Empirical Analysis and Related Strategies of Futures Commodity Cross – Variety Arbitrage

Yilin Li*

Shandong Weifang No.4 Middle School, Weifang, 261206, China

*Corresponding author. Email:18366382468@163.com

Abstract. Applied to futures market about rebar RB2305, coke J2305, iron IM2305, in three different types of products, meanwhile carrying on Multi-variety arbitrage, this essay will use some knowledge about econometrics, utilize futures market data analysis from about 2022.6.1 to 2023.4.27, which is carried out in areas such as rebar, iron and coke future price in many ways. Besides, it employs many inspection forms such as anisotropic variance test,T test and so on. Thus it can be proof of the future market which includes rebar ,coke ,iron cross-breed 2305 contract range arbitrage model and there is room for feasibility. It can foresee and determine the liner model relationship of the three parts, RB=910.4844+0.862298J+0.939140IM. On 2023. 4 .28, the three closing prices of rebar, coke, iron in the Contract 2401 respectively are 3583 yuan/ ton, 2174yuan/ton,862yuan/ton. If the prices of iron and coke agree with their inner price, then the estimated price of steel bar should be at 3596 yuan/ton, and there will be a gap of 13 yuan compared with the actual price, which means there will be profit margin and arbitrage can be proceeded. In this paper, empirical analysis of the legend is applied with Eviews9 and SPSS 26.

Keywords: rebar, coke, iron, linear regression, cross-variety arbitrage, factor analysis

1 Introduction

It is said that the black system goods(iron and so on)hedging transactions have some difficulty ,at the same time can also provide many nice chance for future means of exchange, for example many kinds of goods deal and cross-variety arbitrage deal. And“the aim of this paper is to investigate the causal relationship betweenfutures trading and spot market volatility in the Chinese commodity futures market [1].Author innovation point: about coke, iron, rebar three type of goods data use the way of linear regression to analyse, three parts have sharp changes .At the same time to write the essay reasonably to avoid many different types of the main contract delivery month is different, and it uses 2305 contract to discover the inner linear relationship ,and at the same time it can provide some knowledge about cross-variety arbitrage for people who use it and useful theoretical basis. This paper investigates the portfolio selection of cross-market steel futures contracts by focusing on the interdependence among the

B. K. Kandel et al. (eds.), *Proceedings of the 2023 3rd International Conference on Business Administration and Data Science (BADS 2023)*, Atlantis Highlights in Computer Sciences 19,
https://doi.org/10.2991/978-94-6463-326-9_53

futures of rebar, iron ore and coke, and in this paper we propose a risk-based model for the selection of cross-market arbitrage strategies and apply it to China's iron ore market, which includes three financial assets. iron ore, rebar, and coking coal [2,3].

2 Cross-variety arbitrage

2.1 Cross-variety arbitrage theoretical basis and summary

Cross-variety arbitrage is a basic interesting way in future market or finance market. Arbitrage includes easy deal and complex interest at the same time. This paper investigates the market efficiency of cross-commodity futures arbitrage strategy future goods cross-variety arbitrage is told about two types of future goods or many type of future goods in statistic and economic means which have some interest deal cross-variety arbitrage is the investor want to make profit in future market then to conduct two type of the goods or many type of goods have a high relationship to conduct a goods deal it use the characteristic of price variance, to make the transaction price gap[4]. In other words, it means buying in(sale out)in every delivery month contract at the same time and selling out(buy in)same delivery month and about another one type or many types contract, finally achieving the goal it shows in hedge closing or reverse closing that can make people get profit. Cross-variety arbitrage need to deal with all types of goods, have order and relevance. This paper investigates the cross-market efficiency of commodity futures markets, and examines the interdependence and information transmission among different futures markets, including cross-commodity futures markets [5]. For example, iron and coke are materials to manufacture steel bar. But, different type of future goods have different price in trend and their value is difference and it illustrates different conditions and data, illustrating different ways in increase and decline. Despite high relationship goods in many conditions their prices are similar and it has some different number too they have some deviation, and“This paper examines the hedge effectiveness of many type of commodity futures using many tests, and prove the potential benefits of cross-commodity hedging strategies.”[6].So it have many space to make profit and interest arbitrage chances, the main elements should consider about supply quantity and need situation ,repertory levels, macro policy and so on the basic elements.

