

China's Gold Statistical Arbitrage

JiuHong Yu

Center for financial engineering
school of mathematical science Soochow University
Suzhou, China
E-mail: yjh@suda.edu.cn

ZhiBo Wang

Center for financial engineering Soochow University
Suzhou, China
E-mail: zhibolovexian@outlook.com

Abstract—This article tries to find an arbitrage strategy in the gold prices market of China. And the method is the linear regression based on the time serious date, the gold price, the gold companies' stock prices and other metals prices from 2002 to 2014, and the data of different period, so we tried to test every possible relationship which based on the financial theory between these dates. At last we find there really is a relationship which is statistic tested significantly between the gold price and gold company stock price and based on a special time period date, so we can base the relationship to build an arbitrage strategy, then we test it by the real world dates, and we found it is good, what is more, this method we can use it in the other similar market, for example, silver or oil. And this is not just a method, we can really use it.

Keywords-Stock Price; Gold Price; Arbitrage strategy; Time series; Gold price volatility

I. INTRODUCTION

All over the world, the traders are always looking for new ideas to avoid the risk and increase the returns on their investment. Although the researchers have done a lot of research about this. And the traders now have some way to get it. But the higher the return is, the higher the risk will be. So that, we have to find a way can produce a higher return and more stable predictions and control.

So, in this article, we will examine the relationship between the commodities, oil price changes and stock index returns in China. And then we select single stocks to set up an arbitrage trading strategy for the stock.

As all we know, commodities and oil are the basic elements of pushing economic growth; they are the basic input materials of all industries. So the price changes both of commodities and oil to affect the world economy straight away. The whole commodity price changes combined with crude oil price changes, will have to yield more accurate forecast.

For the commodity and stock return relations, due there are many kinds, like silver, zinc, gold, nickel etc, copper... These relationships may not be all the same. So what we need is to verify all relationships for each one, so that we may get a more accurate result.

In this case, we have China stock returns selected for the forecast by commodity and oil price changes.

In the article, we have discussed some articles about similar topics done by other researchers. These articles showed that, there are some kinds of relationships between commodity price changes with stock returns and oil price changes with stock returns. But in this article, our target is to combine commodities and oil together to produce an accurate forecast of stock returns.

For the methodology section, we will introduce our data and where we get them from and also use E-views software to test if the relationship is significant or not. In the end of the section, we also do a robustness test for all and select the most correlated items and that can have the greatest accurate forecast power for stock returns.

Then in the analysis section, we will use the methodology with the real world data, to test the twelve years of data for a significant relationship. Then we will select the most significant commodities and countries, keep these for further analysis and develop arbitrage trading strategies.

In the section developing arbitrage trading strategy, by using results from the analysis section, we select the most significant commodities and select three single company shares in that industry, and calculate the relationship between the company share returns and commodity price changes.

II. METHODOLOGY

In this research article, the data period spans twelve years from 2002 to 2014. Also the period average and average return have been calculated for all the data as well. But unfortunately, the date of every element is not the same period, but we can compare them in a same period automatically by the software.

I use the following daily and weekly data in the research report:

Gold au-99.5	2002-2014
Copper futures index	2008 - 2014
Oil futures index	2010 - 2014
Lead futures index	2011 - 2014
Silver futures index	2012 - 2014
Zinc futures index	2008 - 2014
Aluminium futures index	2008 - 2014
SSE constituent index	2002 - 2014

All data above are downloaded from <http://www.gw.com.cn/>. Here, we use the futures index, not the spot price, because we find that, Most of the operations of the commodities companies depend on the futures hedge for reducing the risk in the future, so that the futures index is more and more important, and there is a really correlation between the spot price and the futures index.

The SSE constituent index is designed to show the most objective situation of the investment environment for the investors in the Shanghai stock exchange from July 1, 2002.

