Original Research Article

Stock Index Futures in China; Hushen 300 Index

Authors:

1) Thian Cheng Lim

BEM department, Xi’an Jiaotong-Liverpool University

111 Ren’ai Road, Dushu Lake Higher Education Town, Suzhou Industrial Park 215123, China

2) Xue Feng

Swansea University

3) Zhengjiang Hu

National University of Singapore
Abstract

This paper examines the arbitrage-induced in regimes (upper, inner and lower regime) price dynamics between Hushen 300 index spot and futures markets using a threshold cointegration analysis and Error Correction Model (ECM) on china stock market index. The studies are carried out from 426 observations with samples selected from 04/16/2010 to 03/16/2012. We are interested to know what extent of mispricing would represent for profitable arbitrage opportunity. Futures prices would have a fluctuating effect within lower and upper thresholds by the factors of transaction costs and arbitrage risks, but no profitable arbitrage opportunities within this area.

Keywords: Chinese Stock; Index Futures Market; Intraday; Cointegration Analysis; Error Correction Model (ECM)

I. INTRODUCTION

China officially started trading on its stock index in April 16, 2010. China Shanghai Shenzhen 300 Stock Index Futures, is often abbreviated to "Hushen 300 Index" in an effort to push the reform of capital market. The Hushen 300 tracked index of 300 Shanghai- and Shenzhen listed class A-shares and is on China Financial Futures Exchange (CFFEX) (Reuters, 2010). The reason for trading in stock index futures is price discovery and risk management although a stock index futures can also be used for speculation. China has no major financial derivative products, leaving investors without hedging tools and limiting the country's pricing power in the international arena.

This study is an empirical study following deductive logical, investigates the no-arbitrage pricing band and price dynamics on the upper and lower regimes in Hushen 300. From the graph below (Graph 1), we can roughly see that the prices between the spot market and the future market are cointegrated together which gave us an opportunity to find mispricing for arbitrage. Error correction model (ECM) is the classic model for analyzing the relationship between spot and futures. However, it assumes no transaction costs and arbitrage risks, which gives arbitrage a frictionless process. From empirical evidence, index futures price are not continuously moving with the cost of carry benchmark. Of course, when the arbitrage profit is less than transaction costs, the arbitrageur will not enter into the market. Therefore, it implies that the arbitrage opportunity is not continuous. Combined with threshold vector error correction model (TVECM), we can overcome the limitation of applying ECM if the mispricing process is in fact discontinuous through Tsay Test (Tsai, 1989).

In this paper, a multivariate threshold error correction model is used to study the regimes (upper, inner and lower regime) price dynamic between the Hushen 300 spot and futures prices. In the inner regime, the mispricing is not large enough to make profitable arbitrage (the "no-arbitrage window"). Therefore, arbitrageur will not enter into the market to profit from the dynamic relationship between futures and spot prices. The causes including transactions costs, capital constraints, execution risk, interest rate and dividend risks could also allow futures prices to fluctuate without a band, we will examine how much the future price is mispriced. By using the percentage mispricing as the threshold variable, it
helps us to identify three regimes of arbitrage band.

II. LITERATURE REVIEW

The most famous model for pricing stock index futures is undoubtedly the cost of carry model by Cornell and French. The cost of carry model implied that a pair of spot and futures should be cointegrated in the long run. They assume the capital markets are perfect. That means there is no taxes and transaction costs; no short selling restrictions; and the assets can be divided infinitely. They derived the general stock index futures pricing formula which base on no dividend assumption. After the empirical research on S&P500 index futures, they found that —the stock index futures prices are generally below the level predicted by simple arbitrage models. And this difference between the actual and predicted prices is caused by taxes. (Cornell & French, 1983).

However, various studies working on cost-of-carry have found that this model consistently produces pricing errors with respect to the actual price. As to the mispricing puzzle, many literatures have proposed the explanations: by using S&P 500 stock index, Cornell and French (1983b) empirically observe that actual futures prices are below the corresponding values predicted by the cost of carry model and suggest this overestimation is caused by timing option; nevertheless, Cornell (1985) defines this conclusion and indicates that the transaction costs and other factors may lead to mispricing as well; Modest and Sundaresan (1983), Modest (1984) and Klemkosky and Lee (1991) then take transaction costs into account but mispricing still exists; Besides, Stoll and Whaley (1990) document the interdependence between stock index spot and future market and therefore the assumption of cost-of-carry is therefore biased and cause mispricing; what’s more, Cox, Ingersoll, and Ross (1981), Richard and Sundaresan (1981) and French (1982) examine the theoretical difference between forward and futures prices in a variety of contexts and suggest this may lead to mispricing, however, Cornell and Reinganum (1981) and Elton, Gruber and Rentzler (1983) using simulation and empirical test show that the difference is economically insignificant; In addition, evidence in Resnick and Hennigar (1983), Kamara (1988) suggest that that the pricing deviations from cost-of-carry are related to interest rate or the volatility of the underlying security markets in some different financial markets. Motivated by this consideration as well as the unrealistic assumption behind the cost-of-carry, Helmer and Longstaff (1991) develop a closed-form general equilibrium model of stock index future prices in a continuous-time economy characterized by interest rate and market volatility. Their empirical result is consistent with their general equilibrium model but not the cost-of-carry. Considering the market imperfection and the degree of limitation on arbitrage, Hsu and Wang (2004) propose another general equilibrium pricing model of stock in imperfect markets. Hsu and Wang (2006) find Hsu-Wang model provides the best performance among others in Taiwan market, but it shows no improvement in US market.