2.2 Cross-variety arbitrage type and risks

Common cross-variety arbitrage strategies in the market include commodity arbitrage, rotation arbitrage, correlation arbitrage, time value arbitrage, and cross-market arbitrage. To successfully complete futures trading through arbitrage, market conditions need to be studied and analyzed with technical tools. Futures cross-variety arbitrage is characterized by low risk and high return, but it requires investors to have rich experience and high market sensitivity. When carrying out futures cross-variety arbitrage, investors should follow reasonable risk control principles, strictly control positions, and keep calm and rational [7].

3 Data choice

3.1 Data source

The data used in this paper are from the official websites of Dalian Futures Exchange and Shanghai Futures Exchange. RB2305, IM2305 and J23053 commodity futures contracts traded from June 1, 2022 to April 27, 2023 are selected. The reasons for choosing these three contracts are: (1) Because these three contracts are major contracts; (2) Because these three main contracts are concurrent, they are manifested in the same trading time (the distinction between day and night trading) and the same date interval (starting on June 1, 2022 and ending on April 27, 2023).

3.2 The elements effect rebar price

①iron price ②coke price ③other element, in econometric model use u to show that is the reason why we do not consider many type of problems and can not calculate elements to put them in interference factor, for instance the staff salary or insurance or monthly pay, yearly pay changes situations, price instability the price move up or down, inflation rate, finance market changes and politics element effects and so on.

4 Model regression analysis

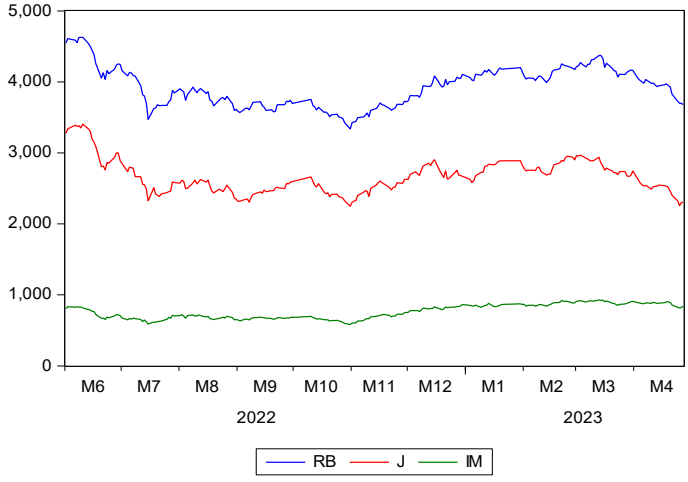

Fig. 1. RB, IM, J price line chart[owner-draw]

Use Eviews9 app made RB and J, IM line chart as shown in Figure 1. Figure1 shows we can see RB and IM, J all have the positive relationship in the figure, so, regression model should be: RB=β0+β1J+β2IM+u; u=random interference item. c, β1, β2—parameter to be estimated, RB=rebar closing price of the day(yuan/ton); IM=iron closing price of the day(yuan/ton); J=coke closing price of the day(yuan/ton). T=01Jun2022---

27Apr2023.there are weekend, Mid-autumn festival, lantern festival and other traditional festival and statutory holidays, do not have deal at these days. I uses the method of arithmetic sequence to deal with the price difference of adjacent trading day. This prevents post-holiday spikes and plunges from interfering with regression analysis.

5 Correlation model test

Assuming that the random disturbance item u in the econometric model satisfies the assumption, the least square method is used to estimate the parameters of the econometric model, using Eviews9 to regression analysis can get figure 1show.

The equation of the binary linear regression model is estimated as:

RB=910.4844+0.862298J+0.939140IM

t=(17.00923)(36.64761)(16.95705)

R^2=0.908818; F=1634.592; D.W.=0.084202

Table 1. Basic on RB The estimated result of futures price data[owner-draw]

Variable	Coefficient	Std.Error	t-Statistic	Prob		
C	910.4844	53.52884	17.00923	0.0000		
J	0.862298	0.023529	36.64761	0.0000		
IM	0.939140	0.055383	16.95705	0.0000		
R-squared		0.908818	Mean dependent var		3927.003	
Adjusted R-squared		0.908262	T. D. dependent var		278.5682	
S. E.of regression		84.37374	Akaike info criterion		11.71741	
Sum squared resid		2335008.	Schwarz criterion		11.75287	
Log likelihood		-1936.232	Hannan-Quinn criter		11.73116	
F-statistic		1634.592	Durbin-Watson stat		0.084202	
Prob(F-statistic)		0.000000				

