

Research of The Statistical Arbitrage Model of Gold Futures Contract Based on Cointegration

Congwen Ding^{1,a}, Kai Li^{2,b,*} and Ran Liu^{3,c}

¹University of Technology and Science Beijing, China

² University of Technology and Science Beijing, China

³ University of Technology and Science Beijing, China

^adingcongwen921@163.com, ^blikai2000@ustb.edu.cn, ^calisa11052@163.com

*Corresponding author

Abstract. The basic of our study is co-integrate theory. We use Augmented Dickey-Fuller (ADF) test and Engle-Granger (EG) co-integration test to analyze the relationship of closing price of AU1512 gold futures and the price of gold in the same period. According to our analysis, we establish error correction model (ECM) and the statistical arbitrage model. Based on the deviation of short-term price and long-term equilibrium price of gold futures, we make paper-trading to verify the feasibility of the statistical arbitrage model. In addition, we predict the possible trend of gold futures' price and if there will exist arbitrage opportunity.

Keywords: Gold Futures, Stationarity, Co-integrate theory, Statistical Arbitrage

基于协整的黄金期货合约统计套利模型研究

丁丛文^{1, a}, 李凯^{2, b, *}, 刘然^{3, c}

¹北京科技大学东凌经济管理学院, 海淀区, 北京市, 中国

²北京科技大学东凌经济管理学院, 海淀区, 北京市, 中国

³北京科技大学东凌经济管理学院, 海淀区, 北京市, 中国

^adingcongwen921@163.com, ^blikai2000@ustb.edu.cn, ^calisa11052@163.com

*通讯作者

中文摘要. 本文以协整为基础, 通过对AU1512黄金期货合约的收盘价和黄金现货价格进行ADF检验、Engle-Granger协整检验, 建立误差修正模型(ECM), 得出统计套利模型。利用短期黄金期货价格与长期均衡价格的偏离进行期现套利交易模拟, 实

证检验统计套利模型的可行性; 通过分析历史期货实际价格与理论价格的偏差, 预测未来一段时间内期货价格的走势与可能存在的套利机会。

关键词: 黄金期货; 平稳性; 协整; 期现统计套利

1. 引言

统计套利方法来源于摩根斯坦利Tartaglia研究团队上世纪80年代的创新,是一种通过对历史数据进行统计分析,从而得出实际价格偏离理论价格的模型,并通过进一步建立相应金融资产的多头或空头而获得利润。

在国外学者的研究中,Mackinlay等(1988)基于S&P指数期货与其标的指数数据的分析,证明了随着交割日的临近二者间存在路径依赖性的错误定价关系。^[1]

Burgess等(1999)采用了逐步回归法和误差修正模型ECM方法研究FEST100与其成分股,将协整的概念定义为了统计套利模型的基础。协整方法给出了两种资产的长期均衡关系,误差修正模型则提供了短期的定价偏离,价差的存在给市场中投资者提供了套利空间。^[2]Hogan等(2003)取样于1965年1月至2000年12月美国证券交易所、纽约证券交易所和NASDAQ证券交易所上市公司股价数据,测试了市场效率使用统计套利与应用动量和价值策略。^[3]

黄金期货是以国际黄金市场未来某时点的黄金价格为交易标的的期货合约,交易采取多空双向交易机制,交易标的是符合国标GB/T4134-2003规定、含金量不低于99.95%的金锭。

我国贵金属期货合约交易集中在上海期货交易所。其中黄金期货于2008年1月9日正式在上海期货交易所上市,现采用T+0的交易制度,即当日可进行买入卖出,具有较强的流动性,为套利提供了有利条件。经过7年多的发展,我国黄金期货市场不断发展,投资者数量日益增加。

