

# 配对交易策略在A股市场的应用与改进

上海大学悉尼工商学院 丁涛

**摘要:** 本文以在融资融券推出的前提下所诞生的配对交易策略作为主要研究内容,通过总结已有策略,结合我国交易规则与现状,设计出一个具备实用价值的配对交易策略。在实证检验中,本文以2008年1月2日至2010年12月31日的向前复权收盘价数据作为样本区间数据,选取了我国A股市场中配对关系最为显著的中国银行与中信银行作为配对交易标的,利用改进后的延后开仓配对交易策略在样本外进行模拟交易,获得了6.01%的年收益率,这个结果无论从收益率还是从基于风险调整思想下的绩效评价体系角度出发,都强于同期市场收益,证明了经过改良后的配对交易策略在我国市场的可行性与获利性。

**关键词:** 配对交易 可配对指数 协整 延后开仓

**中图分类号:** F832

**文献标识码:** A

**文章编号:** 1005-5800(2013)02(b)-024-04

## 1 研究背景与意义

融资融券交易,是指投资者向证券公司提供担保物,借入资金买入上市证券或借入证券并卖出的行为。融资融券业务是目前我国证券行业推进较快,规模初现且最有盈利前景的创新业务,截至2012年6月底共有278只股票和9只ETF作为融资融券标的,44家证券公司已开展融资融券业务,两融规模从2010年3月31日的659万元扩张到2012年11月9日的752亿元。融资融券提供了一种做空获利的方式,在这种可以双向交易的情形下,能够衍生出许多新的投资方法和投资策略,为包括配对交易在内的统计套利模型在国内市场中的应用扫除了障碍。因此,将被国外对冲基金广泛应用而较少被国内投资者所讨论的配对交易策略应用于国内A股市场,并验证此策略在中国市场上在扣除交易成本后的获利性及可行性,可以在一方面为投资者从事相关交易及研究者进行相关研究提供参考,另一方面也为验证配对交易策略在我国现实市场中的盈利能力和应用前景提供一定的依据。

## 2 配对交易策略与数据选取

配对交易策略是一种市场中性的交易策略,其基本原理是选择两个相关性较高的股票,他们的历史价格之间存在某种确定的长期均衡关系,如果未来一段时间能保持这种良好的相关性,那么一旦两者之间出现背离均值的走势,我们预期这种背离在未来将会得到纠正,从而出现套利机会。股票配对交易策略的顺利实施,涉及两大关键环节:一是如何找寻具有较高相关性并保持某种均衡关系的两只股票,二是如何量化所选股票对的均衡关系或者相关性由于受到某些市场因素的作用而发生异变的程度,也即配对交易中建仓与平仓阈值的设置。在确定了配对标的股票、标的股票具有协整关系以及建仓和平仓阈值之后,可按如下程序进行配对交易操作:当价差大于/小于建仓阈值 $k_1\delta_1/-k_1\delta_1$ ( $\delta_1$ 为价差序列标准差)时,卖出/买入一单位 $y_1$ 股票,买入单位的 $y_2$ 股票,当价差回落至均值时进行反向操作:买入一单位 $y_1$ 股票,卖出相应单位的 $y_2$ 股票进行平仓,最终获利了结。而在交易过程中,当出现价差落在止损阈值( $2.5\delta_1$ )之外时,必须止损平仓从而避免出现更大的损失。

本文所选取的配对标的资产均为来自A股市场现有的278只融资融券标的证券,数据样本为我国A股市场2008年1月2日至2012年1月4日共977个交易日的日收盘价格数据,收盘价均采用向前复权的处理方式。其中,2008年1月2日至2010年12月31日的收盘价数据为构造配对交易策略模型的样本区间数据,本文将利用此期间

的历史数据来选择合适的配对标的资产并构建最优套利策略。而2011年1月4日至2012年1月4日为样本外区间,在此期间内将按照样本区间内确定的配对资产与策略进行模拟交易,通过对模拟交易的风险与收益进行分析最终验证配对交易策略的可行性。在交易成本方面,印花税率定为单边卖出时缴纳0.1%,交易佣金定为双边0.05%,融券年利率定位8.5%,冲击成本固定为双边0.1%,折算率在本文中一律按照60%计算,最低保证金比例设定为60%。此外,由于在实证内容中所交易股票均为不超过1手,可以假设初始资金十分充足,不需考虑资金分配问题。

