The Profitability of Crush Spread: Statistical Arbitrage Method

Quan Gu[*], Xinghui Lei

School of Economics and Management, Tongji University, Shanghai, 200092, China

*11guquan@tongji.edu.cn

Keywords: Spread trading, Cursh spread, GARCH model

Abstract: This paper examines the crush arbitrage of soybean and rapeseed futures, which is based on inter-commodity arbitrage in the oil processing process. The paper takes soybean, soybean meal and soybean oil futures listed on the Dalian Commodity Exchange, rapeseed, rapeseed meal and rapeseed oil futures listed on the Zhengzhou Commodity Exchange as research objects. Based on the statistical arbitrage model, empirical studies on two kinds of crush arbitrage are conducted. The improvement of this paper is to examine the characteristics of price volatility from the perspective of time-varying fluctuations, and then improve the setting of trading thresholds. The empirical results show that the use of GARCH model to solve the time-varying variance and set the feasibility of the threshold improvement, especially, rapeseed crushing arbitrage strategy portfolio returns better.

1. Introduction

Nowadays for oil processing enterprises, a more simple and effective mode of futures trading, which is based on oil processing process, can be less riskiness and more stable profitable. That iscrush spread arbitrage: buying (selling) oilseed futures, meanwhile selling (buying) plant oil and meal futures. It is similar to a spot of the production process, and therefore is called crush spread arbitrage. The current research is mainly focused on soybean futures, many studies are based on the soybean futures listed on the Chicago board of trade (CBOT)[1-3].These empirical studies are based on the idea of trend theory, and compare the different average moving averages to form trading signals for buying and selling. This method is simple, but it lacks rigorous theoretical arguments and is easily biased, thus affecting the empirical results of arbitrage. Some studies have also proposed improvements to apply the statistical arbitrage model to empirical research on crush arbitrage [4-8]. However, most of the theoretical premise of the study is that the cointegration regression residual obeys the normal distribution. This view is debatable. Therefore, it is necessary to introduce the GARCH model, describe the time-varying fluctuation characteristics of regression residuals, and set transaction thresholds based on this, and then use the statistical arbitrage model to improve crush arbitrage.

2. Method

2.1 The principle of crush arbitrage

The crush arbitrage is derived from the upstream and downstream production relations of plantoil processing, and there is a certain rule between soybean crushing arbitrage, soybean meal and soybean oil.

$$\text{Crush Profit} = M * \Phi_m + Y * \Phi_y - C - A \tag{1}$$

Among them, Crush Profit is the profits of plant oil processing,M, Y, A represent the price of soybean meal, soybean oil and soybean.$\Phi_m, \Phi_y$are the extraction rate and oil yield respectively. C represents the comprehensive cost of processing. When high spot crush profits, oil processing enterprises will tend to expand production and increase soybean demand, the soybean meal and soybean oil supply relative increase, this will lead to the relative soybean prices and soybean meal,

soybean oil prices relatively low. At this point, the futures market has the arbitrage opportunity to buy soybean meal and soybean oil, which is the forward squeeze arbitrage. On the other hand, when the spot crush profits are too low or even loss, the oil processing enterprises tend to reduce the production scale, and the demand for soybeans will be reduced. The supply of soybean meal and soybean oil, as a product, is also decreasing. This has led to lower soybean prices and higher prices for soybean meal and soybean oil. At this time, the futures market has the arbitrage opportunity to buy soybean meal and soybean oil and sell soybean, namely reverse squeeze arbitrage. In the futures statistical arbitrage model, the cost C represents the transaction cost and the cost of the deposit,Φ_m, $\Phi_y$often expresses in the form of co-integration coefficient.