In this article, we use E-views to test all the relationships in the data, to see if they are significant. And I have created an equation for it.

$$R_t = \beta_0 + \beta_1 r_{t-x}^{\beta_1} + \beta_2 r_{t-x}^{\beta_2} + L + \beta_n r_{t-x}^{\beta_n} \text{ Equation (1)}$$

As we see from the equation:

R_t : return on stock; R : return on commodity; β_0 : regression constant; β_n : the slope of the regression ($n \geq 1$); x : the Lagging days

There can be more elements to represent more goods. We will add more to this equation to test them and get a significant result. To use (t-1) period is used by the previous day's price change in the commodity and oil to predict period (t) stock returns. But in the real world, there comes a time difference problem; the (t-1) may be insignificant for the test. We may need to move to the next (t-10) and (t-30) period to avoid the time difference problem.

Step One:

In our equation, we test all relationships between commodities and oil with stock returns using E-views. If there's any element insignificant in the first step, we will take it out and continue to the next test until we have the entire set of commodities that were significant with stock returns in the China.

$$\text{E.g. } R_t = \beta_0 + \beta_1 r_{t-1}^{\beta_1} + \beta_2 r_{t-1}^{\beta_2} + L + \beta_n r_{t-1}^{\beta_n} \text{ Equation (2)}$$

Step Two:

For the robustness test, to avoid a "time difference" situation, the entire set passing the first test will go on to the next step. The second step is to test them with the (t-7) period returns.

$$\text{E.g. } R_t = \beta_0 + \beta_1 r_{t-7}^{\beta_1} + \beta_2 r_{t-7}^{\beta_2} + L + \beta_n r_{t-7}^{\beta_n} \text{ Equation (3)}$$

Step Three:

The third step is to test them with the (t-30) period returns for avoiding the "time difference" situation also.

$$\text{E.g. } R_t = \beta_0 + \beta_1 r_{t-30}^{\beta_1} + \beta_2 r_{t-30}^{\beta_2} + L + \beta_n r_{t-30}^{\beta_n} \text{ Equation (4)}$$

Step Four:

For the next step, I develop a trading strategy for this relationship between the significant sets and stock returns. The strategy will be "the price changes of set A in period (t-x) will affect positively (negatively) to stock returns B". So we will need to buy or Sell stock B at period (t), if the set A rises or falls at period (t-x).

$$\text{E.g. } R_t = \beta_0 + \beta_1 r_{t-x}^{\beta_1} + \beta_2 r_{t-x}^{\beta_2} \text{ Equation (5)}$$

At last, the remaining commodities will leave a significant strategy.

For the trading strategy, let's assume element A and Stock B have a positive relationship, so if A increases in price in period (t-1) or another period, we invest in stock B in period (t) to get a positive return. But if A decreases in price, we sell stock B and invest into SSE constituent index.

To test this strategy, we need to set a "hold or change" Strategy, which means to invest an amount in the beginning and leave it as it is until the end, and see how much you will get. Comparing the above strategy with

"hold or change" strategy, the higher return strategy will be the better one.

If we have any significant result for this equation, that means we can exploit the relationship between the commodity and the stock, but that is all on a theoretical basis, in the real world, we need to consider transaction costs, to buy and sell the stock if there are any changes in element A at (t-1), we need to add up transaction costs for stock position changes. After subtract transaction costs from the return of the above strategy, we can compare it again to get the best result.

III. DISCUSSION

In this Discussion section, we analysis the data analysis and develop trading strategies. We test the data by using different time periods. And we use the most significant commodities to create trading strategies.

Data Analysis

In the first step, I have tested all commodities with the stock indexes in China, by using the equation:

$$R_t = \beta_0 + \beta_1 r_{t-x}^{\beta_1} + \beta_2 r_{t-x}^{\beta_2} + L + \beta_n r_{t-x}^{\beta_n} \text{ (6)}$$

Daily Data Test Result:

