能采用持有成本套利法,只能用协整统计套利才能发掘到套利机会,并且套利失败的概率较大。一是因为嘉实沪深 300ETF 的单向交易成本比华泰柏瑞更容易确定;二是因为华泰柏瑞沪深 300ETF 的流动性较高,流动性越高的 ETF 套利机会就越少。

在本文中,仍有大量后续工作可以去做。后来者可以基于本文的套利结果计算出套利收益率,并找到套利收益率的分布,继而利用在险价值(VaR)的方法去研究套利的风险。

#### 参考文献:

- [1] 陈守东,韩广哲,荆伟. 主要股票市场指数与我国股票市场指数间的协整分析[J]. 数量经济技术经济研究, 2003(5): 124 - 129.

- [2]仇中群,程希骏.基于协整的股指期货跨期套利策略模型[J].系统工程,2008(12):26-29.
- [3]葛翔宇,吴洋,周艳丽.门限协整套利:理论与实证研究[J].统计研究,2012(3):79-87.
- [4]黄晓坤,侯金鸣.利用ETF类基金进行股指期货套利的方法研究[J].统计与决策,2009(18):132-134.
- [5]刘伟,陈敏,梁斌.基于金融高频数据的ETF套利分析[J].中国管理科学,2009(2):1-7.
- [6]Burgess N. Statistical Arbitrage Models of the FTSE 100 [J]. Computational Finance, 1999, The MIT Press, 2000: 297-312.
- [7]Board J, Sutcliffe C. The Dual Listing of Stock Index Futures: Arbitrage, Spread Arbitrage, and Currency Risk [J]. Journal of Futures Markets, 1996(2): 29-54.
- [8]Park, T. H. and Switzer, L. N. Index Participation Units and Performance of Index Futures Market: Evidence from the Toronto 35 Index Participation Units Market [J]. Journal of Futures Market, 1995, 15(2): 187-200.
- [9]Robert Engle. Pricing Exchange-Traded Fund [R]. Working Paper, 2002.
- [10]Switzer, L. N., Varson, P. L. and Zghidi, S. Standard and Poor's Depository Receipts and Performance of S&P 500 Index Futures Market [J]. Journal of Futures Market, 2000(20): 705-716.

## An Empirical Analysis of CSI 300 ETF Cash – Futures Arbitrage Opportunity Differentials

Ding Tingli

(Information Management College, Jiangxi Financial University, Nanchang, Jiangxi 330013, China)

**Abstract:** This article compares the differential of the cost-of-carried arbitrage methodology and cointegration-based statistical arbitrage strategy and makes an empirical analysis on both ETFs. This research findings indicate that cointegration-based statistical arbitrage strategy has greater scope of applicability and enjoys a bigger arbitrage opportunity. This comes exactly with the same conclusion based on simulation trading data and other ETF statsits. Meanwhile, on the base of cointegration-based statistical arbitrage strategy, Huatai-PineBridge CSI 300 thus has a smaller chance to fulfill arbitrage, while Harvest CSI 300 enjoys a much bigger opportunity.

**Key words:** CSI 300 ETF; arbitrage opportunity; differential; empirical analysis

(责任编辑: 沈 五)