2.2 Crush arbitrage modeling

Assuming thaty_t, x_{1t}, x_{2t}arethe price of the arbitrage composite futures contract,$\beta_0, \beta_1, \beta_2, e_t$The cointegration regression equation and its ductions are as follows:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + e_t \tag{2}$$

$$\text{spread}_t = y_t - \beta_1 x_{1t} - \beta_2 x_{2t} = \beta_0 + e_t \tag{3}$$

$$\text{Mean}(\text{spread}_t) = \text{Mean}(\beta_0 + e_t) = \beta_0 \tag{4}$$

$$e_t = \text{Mspread} = \text{spread}_t - \text{Mean}(\text{spread}_t) = y_t - \beta_0 - \beta_1 x_{1t} - \beta_2 x_{2t} \tag{5}$$

According to the volatility of the arbitrage portfolio residual time series with time-varying characteristics, can be used GARCH model to further volatility accurately depict the regression residuals sequences, thus improve transaction threshold setting. The regression of statistical arbitrage is followed by the perturbation term.$\{\varepsilon_t\}$In general, GARCH (p,q) model is established, and the conditional mean equation and conditional variance can be expressed as:

$$y_t = x_t' \cdot \gamma + \varepsilon_t \tag{6}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \cdot \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \cdot \sigma_{t-j}^2 \tag{7}$$

In the transaction practice, the above mean equation and variance equation of GARCH model are calculated, and the value of p and q is determined according to the actual effect of regression of the model.σ_t^2 can be achieved byε_t^2It is estimated that the transaction threshold will be determined.

2.3 Trading rules setting

By introducing the GARCH model, we can solve the residual difference of the centralization spread sequence, and then complete the transaction threshold setting, namely, the basic arbitrage trading strategy can be constructed:

1) $\text{Mspread} \geq k_1 \sigma_t$, it is considered that soybean (or rapeseed) futures is relatively overvalued, sell one soybean futures contract and buy at the same time. $\text{Mspread} \leq -k_1 \sigma_t$, it is believed that soybean futures are relatively undervalued and carry out the opposite operation.

2) When$-k_2 \sigma_t \leq \text{Mspread} \leq k_2 \sigma_t$, it is believed that the price difference is balanced to carry on the opposite transaction operation.

3) In order to control the transaction risk, avoid excessive fluctuation to cause loss, when$\text{Mspread} \geq k_3 \sigma_tor\text{Mspread} \leq -k_3 \sigma_t$,it means the fluctuation exceeds a certain range, the portfolio head is forced to liquidate.

Aboutk_1, k_2, k_3value, referring torelated research [9-11], the relevant transaction threshold is obtained as $\pm 0.75\sigma_t, \pm 0.05\sigma_t$ and $\pm 1.5\sigma_t$ respectively.

3. Results

This paper selects the continuous futures contracts of soybean, soybean meal, soybean oil, rapeseed, rapeseed meal, plant oil listed on the Dalian Commodity Exchange (DCE) and Zhengzhou Commodity Exchange (ZCE) prepared by Wind from January 4, 2013 to December 2016. The daily closing price of each species on the 30th of the month, excluding the suspended trading dates of

some varieties, a total of 971 samples were taken as sampling intervals, and a total of 483 samples from January 2013 to December 2014 were used as intra-sample training periods. 488 samples were taken as the out-of-sample test period. According to the model method proposed in the previous section, we performed the cointegration test and cointegration regression. The residual characteristics of the cointegration regression can be seen in Table 1.

Table 1 The statistical characteristics of the crush arbitrage spreads

Portfolio	AMY	RMO
Mean	0.0000	0.0000
Median	14.8980	9.9270
Maximum	258.1371	748.6297
Minimum	354.8456	350.5673
Standard deviation	121.3882	215.0499
Skewness	0.3564	0.4472
Kurtosis	2.6937	2.8090

Note: AMY, RMO stand for soybean and rapeseed arbitrage portfolio respectively.

We further perform the autocorrelation test on the above residual sequences. The relevant details are shown in Table 2. The autocorrelation coefficients of the lag order of the lag order of the residual sequence are all relatively high. Judging from the test result output map, the partial autocorrelation coefficient is from the first order. After that, there is no significant difference between 0 and 0. It can be initially judged that the residual sequence is a first-order autocorrelation process, AR (1), and the parameter estimation is performed.