Here's the result for commodity futures price and SSE CONSTITUENT INDEX : (Exhibit 1-1, Appendix One)

$$R_{SSE} = -0.000314 + 0.221460 r_{t-1}^{\text{aluminum}} - 0.036956 r_{t-1}^{\text{copper}} + 0.092734 r_{t-1}^{\text{gold}} + 0.003806 r_{t-1}^{\text{lead}} + 0.029520 r_{t-1}^{\text{oil}} - 0.013422 r_{t-1}^{\text{silver}} - 0.032634 r_{t-1}^{\text{zinc}} \text{ (7)}$$

From the test result above, we can see that all the commodities are insignificant in the test for period (t-1) using daily data. We think the reason is the affection of the future price must be a longer time than just one day. So this time, I test all data with a (t-7) period. We have the following result:

SSE CONSTITUENT INDEX with (t-7) period (Exhibit 1-2, Appendix One)

$$R_{SSE} = -0.000589 - 0.111746 r_{t-7}^{\text{aluminum}} + 0.034722 r_{t-7}^{\text{copper}} + 0.151706 r_{t-7}^{\text{gold}} + 0.127218 r_{t-7}^{\text{lead}} - 0.069787 r_{t-7}^{\text{oil}} - 0.066070 r_{t-7}^{\text{silver}} - 0.071050 r_{t-7}^{\text{zinc}} \text{ (8)}$$

SSE CONSTITUENT INDEX with (t-30) period (Exhibit 1-3, Appendix One)

From the result, we see that gold is significant in result from (t-7) period test, but the other commodities are insignificant. Then we will test all date with a (t-30) period.

$$R_{SSE} = -0.000343 + 0.038452 r_{t-30}^{\text{aluminum}} - 0.033028 r_{t-30}^{\text{copper}} + 0.023879 r_{t-30}^{\text{gold}} + 0.135474 r_{t-30}^{\text{lead}} - 0.020562 r_{t-30}^{\text{oil}} - 0.089000 r_{t-30}^{\text{silver}} - 0.082289 r_{t-30}^{\text{zinc}} \text{ (9)}$$

From the result, we can see all the commodities are insignificant in the test for period (t-30) using daily data. So the 30 days may be much longer than the "time difference".

From all the test results above, we can see that most commodities are insignificant in the test except gold in the test (t-7) period. So we can create arbitrage trading strategies about gold, and the notion is that the gold price is not futures price. The trading strategy is about daily price, so we don't and it isn't necessary to test in weekly date.

IV. TRADING STRATEGY

To continue the analysis and create arbitrage trading strategy for gold prices, I select the one of the biggest

company in the gold industry in China. ZiJin mining group company limited.

ZiJin mining group company limited Founded in 1993 and publicly traded on the ShangHai stock exchange since 2003, and it is one of the China's leading gold companies. With a market capitalization of 26.2 Billion RMB. Headquartered in FuJian Province, China.

In this case, I select daily share prices for ZiJin from 2003 to 2014. And test them with the (t-1) and (t-10) period gold prices.

Gold price and ZiJin share price with (t-1) period (Exhibit 2-1, Appendix Two):

$$R_{ZiJin} = 0.000405 + 0.045097r_{t-1}^{Gold} \quad (10)$$

But in this result, gold is insignificant, so I will test the date with (t-7) period.

Gold price and ZiJin share price with (t-10) period (Exhibit 2-2, Appendix Two):

$$R_{ZiJin} = 0.000471 - 0.172408r_{t-10}^{Gold} \quad (11)$$

Here, you may find a question that the period is not (t-7), because we find that from (t-2) to (t-15), only at the point (t-10), gold is significant.

In this result, gold is significant and the slope of the regression is different with the slope of the regression between the gold and SSE constituent index, so that there is a chance to build an arbitrage trading strategy.

From these results, we have a question. Why is there a negative relationship between gold industry shares?

Most people would think there's a positive relationship between gold prices and gold industry shares, the higher the gold price, more value for the shares. But we find it doesn't happen like that.

Publicly traded gold mining firms produce a commodity output whose price is highly volatile. (P, Tufano, 1996) Under this situation, most gold industry companies have chosen futures or forward contracts to protect themselves from the price volatility.