## 3 配对交易策略的实证检验

### 3.1 配对标的的选择

从配对交易定义出发可以发现:满足要求的股票应该是主营业务类似、规模相近、股价驱动因素大致相同的股票。因此,在初选过程中,可以基本确定满足配对条件的股票应该是出自同一行业。为此,本文根据申万二级行业分类,引入“行业可配对因子”这一指标来考察A股市场各行业的可配对性。该指标的具体计算方法如下。

首先对A股市场现有的278只融资融券标的证券进行行业归类:融资融券标的证券共涉及62个行业(按申万二级行业分类),剔除只包括一只股票的行业,剩下40个行业,共256只股票。考虑到实际操作上的便利以及最后统计结果的可视性,本文仅从40个行业中选取了包含股票数量最多的前8个行业进行跟踪实证,样本区间为从2008年1月2日到2010年12月31日,共352个正常交易日。其后,对排名前八的行业中的各支股票进行行业内的两两配对进而计算其价格的相关系数。最后,对各行业两两配对股票所取得的相关系数取均值即可得到各行业的“行业可配对因子”。该指标处理后在将会在区间 $[-1, 1]$ 的范围内波动,越靠近1说明行业整体的可配对情况越好,或者说行业内股票的集群(或聚类)效应越好,反之则反。对上述8个行业的行业可配对因子这一指标进行统计分析后结果如表1所示。

表1 样本区间各行业整体可配对情况

行业名称	均值	标准差	变异系数
房地产开发	0.585	0.121	0.207
有色金属冶炼与加工	0.642	0.196	0.305
煤炭开采	0.787	0.116	0.147
证券	0.774	0.164	0.211
银行	0.830	0.118	0.142
电气设备	0.514	0.182	0.354
专用设备	0.589	0.212	0.360
饮料制造	0.498	0.120	0.240

从表1中可以看出,银行业股票之间相关性最大,标准差也相对较小,这说明配对的可行性和稳定性均属于各行业前列,存在协整关系的可能性最大。因此,本文选取现有278只融资融券标的证券中的14支银行业股票:浦发银行、华夏银行、民生银行、招商银行、兴业银行、北京银行、交通银行、工商银行、建设银行、中国银行、平安银行、宁波银行、中信银行、南京银行作为研究对象。同属银行业融资融券标的的农业银行与光大银行由于上市时间较晚,无法取得足够样本区间数据,因此将其剔除于研究对象之外。在通过行业可配对因子分析确定了配对股票所处行业之后,分别计算两两之间的相关系数,所得结果如表2。

表2 银行业股票对之间Pearson相关系数矩阵

	PFYH	HXYH	MSYH	ZSYH	XYYH	BJYH	JTYH	GSYH	JSYH	ZGYH	PAYH	NBYH	ZXYH	NJYH
PFYH	1													
HXYH	.822	1												
MSYH	.916	.886	1											
ZSYH	.903	.724	.800	1										
XYYH	.759	.805	.863	.632	1									
BJYH	.616	.691	.739	.471	.959	1								
JTYH	.913	.827	.939	.830	.917	.820	1							
GSYH	.892	.835	.903	.898	.866	.754	.947	1						
JSYH	.887	.832	.904	.900	.867	.762	.951	.939	1					
ZGYH	.926	.845	.946	.882	.868	.753	.961	.955	.942	1				
PAYH	.887	.890	.932	.827	.927	.828	.956	.952	.954	.948	1			
NBYH	.649	.846	.814	.534	.910	.885	.792	.788	.793	.791	.870	1		
ZXYH	.782	.889	.906	.691	.896	.819	.876	.880	.876	.965	.931	.943	1	
NJYH	.575	.576	.696	.385	.913	.938	.783	.670	.667	.703	.744	.784	.735	1

通过对相关系数的测量我们可以发现中国银行与中信银行的日收盘价格相关性最高,达到了0.965。因此,本文在下面的研究中讲选取中国银行与中信银行作为配对交易的标的资产。

### 3.2 协整关系的确定

由于计量经济学中的协整概念能够很好地体现两时间序列之间的长期均衡关系。因此,协整检验成为了验证股票之间是否存在长期均衡关系的常用方法。在运用协整理论判定股票价格序列存在的相关性,首先需要股票价格序列进行平稳性检验,即检验其是否存在单位根。在此,我们采用ADF检验法分别对中国银行与中信银行的日收盘价格序列进行单位根检验。检验结果如表3。