目前国内对于黄金期货合约的研究主要关注于,波动性特征分析和价格发现分析。王兆才等(2012)基于GARCH模型对中国黄金期货市场的量价关系进行实证研究,结果发现,期货市场的交易量与价格波动之间存在正相关关系,而持仓量对价格波动有负相关的影响关系。^[4]徐伟等(2009)借助蛛网理论、持有成本理论和理性预期理论分析了期货市场价格发现功能的机理。分析了期货市场价格发现的制度基础,并且将影响期货市场价格发现效率的因素期货市

场和现货市场两个方面进行了阐述,认为期货市场自身的因素包括交易成本、交易主体结构和市场的流动性对期货市场价格发现效率有很大影响。^[5]

从统计套利角度出发对黄金期货套利的研究较少,强盛等(2013)利用了协整分析方法,基于上海交易所黄金和白银主力合约历史数据,对统计套利策略在中国贵金属市场上应用进行了实证分析。整体上对于商品期货的统计套利研究较少,多数学者的研究集中于股指期货合约当中。^[6]李乐等

(2014)基于沪深300股指期货合约数据,进行了高频日内跨期统计套利策略研究,实证了其套利策略算法有效性。^[7]

本文选取上海期货交易所黄金期货合约和黄金现货为研究对象,基于统计套利模型,对二者的日价格时间序列进行协整检验,并以检验结果为依据建立误差修正模型(ECM)。验证两者的价格在长期存在均衡关系,建立短期均衡价格模型,即期货理论价格模型,利用短期黄金期货价格与均衡价格的偏离进行期现套利交易模拟,并对套利结果进行分析。同时我们也通过期货理论与实际价格的偏离,得出了未来短期内黄金期货的价格走势,对投资者的套利投机等具有一定指导意义。

本文的创新之处有三点:1、以协整为基础,对黄金期货和现货进行统计套利的研究。2、通过一个月多次套利实证分析,验证了短期均衡价格模型的可靠性。3、利用历史期货实际价格与理论价格的偏差,预测未来一段时间内期货价格的走势及可能存在的套利机会,为投资者提供了一定的建议。

2. 数据来源

目前,上海期货交易所上市交易的黄金期货合约共有7种,如表1所示。为了分析的严谨性和准确性,我们观察了一周内各黄金期货的成交量和持仓量的变化情况,通过对比分析,我们选择了黄金1512期货进行数据统计及相关分析,得出其统计套利模型。

表1 黄金期货的成交量和持仓量

名称	成交量	持仓量
黄金1506	103940	101680
黄金1507	16	40
黄金1508	28	152
黄金1510	18	92
黄金1512	50306	103120
黄金1602	8	60
黄金1604	6	34

数据来源：WIND数据库

表1中还反映了2015年5月5日不同交割日期黄金期货的成交量和持仓量情况，从表中数据，我们不难看出，黄金1506和黄金1512的成交量和持仓量均处于较高水平，更具有统计套利的研究价值。然而黄金1506期货合约于2015年6月就到期，对其进行统计套利模型研究的实际意义不显著，于是，我们选择了黄金1512期货合约作为我们的研究对象。

黄金1512期货合约2015年1月上市，将于2015年12月到期。我们对该期货合约2015年1月26日至4月30日的期货收盘价进行了统计，并整理了同时期的黄金现货价格，共得到63组数据。

表2 黄金期货合约主要条款

交易品种	黄金
交易单位	1000克/手
报价单位	元（人民币）/克
最小变动价位	0.05元/克
每日价格最大波动限制	不超过上一交易日结算价±3%
合约交割月份	最近三个连续月份的合约以及最近11个月以内的双月合约
交易时间	上午9:00—11:30，下午1:30—3:00和交易所规定的其他交易时间
最后交易日	合约交割月份的15日（遇法定假日顺延）
交割日期	最后交易日后连续五个工作日
交割品级	金含量不小于99.95%的国产金锭及经交易所认可的伦敦金银市场协会（LBMA）认定的合格供货商或精炼厂生产的标准金