Gold companies usually follow a policy of hedging a certain amount of their production over the short term (often three years). This provides income stream stability and allows them minimum debt coverage. (G, Twite, Dec 2002)

This phenomenon has made these gold industry companies sell their product in the future at a fixed price. So the higher the gold price is, the lower the profit they will make from these contracts, some companies may have an over hedging problem, so the price volatility will affect them more than others.

Hedging on products, it is common in the gold industry, and it is the reason for the negative relationship between gold prices and their shares.

From all the data we have above, we will create an arbitrage trading strategy for both ZiJin and SSE constituent index.

The trading strategy is that when r_{t-10}^{gold} is greater than zero, we will buy SSE constituent index at time t, otherwise we will buy ZiJin, you may ask in our regression equation, the period is (t-7) not (t-10) about

r^{ZiJin} , the reason is that we just want to find out a linear relationship.

I have created an Excel equation sheet to calculate the returns and the results.

And I show it on the next page, and there are not date of weekends and festive

The ZiJin Index and SSE Index represents the investment, 1 means invest, 0 means not. As I mention in methodology part, if gold price decrease in period t-10, then the stock price will increase in period t.

ZiJin return shows the stock price changes. I have set the transaction Inflation rate 10%/365, obviously, it is really a high rate.

The arbitrage trading strategy, we assume that we invest \$100 in SSE index at the date 2014/1/14, and by using equation:

$$=H12*(1+E4)*(1-10\%/365).$$

$$\text{Total Return} = \frac{H_{t+1} - H_t}{H_t} = H_t * \text{IF}(B_{t-9} > 0, (1+G_{t+1}), (1+B_{t+1})) * (1-10\%/365)$$

$$\text{ZiJin Return} = \frac{E_{t+1} - E_t}{E_t}$$

$$\text{Gold Return} = \frac{B_{t+1} - B_t}{B_t} \quad (11)$$

From all this trading strategy, we can see that it is significant after calculating with real world data, if we just invest in Gold or SSE index, we may lost a lot in the end, but our strategy can increase the return much, even the the whole market is not good and the inflation rate is really high. The relationship between gold prices and gold industry shares can be a solution to decrease the risk by hedging.

V. CONCLUSION

In this article, we have discussed the relationship between commodities and stock returns. And we also have run tests to verify significant results between them. We found that gold may have effect on SSE at (t-7) period, and we also show the relationship between gold price changes and the gold industry share price changes. And we have used this relationship to create an arbitrage trading strategy by SSE index and gold industry share price changes.

For most gold industry companies, to protect themselves from product price volatility, hedging is a common phenomenon. This phenomenon has made a negative relationship between gold prices and share prices. Even some gold industry companies have an over hedging problem. This negative relationship will affect them much more than others. And we also can see that the gold price changes may affect SSE index, because gold is one of the most important commodities.

	A	B	C	D	E	F	G	H		
1	Date	Gold return	SSE Index	Zijin Index	Zijin return	Inflation rate	SSE return	Total Return	SSE return	Gold total return
2	2003/12/29	0.45%			6.72%	10%/365	-1.04%	100	100	100
3	2003/12/30	0.52%			4.72%	10%/365	0.66%	100	100	100
4	2003/12/31	0.56%			1.65%	10%/365	-0.96%	100	100	100
5	2004/1/2	0.25%			1.63%	10%/365	0.92%	100	100	100
6	2004/1/5	0.36%			-1.60%	10%/365	3.47%	100	100	100
7	2004/1/6	1.41%			-1.63%	10%/365	1.01%	100	100	100
8	2004/1/7	0.24%			2.48%	10%/365	1.06%	100	100	100
9	2004/1/8	0.37%			0.00%	10%/365	1.52%	100	100	100
10	2004/1/9	0.17%			-3.23%	10%/365	-1.50%	100	100	100
11	2004/1/12	0.98%	1	0	-6.67%	10%/365	2.60%	102.5719	93.30776	102.5719
12	2004/1/13	0.03%	1	0	-4.46%	10%/365	0.45%	102.0781	89.11781	102.0781
13	2004/1/14	0.42%	1	0	8.41%	10%/365	1.48%	100.5348	96.58724	100.5348
14	2004/1/15	1.06%	1	0	-2.59%	10%/365	0.56%	99.9412	94.06351	99.9412
15	2004/1/16	2.54%	0	1	-2.65%	10%/365	0.85%	97.26408	91.54116	100.7685
16									
17	2014/2/7	0.74%	1	0	-2.38%	10%/365	0.33%	235.8024	14.32478	83.3883