表3 单位根检验结果

变量	Test critical values			t-statistic	Prob	平稳性
	1% level	5% level	10% level			
ZGYH	-3.439068	-2.865278	-2.568816	0.467959	0.9856	非平稳
ZXYH	-3.439068	-2.865278	-2.568816	-0.157701	0.9410	非平稳
$\Delta ZGYH$	-3.439081	-2.865283	-2.568819	-27.50901	0.0000	平稳
$\Delta ZXYH$	-3.439081	-2.865283	-2.568819	-28.81608	0.0000	平稳

从表3中可以看出两只股票的收盘价格序列在经过一阶差分之后,可以在1%的显著性水平下通过检验,即都为二阶单整序列,可能存在长期协整关系。

确认了中国银行与中信银行的收盘价格序列的一阶单整性之后,可以运用EG检验法对两者之间的协整关系进行协整检验。

第一步,运用最小二乘法(OLS)对两只股票的收盘价格序列进行一元回归分析,估计协整回归方程: $ZGYH_t = \beta_1 + \beta_2 ZXYH_t$

运用Eviews7.0估计出的结果如下:

表4 回归估计结果

Variable	Coefficient	Std. Error	t-Statistic	Prob.
$\beta_1$	1.047148	0.058803	17.80777	0.0000
ZXYH	0.494064	0.009675	51.06561	0.0000
R-squared	0.970071	Adjusted R-squared	0.969948	

根据表4可知,两只股票的协整关系(长期均衡关系)为:

$$ZGYH_t = 1.047148 + 0.494064ZXYH_t$$

其后,对其残差序列

$\mu_t = ZGYH_t - 1.047148 - 0.494064ZXYH_t$  进行平稳性检验。

表5 残差序列 $\mu_t$ ADF检验结果

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.687756	0.0287
Test critical values:		
1% level	-3.9	
5% level	-3.34	
10% level	-3.04	

从表中可以看出,残差序列 $\mu_t$ 在5%的置信水平下为不存在单位根的平稳序列。这说明了中国银行与中信银行之间存在着协整关系。

### 3.3 交易信号的确定

确定建仓阈值 $k\delta_1$ 可采用非参数法:以0.01为步长对k从0至2共取200个值,分别计算出不同k值对应的期望套利收益,最终取得期望套利收益最大的一个值作为最终确定的建仓阈值。对K取不同数值时的套利收益计算将采用Matlab编程实现。

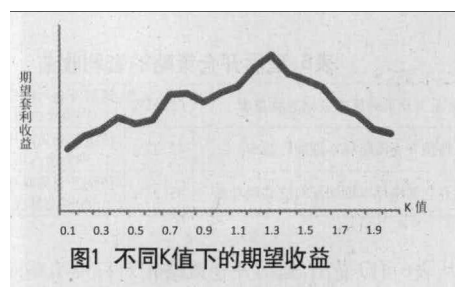


图1 不同K值下的期望收益

由分析可知,当k值在1.2至1.3之间时期望套利收益有最大值,因此,k值可大致确定为1.25。而经测算价差序列的标准差 $\delta_t = 0.3439$ ,因此本文将选取 $1.25\delta_t = 0.4299$ 做为建仓阈值。平仓阈值选取2.5倍标准差即 $2.5\delta_t = 0.85975$ 。

### 3.4 样本期间内套利收益分析

在确定了配对标的资产、资产间的协整关系、交易阈值与平仓阈值之后,可以在样本期间内实施具体的配对交易策略。交易时机图见图2。

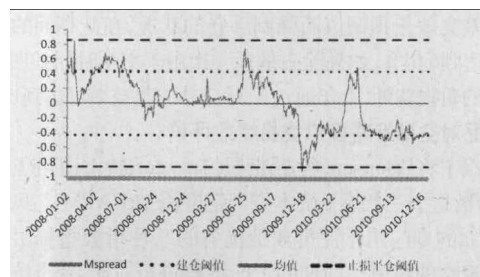


图2 样本期间内交易时机图

从图2中可以看出,在样本期共732个交易日共进行了7次交

易,其中第四次交易由于在2009年12月15日价差序列触及了止损阈值,需要止损平仓,而最后一次由于交易期结束而强制平仓。除了以上两次交易造成了一定损失之外,其余5次交易均实现盈利,累计收益率达到42.27%,年化收益率约为14.09%。