锭	
交割地点	交易所指定交割金库
最低交易保证	合约价值的4%
金	
交割方式	实物交割
交易手续费	10元/手
交易代码	AU

3. 黄金期货理论价格模型

3.1 黄金期、现货价格平稳性检验

记黄金期货收盘价为F，其时间序列为{F_t, t=1, 2, 3……}, t是时间变量，代表了时间序列随时间变化的某种趋势。记黄金现货价格为P，其时间序列为{P_t, t=1, 2,3……}。若要对黄金期货、现货价格进行经典回归分析，那就要求两个时间序列均是平稳的，否则会有存在伪回归的问题。因此，为了对二者的相关关系进行研究，我们必须先探究两个时间序列的平稳性。

本文采用ADF检验，对两个时间序列进行了平稳性检验。我们建立了以下三个模型。

模型1: $\Delta X_t = \delta X_{t-1} + \sum_{i=1}^m \beta_i \Delta X_{t-i} + \varepsilon_t$

模型2: $\Delta X_t = \alpha + \delta X_{t-1} + \sum_{i=1}^m \beta_i \Delta X_{t-i} + \varepsilon_t$

模型3: $\Delta X_t = \alpha + \beta t + \delta X_{t-1} + \sum_{i=1}^m \beta_i \Delta X_{t-i} + \varepsilon_t$

其中 $\Delta X_t = X_t - X_{t-1}$ ，我们分别对时间序列{F_t}、{P_t}进行三个模型的检验。通过t统计量的检验，我们可以判断时间序列的平稳性。若t统计量的值小于临界值，则认为序列平稳，反之亦然。当三个模型的检验结果t统计量的值均大于临界值，则认为时间序列非平稳。

图1、图2、图3，是我们对{P_t}时间序列进行ADF检验的结果，t统计量均大于临界值，因此{P_t}序列非平稳。图4、图5、图6，是我们对{F_t}时间序列进行ADF检验的结果，同理可知，{F_t}序列也是非平稳的。

Null Hypothesis: P has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic based on SIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.784089	0.7001
Test critical values: 1% level	-4.118444	
5% level	-3.486509	
10% level	-3.171541	

图1 Pt的模型3 ADF检验结果

Null Hypothesis: P has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic based on SIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.206756	0.2061
Test critical values: 1% level	-3.544063	
5% level	-2.910860	
10% level	-2.593090	

图2 Pt的模型2 ADF检验结果

Null Hypothesis: P has a unit root
Exogenous: None
Lag Length: 0 (Automatic based on SIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.058963	0.2585
Test critical values: 1% level	-2.604073	
5% level	-1.946348	
10% level	-1.613293	

图3 Pt的模型1 ADF检验结果

Null Hypothesis: F has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic based on SIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.213532	0.4735
Test critical values: 1% level	-4.118444	
5% level	-3.486509	
10% level	-3.171541	

图4 Ft的模型3 ADF检验结果

Null Hypothesis: F has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic based on SIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.268119	0.1855
Test critical values: 1% level	-3.544063	
5% level	-2.910860	
10% level	-2.593090	

图5 Ft的模型2 ADF检验结果

Null Hypothesis: F has a unit root
Exogenous: None
Lag Length: 0 (Automatic based on SIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.060113	0.2581
Test critical values: 1% level	-2.604073	
5% level	-1.946348	
10% level	-1.613293	

图6 Ft的模型1 ADF检验结果

3.2 黄金期、现货价格协整检验

由上述平稳性检验我们可知，期货、现货价格时间序列均是非平稳的。为了得到期、现货价格之间的相关关系，我们分别对两个时间序列进行差分并对差分后的序列进行ADF检验，我们发现，{Ft}、{Pt}序列均是一阶单整序列，两个时间序列可能存在协整关系。

图7 Ft 1阶差分

Null Hypothesis: D(F) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic based on SIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-9.160799	0.0000
Test critical values: 1% level	-4.127338	
5% level	-3.490662	
10% level	-3.173943	