Table 1 shows prob less than 0.05 means that it passes the 5% significance test, and the value in parentheses below the coefficient is the robust standard error. Dependent variable index for iron ore and coke futures prices. The test of coefficient of determination R^2 is 0.908818 > 0.8, indicating that the model fits the samples very well and has a high degree of fitting. The F test was output through the analysis of variance table, and the significance level was used to test whether the linear relationship of the regression equation was significant. Generally speaking, the significance level above 0.05 was significant. When the F test passes, it means that at least one regression coefficient in the equation is very significant, and the F test also passes the test. For H0: $\beta1=\beta2=0$, given the significance level about $\alpha=0.05$, F=1634.592, the original hypothesis H0: $\beta1=\beta2=0$ should be rejected, indicating that the regression equation is significant, that is, "iron ore price" and "coke price" and other variables association has a significant impact on the "rebar price"; D.w =0.084202 < 2, indicating that the prices of rebar, coke and iron ore will have an impact to the other type of element, but it not in dependent .

5.1 Economic means test

Figure 2 shows from the regression estimation results that iron ore prices, coke prices and rebar prices are linearly positive correlation, and that have a significant impact." Iron ore price "every 1% increase, rebar price increase 0.939140I%;" For every 1% increase in coke price, rebar price increases 0.862298%. The price elasticity of rebar to iron ore is greater, which is consistent with the nowadays economy significance theory means.

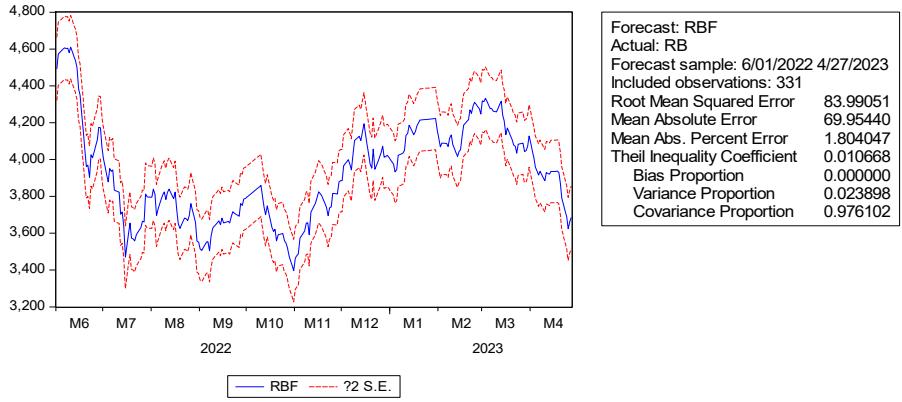

Fig. 2. Model forecast result[owner-draw]

Figure 2 shows can know, The solid line represents the predicted value of the dependent variable, and the two dashed lines above and below the solid line give the region of the can be trust band with a trust level of 95%. When CP values are large but BP and VP values are small, the result of prediction is very nice. As can be know from the figure, CP=0.976102, BP=0, VP=0.023898, indicating that the prediction is practical and the result in ideal.

5.2 Econometrics test

Test of heteroscedasticity. Using Eviews9 table 2 shows ordinary least square method to estimate the result: RB=95.18445+2.817727IM+1.174094J. It can be considered that the iron ore price changes affect the steel bar price changes to a greater extent. So if there is heteroscedasticity, that is because IM and J.

Table 2. RB and J regression result RB and IM regression result**[owner-draw]**

Variable	Coefficient	Std.Error	t-Statistic	Prob
C	1049.896	72.34935	14.51148	0.000000
J	1.078933	0.027027	39.92060	0.000000

S-squared	0.828882	Mean dependent var	3927.003
Adjust R-squared	0.828362	S. D.dependent var	278.5682
S.E.of regression	115.4085	Akaike info criterion	12.34086
Sum squared resid	4381994.	Schwarz criterion	12.36383

Log likelihood	-2040.412	Hannan-Quinn criter	12.35002
F-statistic	1593.654	Durbin-Watson Stat	0.062395
Prob(F-statistic)	0.000000		