3.5 交易策略的改进

从上文的样本期间配对交易实证分析可以看出,配对交易的亏损主要是由于价差发生剧变,在一定时期内并未发生回归而是直接突破止损阈值导致强制平仓造成的,为了尽量减少这种情况的发生,本文提出了采用等价差突破阈值后,反向回归穿越阈值时再建仓这一延后开仓策略(见图3)。延后开仓的好处在于一方面可以更加准确的判断出价差向均值回归的趋势;当价差出现单边走强(或走弱)的情形时,延后开仓策略不会发出建仓信号,从而可以避免一部分损失,提升配对交易的收益。另一方面,延后开仓策略也可以使得交易期限缩短,从而降低融券息费,提高资金使用效率与收益率。

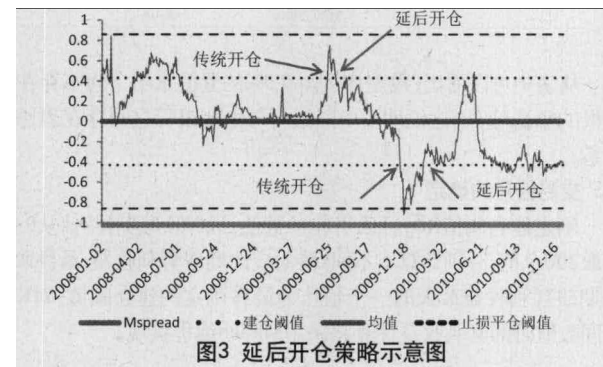


表6 延后开仓策略的套利收益

延后开仓策略样本期间总收益率	42.61%	延后开仓策略样本期间内资金占用天数	433
传统开仓策略样本期间收益率	42.27%	传统开仓策略样本期间内资金占用天数	483
延后开仓策略样本期间配对交易成功率	83.33%	传统开仓策略样本期间配对交易成功率	71.43%

从表6可以看出,延后开仓策略相对标准策略而言虽然总收益率只是略微高出,但是交易期限却缩短了11.55%,也没有发生价差突破止损阈值而导致强制止损的状况发生,从而节约了资金的时间成本,提升了配对交易的交易成功率。

3.6 配对交易模型样本外实证测试

为了检验根据这些参数确定的配对交易策略在样本外的表现,我们取2011年1月4日至2012年1月4日的中国银行与中信银行日收盘价数据进行样本外交易模拟。采用延后开仓策略后,中国银行与中信银行的样本期间外模拟交易时机见图4。

由图4可以看出,在模拟期间内一共存在4次交易机会,没有发生价差突破止损阈值而强制平仓的状况,在此期间的模拟交易收益率达到6.01%,如果除去最后一次由于交易时间到期强制平仓所带来的负收益外,全年通过配对交易的收益率可达到10.41%。

3.7 配对交易策略模拟交易绩效评价

以下将从Jensen(詹森)指数、Trenor(特雷诺)指数和Sharp(夏普)指数这个三个基于资本资产定价模型(CAPM)与风险调整法思想建立的专门用于评价业绩优劣的三种指数指标出发来将配对交易策略的绩效与同期的上证指数进行对比。令市场无风险收益率为一年期定期存款利率——约为3%。表7揭示了这些指标的对比情况。

比情况。

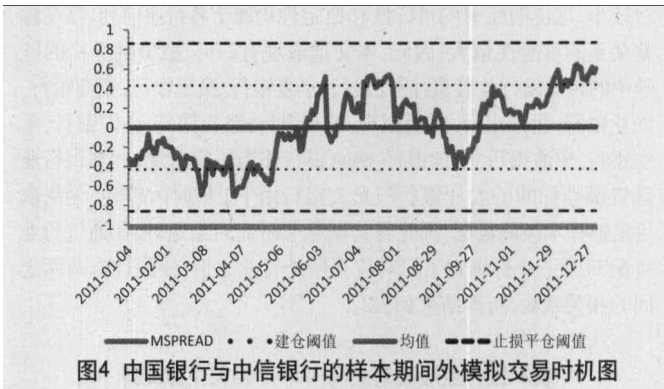


表7 模拟期间配对交易绩效分析

策略	年收益率	贝塔系数	詹森指数	特雷诺指数	夏普指数
配对策略	6.01%	0.004	3.12%	7.49	8.62
上证指数	-23.95%	1	0%	-0.2695	-23.31

从收益率情况来看,在行情走势低迷的2011年,配对交易策略仍能获得6.01%的较为稳定的收益率,远远高于上证指数。而从詹森指数、特雷诺指数与夏普指数来看,配对交易所得数值分别高于甚至显著高于上证指数,这说明即使考虑入交易策略所承担风险大小后,配对交易策略的表现仍然优于市场,获利效率更高,收益稳定性更好。此外,在模拟交易期中,配对策略的贝塔系数仅为0.004,这也体现了这种交易策略市场中性的特性:所获收益与市场风险相关性较弱。