图8 Pt 1阶差分

Null Hypothesis: D(P) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic based on SIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-9.673287	0.0000
Test critical values: 1% level	-4.127338	
5% level	-3.490662	
10% level	-3.173943	

本文采用EG检验判断P、F的协整关系，首先，我们对F与P进行如下协整回归：

$$\ln \hat{F}_t = 0.580 + 0.897 \ln P_t$$

(4.48) (38.06)

$$R^2 = 0.9596 \quad D.W = 1.77$$

通过该式计算的残差 e_t 序列进行ADF检验，使用模型3，得到相应的检验模型为

$$\Delta e_t = 0.001 - 1.180 e_{t-1} + 0.277 \Delta e_{t-1} + 0.352 \Delta e_{t-2}$$

(0.85) (-5.45) (1.56) (2.63)

在1%的显著性水平下，协整的ADF检验的临界值为-4.137， e_{t-1} 前参数的 t 值-5.451，因此拒绝存在单位根的假设，表明残差项是稳定的。据此判断出，黄金期货合约价格的对数序列 $\ln F$ 和黄金现货价格的对数序列 $\ln P$ 是(1,1)阶协整的，由此可以得出这两个变量的对数序列之前存在长期均衡关系。

3.3 黄金期货理论价格模型

上述检验了黄金期货价格F与黄金现货价格P的对数序列间呈现协整关系。本文选择误差修正模型（ECM）得出黄金期货的理论价格。以lnF关于lnP的协整回归中的稳定残差序列et作为误差修正项，可建立如下的误差修正模型：

$$\Delta \ln \hat{F}_t = 0.951 \Delta \ln P - 0.932 e_{t-1}$$

(10.31) (-6.64)

$$R^2 = 0.653$$

由误差修正模型，我们可得黄金期货理论价格模型：

$$\hat{F}_t = \exp(0.951 \Delta \ln P_t - 0.933 e_{t-1} + \ln F_{t-1})$$

而误差修正项系数反映了短期波动偏离长期均衡时的调整力度，根据Madhavan和Smidt的回复半周期计算公式

$$\frac{T}{2} = \left| \frac{\ln 2}{\ln(1 + \gamma)} \right|$$

可得短期偏差向长期均衡回归半周期是2.05交易日，回复速度快。

4. 实证检验

根据上述所得的黄金期货理论价格模型，我们计算得到2015年3月的黄金期货理论价格，如表3所示。若理论期货价格高于期货实际价格，则期货价格被低估，期货价格将会呈上升趋势，现货价格将会下跌，会存在正向套利机会，反之，则可以进行反向套利。

从表3中，我们发现，期货价格大部分时间处于被高估状态，因此期货价格将在较长一段时间内处于下跌趋势，从我们所统计的数据中，我们也不难看出这样的趋势。在这样的情况下，市场可能存在较多的反向套利机会。2015年3月，期货价格也存在着短期被低估的情况，因此也可能有正向套利的机会。通过分析，我们一共发现了3次套利机会，1次正向套利，2次反向套利。（本文中忽略了手续费、利息等各项费用支出）

表3：黄金期货实际价格和理论价格对比

日期	现货价格P/元	期货理论价格F/元	期货价格F/元	套利空间
20150302	247.30	247.8944	251.15	-3.2556
20150303	244.60	245.0673	249.20	-4.1327
20150304	244.60	244.3467	248.60	-4.2533
20150305	244.60	244.1766	247.75	-3.5734
20150306	242.33	242.6259	246.90	-4.2741
20150309	242.33	242.1170	241.85	0.2670
20150310	235.60	239.8002	239.50	0.3002
20150311	235.60	240.4512	240.15	0.3012
20150312	233.03	237.9994	240.65	-2.6506
20150313	234.45	236.7752	239.50	-2.7248
20150316	234.48	236.4766	240.10	-3.6234
20150317	233.60	234.7832	238.50	-3.7168
20150318	233.60	234.6609	235.90	-1.2391
20150319	234.70	238.0101	239.75	-1.7399
20150320	234.44	237.4113	239.45	-2.0387
20150323	236.53	239.1310	242.30	-3.1690
20150324	236.53	237.9781	242.90	-4.9219
20150325	238.00	237.6819	243.20	-5.5181
20150326	239.92	238.7628	245.90	-7.1372
20150327	240.78	237.9820	244.90	-6.9180
20150330	239.05	236.3827	242.00	-5.6173