Variable	coefficient	Std.Error	t-Statistic	Prob
C	2368.437	80.71425	29.34348	0.000000
IM	2.044168	0.104817	19.47363	0.000000

S-squared	0.535457	Mean dependent var	3927.003
Adjust R-squared	0.534045	S.D.dependent var	278.5682
S.E.of regression	190.1532	Akaike info criterion	13.33956
Sum squared resid	11896067	Schwarz criterion	13.36254
Log likelihood	-2205.697	Hannan-Quinn criter	13.34872
F-statistic	379.2223	Durbin-Watson Stat	0.018204
Prob(F-statistic)	0.000000		

unitary linear regression: RB=1049.896+1.078933J

(14.51) (39.92) R^2=0.83 D.W.=0.062 F=1593.65

The coefficient of determination R^2 test is 0.82 > 0.8, indicating that the model is higher degree of fitting. F test 1593.654 also passed the test, And D.W.=0.062 < 2, indicating that for rebar, coke prices between the two will have an impact, not independent of each other unitary linear regression: RB=2368.437+2.041168IM (29.34) (19.47) R^2=0.53 D.W.=0.018 F=379.22

The coefficient of determination R^2 test is 0.53＜0.8, It shows that the fitting effect of the model on the sample is general. F test 1593.654 also passed the test. And D.W.=0.018 < 2, indicating that for rebar, iron ore and iron ore will have an impact on the price of each other. The relationship between the element are not independent.

Using Eviews9 to make J and IM anisotropic variance test

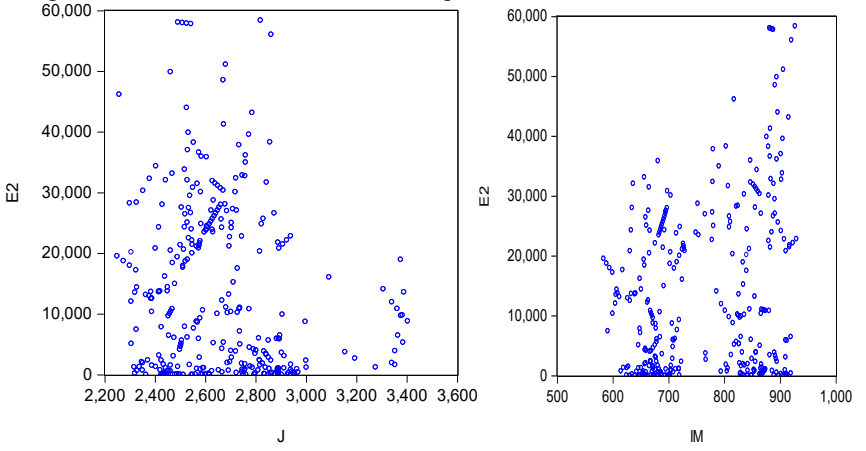

Fig. 3. Coke price and residual scatter plot iron price and residual scatter plot [owner-draw]

Figure 3 shows, it can be roughly seen that the middle and lower parts of the figure are divided into the main areas about distribution of the scatter points of residual square term e2 for explanatory variable J coke. E2 tends to increase first with the change of J,

so the model is more highly likely to have heteroscedasticity. In the same case, it can be seen from figure3 that the residual square term e2 mainly distributed in the middle part of the figure has a tendency to increase with the increase of IM variation first on the scatter plot of explanatory variable IM, so the model may have heteroscedasticity. Then we run another test the Gorieser test.

Table 3. Glesiser test result[owner-draw]

Variable	Coefficient	Std.Error	t-Statistic	Prob.
C	24.32870	11.70908	2.077763	0.0385
LOG(J)	-2.038354	1.484923	-1.372700	0.1708

R-squared	0.005695	Mean dependent var	8.256579
Adjust R-squared	0.002673	S.D.dependent var	2.310135
S.E.of regression	2.307046	Akaike info criterion	4.
Sum squared resid	1751.090	Schwarz criterion	4.538810
Log likelihood	-745.3710	Hannan-Quinn criter	4.525000
F-statistic	1.884306	Durbin-Watson Stat	0.423733
Prob(F-statistic)	0.170780		

Variable	Coefficient	Std.Error	t-Statistic	Prob.
C	-6.444496	6.366670	-1.012224	0.3122
LOG(IM)	2.217556	0.960180	2.309521	0.0215