4 结语

本文采用我国A股市场可配对程度最高的银行业中的两只相关性较好的股票——中国银行与中信银行作为配对交易策略的标的资产,选取其自2008年1月2日至2012年1月4日的日共977个交易日的日收盘价格数据进行了配对交易的实证检验,在原有的交易策略的基础上对开仓时机的选择进行了改进,提出了延后开仓策略,获得了不错的收益。最后,我们还对此次模拟交易的绩效通过Jensen(詹森)指数、Trenor(特雷诺)指数和Sharp(夏普)指数进行了深入分析,结果显示,配对交易策略在我国市场上存在着套利的可行性与获利性。

参考文献

[1] Gatev, e., Goetzmann, W.N., and Rouwenhorst, K.G., Pairs trading: Performance of a relative-value arbitrage rule[J]. Review of Financial Studies, Vol. 19, No. 3, 2006.

[2] Elliott, R., van der Hoek, J. and Malcolm, W., Pairs Trading[J]. Quantitative Finance, Vol. 5, No. 3, 2005.

[3] B. Do, R. Faff, and K. Hamza, A new approach to modeling and estimation for pairs trading[C]. In Proceedings of 2006 Financial Management Association European Conference, Stockholm, 2006.

[4] Engle, R. and Granger, C., Cointegration and Error Correction: Representation, Estimation and Testing[J]. Econometrica, Vol. 55, No. 2, 1987.

[5] 曾俊涵. 权益成对交易策略与台湾股市之实证研究[D]. 高雄第一科技大学, 财务管理所硕士论文, 2004. (下转028页)

结果的影响在短期内不大,因而在短期内比较精确地预测是可行的。所谓的短期是指  $t < 1/M_1$ , 式中,  $M_1$  是系统里亚普诺夫指数中的最大者, 根据文献《上证指数中的奇异吸引子》中关于市场的里亚普诺夫指数约为 0.0107, 也就是说大约在 90 日交易日之后初始条件的误差将会成倍的增长, 超过这个期限的预测是没有价值的<sup>[9]</sup>。黄金是一种具有商品属性、货币属性、金属属性三种属性的特殊商品, 其价格波动受到如通货膨胀率、汇率、利率等方面的影响, 且各种因素对黄金价格的影响都有一定的内在随机性, 这为我们合理把握黄金价格波动范围防范其风险带来困难。正如 2008 年金融危机时期的黄金价格走势。市场普遍认为黄金的价格会因为美元的走软和其特有的避险属性一路升高, 然而一段时间后由于市场缺少资金流动性和大量的获利盘导致许多机构和投资者卖出黄金以持有更具流动性的资金, 导致黄金价格大幅下跌, 不少投资者蒙受了不必要的损失。鉴于市场的短期精确预测是可行的, 投资者要尽量减少黄金持有的时间, 充分利用各种信息对黄金的短期走势做出正确判断, 合理把握利润, 减少长期过程中不确定因素导致的价格波动带来的风险, 避免不必要的损失。

## 2.2 基于混沌初值敏感性的投资策略

混沌的初值敏感性就如“蝴蝶效应”, 但并不是只与不利的事联系在一起。一个好的微小机制, 如果正确引导, 将会产生轰动“蝴蝶效应”。所以, 投资者要充分利用“蝴蝶效应”, 使其起到“四两拨千斤”的效果。

混沌的初值敏感性贯穿于整个黄金价格形成的过程, 因此, 投资者必须找准并控制其形成的关键因子(分岔参数)。及时变更投资策略, 防范由于黄金价格大幅波动所带来损失的风险。总之, 在黄金投资过程中投资者必须找准关键因子, 尽量减少决策上的失误。