3月10日，期货价格被低估，可以进行正向套利。多头1手期货，合约保证金为1000*239.5*4%=9580元，同时卖出黄金9580/235.6=40.66克。3月12日，买入40.66克黄金，同时将黄金期货平仓。套利组合可获利1000*（240.65-239.5）+（235.6-233.03）*40.66=1255.50元。

3月16日，期货价格被高估，可以进行反向套利。空头1手期货合约，合约价格为1000*240.1=240100元，同时买入黄金240100/234.48=1023.97克。3月19日，卖出黄金，同时将期货平仓。套利组合可获利1000*（240.1-239.75）+1023.97（234.7-234.48）=575.27元。

3月26日，期货价格被高估，可以进行反向套利。空头1手期货合约，合约价格为1000*245.9=245900元，同时买入黄金245900/239.92=1024.92克。3月27日，卖出黄金，同时将期货平仓。套利组合可获利1000*（245.9-244.9）+1024.92*（240.78-239.92）=1881.43元。

三次套利周期都较短，平均约两个交易日，与上述所得回归半周期数接近。

综上所述,我们的黄金期货理论价格模型较为准确,可以用以进行短期预测,寻找套利机会。我们也对2015年4月的期货理论价格进行了计算,同样得出期货价格长期被高估的结论,黄金期货的价格呈下跌趋势,存在反向套利的机会。

5. 结束语

本文通过对AU1512黄金期货合约和黄金现货价格的分析,基于协整对该黄金期货进行统计套利模型的研究,并进行了实证研究,得出了以下结论:

1、黄金期现货价格存在长期协整关系,短期黄金期货价格偏离长期均衡价格,符合统计套利的观点,可以寻求套利机会。

2、短期均衡回归半周期为2.05交易日,回复速度快。

3、可以基于期货实际价格和理论价格之差,预测黄金期货是否可以进行相应的正向或反向交易实现期现无风险套利。

目前市场黄金期货的价格长期处于被高估状态,预计黄金期货的价格在未来一段时期内会处于总体下跌趋势,可能会存在着较多的反向套利机会。投资者可以根据市场情况,抓住套利机会,采取适当的措施,从中获利。

致谢

本文为北京市青年英才计划(YETP0396)的阶段性成果之一

References

- [1] Mackinlay A C, Ramaswamy K. Index-futures arbitrage and the behavior of stock index future prices [J]. Review of Financial Studies, 1988, I (2):137-158.
- [2] Burgess A N. Statistical arbitrage models of the FTSE 100 [J]. Computational Finance, 1999, 4(1):297-312.
- [3] Hogan S, Jarrow R, Teo M, et al. Testing market efficiency using statistical arbitrage with applications to momentum and value strategies [J]. Journal of Applied Mathematics Decision Sciences, 2004, 73(3):525-565.
- [4] Zhaocai Wang. Fluctuation Characteristics and risk research of Chinese gold futures market [D] Fudan University, 2012
- [5] Wei Xu. The empirical research of Chinese gold futures' price discovery function [D] Zhejiang University, 2009
- [6] Sheng Qiang. Application of statistical arbitrage in the investment decision of Chinese precious metals futures [D] East China University of Science and Technology, 2014
- [7] Le li, Chunyi Zhang, Zhishu yang. Intraday high frequency intertemporal statistical arbitrage strategies based on CSI 300 index contracts [J] Journal of Tsinghua University (Science and Technology), 2014, 08:1080-1086.