R-squared	0.015954	Mean dependent var	8.256579
Adjust R-squared	0.012963	S.D.dependent var	2.310135
S.E.of regression	2.295113	Akaike info criterion	4.505465
Sum squared resid	-743.6545	Schwarz criterion	4.528439
Log likelihood	5.333889	Hannan-Quinn criter	4.514628
F-statistic	0.021534	Durbin-Watson Stat	0.427253
Prob(F-statistic)	0.021534		

Table 3 shows the index of dependent variable is the futures prices of iron ore and coke. In the T-test, when α=0.1, ta/2(n-k)=a/2(329)=2.5, log(IM) and log(J) are significant in the sense that p value is 0.05, so the IM of iron ore has heteroscedasticity, and the coke also has heteroscedasticity.

Factor analysis model:

$X = \mu + L F + e$, where X is the p x 1 vector of the measured value, μ is the p x 1 vector of the mean value, L is the p × m matrix of the load, F is the m × 1 vector of the common factor, and e is the p × 1 vector of the residual.

Factor analysis uses SPSS26, KMO >0.5, spherical Bartlett to test whether p value Less than 0.05;

KMO value is a statistic to measure the applicability of factor analysis. Its value ranges from 0 to 1, and the closer the value is to. Data is suitable for factor analysis. It is also believed that a KMO value greater than 0.5 indicates that the data is suitable for factor analysis. The spherical Bartlett test is used to test whether the correlation coefficient matrix of the data is suitable for factor analysis. If the p value is less than 0.05, the null hypothesis is rejected, indicating that the data is suitable for factor analysis. KMO and spherical Bartlett tests were performed on the data. As can be seen from the

next table, KMO value was 0.632>0.5 and p value was 0.000 < 0.05, so null hypothesis was rejected and the data was suitable for factor analysis (see Table 4).

Table 4. KMO and Bartlett's Test[owner-draw]

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.632
Bartlett's Test of Sphericity	Approx. Chi-Square	900.546
	df	3
	Sig.	.000

5.3 Principal factor analysis

According to explanatory value (eigenvalue) from largest to smallest order, according to the gravel plot and cumulative total variance interpretation, to determine the number of selected principal factors.

Table 5. Total Variance Explained[owner-draw]

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.467	82.235	82.235	2.467	82.235	82.235
2	.478	15.949	98.184			
3	.054	1.816	100.000			
Extraction Method: Principal Component Analysis.						

Table 5 shows in the characteristic value of one factor selected is 2.467>1, and its cumulative total variance interpretation rate reaches 82.235%

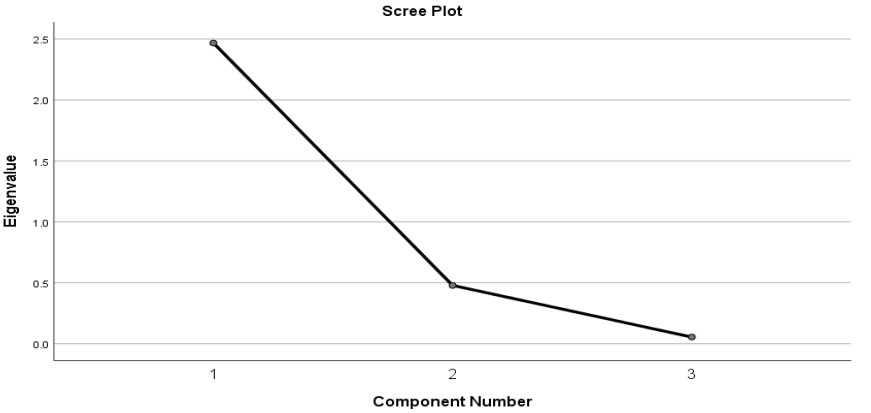

Fig. 4. scree plot[owner-draw]

5.4 Scree plot illustrate and conclusion

Figure 4 shows from the lithotripsy diagram, it can be seen that after the second main element, the lithotripsy soil gradually tends to be flat, indicating that first element is the most appropriate for the data. The explanation of the cumulative variance of the principal factors, and the explanation of the principal factors for each original variable 4. Composition of principal factors: rotated component matrix.