黄金价格的波动是不确定的, 引起其价格波动不确定性的因素有环境和投资者自身两个方面, 其中环境因素包括宏观经济政策、宏观经济形势、国与国之间的关系、敏感地区的繁荣与稳定、信息分享的通畅和准确程度, 投资者自身因素包括心理预期、投资观、消费者独特投资特性、投资者对黄金的信任程度、投资者在投资过程中的相互感染的程度。这里许多都可以成为黄金价格波动的关键因子(分岔参数), 投资者在黄金价格形成的输入端找准关键因子做出及时正确的投资决策, 采取果断的措施, 从而减少黄金价格大幅波动而带来损失的风险。比如在影响黄金价格波动的重大经济数据——美国月度非农数据出炉前, 投资者要充分利用之前的各种经济数据做出对数据的合理判断并根据这种判断分析其对数据出炉后黄金价格波动的影响, 采取一系列的风险防范措施。也可以在数据出炉后, 形势明确后再做出黄金投资决策以规避由于其初值敏感性带来的价格波动风险。又如信息方面, 投资者比较重视与黄金价格波动密切相关的信息的搜索。当下, 互联网被普遍当做一种信息来源, 投资者必须学会利用更加迅捷的信息工具有用的信息及时“流”向自己, 使自己掌握第一手的信息从而利于投资者早于市场做出正确的判断和反应, 防范黄金价格波动所带来的风险。又如对于投资者的投资心理, “知己知彼, 百战不殆”。投资者要准确把握整个市场的投资心理的微小变化, 及时准确地做出分析和判断, 并采取果断的止盈止损措施以规避黄金价格波动所带来的风险。

## 2.3 基于混沌规律性的投资策略

黄金价格波动的形成过程具有规律性, 投资者如果能认识掌握这个过程, 是可以有效防范黄金价格波动的风险并获得可观收

益的。混沌并非混乱, 混沌中有规律。比如, 黄金价格与美元指数的变动有一定的联动关系。在大多数情况下, 美元指数上升, 由美元计价的黄金价格就会下跌。反之则会上升。投资者可以在正确判断美元走势的基础上, 通过对美元对黄金影响规律的认识做出相应的投资策略。由于黄金还具有一定的商品属性, 如果供过于求黄金的价格就会下跌, 如果供不应求黄金的价格就会上升。投资者可以通过各种渠道获知黄金的供求数据, 并可以通过对这些数据的处理分析知道黄金市场的供求关系, 在此基础上可以在较小的风险敞口下获得可观的收益。所以, 投资者可以在把握黄金价格波动的客观规律的基础上做出合理的黄金投资决策, 在较小的风险敞口下获得较为丰厚的收益。

## 参考文献

- [1] 方超逸. 国际黄金价格走势分析和预测方法[J]. 黄金, 2009(11).
- [2] 黄润生, 黄浩. 混沌及其应用[M]. 武汉: 武汉大学出版社, 2007.
- [3] 张治, 韩国照. 市场营销的方法论问题[M]. 北京: 中国经济出版社, 2006.
- [4] 乔一桢. 国际黄金价格与欧元兑美元汇率和中国上证综指三者关系探究[J]. 时代金融, 2011.
- [5] 张绍飞. 伊拉克战争与世界石油形势[J]. 石油化工技术经济, 2003(4).
- [6] 张震. 探析黄金、美元和石油之间的互动关系——基于分量回归模型的再探讨[J]. 财政金融, 2010(11).
- [7] 邱红, 李肖钰. 影响黄金价格的因素分析[J]. 东方企业文化. 公司与产业, 2010: 86.
- [8] 侯荣华. 经济学研究混沌现象的必要性[J]. 上海海运学院学报. 交通学院, 1000-5188(2002).
- [9] 徐前方. 上证指数中的奇异吸引子[J]. 数量经济技术经济研究, 1994(2).

## (上接026页)

- [6] 罗君昱. 台湾股票市场执行套利统计之可行性分析[D]. 国立政治大学经营管理硕士学程硕士论文, 2005.
- [7] 方昊. 统计套利的理论模式及应用分析——基于中国封闭式基金市场的检验[J]. 统计与决策, 2005(6).
- [8] 陈守东, 韩光哲. 统计套利模型研究——基于上证50指数成分股的检验[J]. 数理统计与管理, 2007, 9(5).
- [9] 宋曦. 统计套利研究系列——基于协整的成对交易[R]. 联合证券研究报告, 2007.
- [10] 孙华强, 王红兵. 利用协整提高股指期货跨期套利成功率[R]. 联合证券研究报告, 2007.
- [11] 于玮婷. 基于协整方法的统计套利策略的实证分析[J]. 科学决策, 2011(3).
- [12] 吴先兴. 算法交易研究系列三: 统计套利之股票配对交易策略[R]. 海通证券研究报告, 2011.
- [13] 张连华. 基于高频数据的股指期货期现统计套利程序交易[J]. 计算机应用与软件, 2011(28).
- [14] 魏晓兰. 融资融券与权证行权的无风险套利分析[J]. 商业时代, 2011(10).
- [15] 温予群, 刘洪光. 基于沪深300的统计套利的实证研究[J]. 金融经济, 2011(6).