Table 2 The autocorrelation test of the residual sequence of arbitrage

Lags	0	1	2	3	4	5	6	7	8	9	10
AMY	1	0.957	0.910	0.859	0.810	0.765	0.717	0.670	0.624	0.586	0.541
RMO	1	0.966	0.940	0.919	0.893	0.871	0.842	0.817	0.790	0.762	0.735

Table 3 The ARCH-LM test for the residual sequence of crushing arbitrage

Test indicators	AMY	RMO
F - statistic	9.0099	9.7332
Obs*R- squared	10.0314	18.7370
Prob. F (1422)	0.0021	0.0001
Prob.Chi-Square (1)	0.0014	0.0001

We then perform the ARCH effect test on the autocorrelation equation and adopt the ARCH-LM test method. The test results are shown in Table 3. From the results, it can be judged that the residual sequence rejects the null hypothesis. Therefore, consider applying the GARCH (1.1) model to the residual sequence. According to the above threshold setting method, the transaction is empirically tested. The relevant results are shown in Table 4.

Table 4 Performance of the crush arbitrage portfolios

Portfolio	AMY		RMO	
	CVM	TVM	CVM	TVM
Annual return	27.45%	36.94%	35.45%	22.94%
Sharpe ratio	0.545	0.356	0.752	0.770
Calmar ratio	0.452	0.852	0.452	0.336
Maximum withdrawal	36.25%	42.15%	36.42%	26.33%
Transactions	19	28	21	25
Average trading days	26	15	22	10

Note: CVM is an abbreviation of the constant volatility model, and TVM is an abbreviation of the time-varying volatility model.

4. Conclusion

The research results in this paper show that the time-varying threshold setting based on the residual fluctuation has a certain advantage over the traditional method, mainly reflected in the cumulative yield and the Calmar ratio. Using the GARCH model to improve the statistical arbitrage model and apply it to crush arbitrage is feasible. In addition, from the comparison of the empirical results of two crushing arbitrages, the crushing profit of soybean crushing is lower than that of rapeseed crushing, which is closely related to the popularity of arbitrage model application. There is still a wide space for the future application of rapeseed crushing arbitrage.

References

[1] Johnson, R. L., Zulauf, C. R., Irwin, S. H., & Gerlow, M. E. (1991). The soybean complex spread: an examination of market efficiency from the viewpoint of a production process. Journal of Futures Markets, 11(1), 25-37.

[2] Rechner, D., & Poitras, G. (1993). Putting on the crush: day trading the soybean complex spread. Journal of Futures Markets, 13(1), 61-75.

[3] Mitchell, J. B. (2010). Mitchell, j. b. (2010). Soybean futures crush spread arbitrage: trading strategies and market efficiency. Journal of risk and financial management, 3, 63-96.

[4] Fallahpour, S., Hakimian, H., Taheri, K., &Ramezanifar, E. (2016). Pairs trading strategy optimization using the reinforcement learning method: a cointegration approach. Soft Computing, 20(12), 5051-5066.

[5] Kim, S., &Heo, J. (2017). Time series regression-based pairs trading in the korean equities market. Journal of Experimental & Theoretical Artificial Intelligence, 29(4), 1-14.

[6] Laws, J. (2015). Trading and hedging the corn/ethanol crush spread using time-varying leverage and nonlinear models. European Journal of Finance, 21(4), 352-375.

[7] Yang, Y., Goncu, A., & Pantelous, A. (2016). Pairs trading with commodity futures: evidence from the chinese market. China Finance Review International, 7(3).

[8] Mark Cummins, & Andrea Bucca. (2012). Quantitative spread trading on crude oil and refined products markets. Quantitative Finance, 12(12), 1857-1875.

[9] Li, T. N., & Tourin, A. (2015). Optimal pairs trading with time-varying volatility. International Journal of Financial Engineering, 03(3), 1650023.

[10] Chen, C. W. S., & Lin, T. Y. (2016). Nonparametric tolerance limits for pair trading. Finance Research Letters, 21.

[11] Chen, C. W. S., Wang, Z., Sriboonchitta, S., & Lee, S. (2017). Pair trading based on quantile forecasting of smooth transition garch models. North American Journal of Economics & Finance, 39, 38-55.