Table 6. Component Matrixa [owner-draw]

	Component
	1
RB price	.977
J price	.912
I price	.825
Extraction Method: Principal Component Analysis.	
a. 1 components extracted.	

5.5 Explanatory analysis of cumulative variance of principal components

Table 6 shows, this principal component explains the cumulative percentage of variance for all variables, which is 100% variance. This is because only one principal component has been extracted. Principal component interpretation for each of the original variables: As shown in Figure 3, this principal component has a high positive interpretation for all three original variables. Especially for RB price, J price and I price, the factor loads of main element were 0.977, 0.912 and 0.825, respectively. That means that the principal component explains most of the variance in the original variable, and that the original variable can be represented by a composite factor. In this situation, the main element can be interpreted as a "price factor."

5.6 Preservation principal factor

Figure 5 shows according to the preserved principal factor values, the analysis is carried out using finance. Calculate the price situation according to the above principal component, and make a time series analysis of it. the price has period move in October the price in minimum.

Fig. 5. Analysis of time series [owner-draw]

6 Conclusion

This paper elaborates views in a variety of methods. It utilizes linear regression model, F test, T test, logistic regression model, variance test, factor analysis and other methods as tools to study futures prices and arbitrage. Besides, empirical analysis is used to fully illustrate the views in this paper. The main factors affecting commodity prices include demand, cost and futures. Investment in futures is a good way and excellent choice for investors to allocate assets, which can also increase the value of assets and promote the development of futures market and financial market. At the same time, investors should consider futures investment reasonably and allocate arbitrage varieties reasonably for cross-variety arbitrage. Coke, iron ore, rebar futures price trend are also discussed in the essay. "In this paper, we investigate whether cross- commodity trading strategies exist in the U.S. energy market"[8]. At the same time“This paper compares the effectiveness of many type of futures hedging strategies , and examines the potential benefits of cross-commodity hedging”[9,10]. And in this paper,we,can know about many information about future, “the paper investigates the relationship between futures trading and information dissemination in the Chinese commodity futures market”[11,12].

Reference

1. Zhang, C., Li, L., & Ji, Q. (2019) Does futures trading increase spot market volatility? Evidence from the Chinese commodity futures market. Journal of Futures Markets, 39(6), 677-697.
2. Yu, J., & Yan, J. (2019)"Optimal Portfolio Selection of Cross-Market Steel Futures Based on Interdependence." The Journal of Business and Economic Studies (JBES), 25(1), 146-161.

3. Li, J., & Zeng, Q. (2019). "A Risk- Based Model for the Selection of Cross-Market Arbitrage Strategy: Evidence from China's lron Ore Market." Annals of Operations Research, 274(1-2)125-142.
4. Luo, D., Wang, J., Li, Q., & Zhang, L. (2018). "Cross-Commodity Futures Arbitrage and Market Efficiency." International Journal of Financial Research, 9(2), 203-214.
5. Li, H., & Tang, K. (2018) Futures cross-market efficiency, market interdependence, and information transmission: A multi-frequency approach. Journal of Futures Markets, 40(5), 837-859.
6. Xuan, H., Huang, X., & Zhang, W. (2020). Hedge effectiveness evaluation using VAR-GARCH models for crude oil and agricultural commodity futures. International Review of Economics & Finance, 65,78-95.
7. Wu, S.&Li, Y (2019) cross-variety arbitrage strategy journal of Northeast Aisa University of Finance and Economics, (4) 74-81.
8. Kim, S., & Sung, H. (2005). "Cross Commodity Trading Strategies in the U.S. Energy Market." Energy Economics, 27(2),195-222.
9. Gao, X., An, H., Wang, K., Ma, H., & Li, Y (2021). Comparative study of crude oil futures hedging strategies in the Asian and European markets. Energy Economics, 94,105023.
10. Ewald, C. O., Nawar, R., & Siu, T. K. (2011). Risk minimizing hedging of crude-oil options: theory and empirical performance. Ssrn Electronic Journal.
11. Guo, Y., & Wang, S. (2019). Futures trading and information dissemination: Evidence from the Chinese commodity futures market. Pacific-Basin Finance Journal, 57,101187.
12. Yang, Y, Goncu, A., Pantelous, A., Chen, G., & Zhou, C. (2017). Pairs trading with commodity futures: evidence from the chinese market. China Finance Review International, 00-00.