18	2014/2/10	0.65%	1	0	-1.83%	10%/365	2.31%	241.1936	14.05889	85.29482
19	2014/2/11	0.57%	0	1	8.07%	10%/365	0.98%	260.5974	15.18992	86.10715
20	2014/2/12	0.35%	0	1	1.15%	10%/365	0.06%	263.5206	15.36031	86.13306
21	2014/2/13	0.04%	0	1	2.84%	10%/365	-0.43%	270.9327	15.79235	85.74227
22	2014/2/14	1.66%	1	0	0.55%	10%/365	0.70%	272.749	15.87525	86.31705
23	2014/2/17	1.34%	1	0	-0.55%	10%/365	0.56%	274.2067	15.7837	86.77838

Figure 1. The Gold return represents the percentage of gold price change.

The gold industry shares are not sensitive to the nearly daily changes in gold price, but are sensitive for longer periods, for example, (t-7) or (t-10), the continued increase or average increase will affect the share price, this situation will affect their profit from futures or forward contracts. The share price decrease somehow represents the market trend.

From these arbitrage trading strategies, we found the negative relationship between gold prices and gold industry shares, we can use gold prices in (t-10) to predict the returns on the t period gold industry share returns, otherwise we buy SSE index. It is significant in our test with real world data. So this could be another way to increase return and decrease risk.

There are many other commodities in the world, gold is the most commonly used. The relationship between commodities and stocks is complicated. There will be much more we can discover in the future. But overall, by using the results from this report, hedging risk by gold price changes could be considered and used.

ACKNOWLEDGMENT

This paper is a part of research from project of D/2013/01/111, Jiangsu education Commission and project of graduate textbook 'investment', Soochow University.

REFERENCES

- [1] Robert Raff and Howard Chan. "A multifactor model of gold industry stock returns: evidence from the Australian equity market." Applied Financial Economics, 1998, 8, 21-28.
- [2] Laurence E. Blose and Joseph C.P.Shieh. "The Impact of Gold Price on Gold Mining Stock." Review of Financial Economics, 1995, Vol.4 No.2, 125-139.
- [3] Gerben. Driesprong and Ben Jacobsen and Benjamin Maat. "Striking Oil: Another Puzzle?" 2005.
- [4] Schwert, G. William. "Anomalies and Market Efficiency", the Handbook of the Economics of Finance, 2003, Chapter 15, 939-972.

- [5] J.F.Jaffe, "Gold and Gold Stocks as Investment for Institutional Portfolios," *Financial Analysts Journal*, March/April 1989.
- [6] Chua, Jess H. Woodward, Richard S."Diversifying with Gold Stocks". *Financial Analysts Journal*, Jul/Aug90, Vol. 46 Issue 4, p76-79.
- [7] Twite, Carry. "Gold Prices, Exchange Rates, Gold Stocks and the Gold Premium." *Australian Journal of Management*, Dec2002, Vol. 27 Issue 2, p123.
- [8] Tufano, Peter. "The determinants of stock price exposure: Financial engineering and the gold mining industry." *Journal of Finance*, Jun98, Vol. 53 Issue 3, p1015
- [9] Tufano, Peter, "Who manages risk? An empirical analysis of risk management practices in the gold mining industry." *Journal of Finance* 41, P393-409
- [10] Hatirli S A, Ozkan B, Fert C. An econometric analysis of energy input-output in Turkish agriculture[J]. *Renewable and Sustainable Energy Reviews*, 2005, 9(6): 608