\[ F_t^* = S_t e^{(r-d)(T-t)} \]

\[ X_t = \frac{F_t - F_t^*}{S_t} \]
At first, Engle and Granger (1987) used the error correction model to examine price dynamics between spot and futures markets. The Engle–Granger method involves first the running regression of one variable on another, and second checking whether the regression residual from the first step is stationary using, say, an ADF test. In this sense, the Engle–Granger method is largely the unit root test and will not be deliberated either. Then, Brenner and Kroner (1995) modified this method. Even though there is also limitation with the standard linear error correction model, which is that it assumes that the arbitrage possibilities which creates the tendency to move towards equilibrium are present in every time period, the arbitrage-induced error correction effect is allowed to be inactive inside the no-arbitrage window and only becomes active when the system gets sufficiently far from equilibrium (Balke and Fomby 1997).

The Engle–Granger approach developed for linear time series models which turns out standard tests for detecting cointegration in linear time series are also capable of detecting threshold cointegration. Then, Balke and Fomby (1997) introduced the threshold cointegration which combines the concept of cointegration and the nonlinear error correction effect which should be discontinuous.

The width of the no-arbitrage band can be tested by using substantive transaction costs information. For example, Yadav and Pope (1990, 1994) and Butterworth and Holmes (2000) have studied the mispricing of index futures contracts and the implied arbitrage profitability in the UK market by comparing mispricing with round-trip transaction costs faced by different categories of index arbitrageurs. However, these subjectively estimated transaction costs may not be representative of the true average transaction costs for most arbitrageurs. Therefore, in this study, we use an idea of Mackinlay and Ramaswamy (1998), and introduce the percentage mispricing as the threshold variable in the threshold error correction model. The threshold values yielded by estimating the threshold error correction model should therefore reflect the average transaction costs faced by most arbitrageurs. We believed that the transaction costs implied from estimating the threshold error correction model are more reliable and fair than subjective estimation in reflecting the average transaction costs for most arbitrageurs because they are based on the observed market pattern that the price dynamics switch between regimes depending on the presence or absence of arbitrage activities. Furthermore, the substantive information about transaction costs is the only used to specify a reasonable range for searching the threshold values.

III. MAIN METHODOLOGIES

A. Cointegration Tests

Cointegration is one of the most important developments in time series econometrics in the last quarter-century. A group of non-stationary I(1) time series is said to have cointegration relationships if a certain linear combination of these time series is stationary (Wang, 2001). If two or more I (1) series are cointegrated, there exists a linear combination of these series that is I(0). Brenner and Koroner (1995) argue that according to the cost of carry model, if interest rates have a stochastic trend, spot and futures prices will not be cointegrated by themselves, and the differential should be included in the
cointegrating vector.

There are two major approaches to testing for cointegration, the Engle–Granger two-step method (Engle and Granger 1987) and the Johansen procedure (Johansen 1988, 1991; Johansen and Juselius 1990). The Johansen procedure which is to test the restrictions imposed by cointegration on a vector autoregression (VAR) model:

\[ Y_t = u + A_1 Y_{t-1} + ... + A_p Y_{t-p} + \varepsilon_t \]  

(3.1.1)

Where \( Y_t \) is a \( k \)-dimension vector of variables which are assumed to be \( I(1) \) series (but can also be \( I(0) \)), \( A_i, i = 1 \ldots p \) is the coefficient matrix, and \( \varepsilon_t \) is a \( k \)-dimension vector of residuals. Subtracting \( Y_{t-1} \) from both sides of equation (1) yields:

\[ \Delta Y_t = u + \Pi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + ... + \Gamma_{p-1} \Delta Y_{t-p-1} + \varepsilon_t \]  

(3.1.2)

Where

\[ \Pi = \sum_{i=1}^{p} A_i - I \]

and

\[ \Gamma_i = - \sum_{j=i+1}^{p} A_j \]

We can observe from equation (2) that only one term in the equation, \( \Pi Y_{t-1} \) is in levels, cointegration relations depend crucially on the property of matrix \( \Pi \). It is clear that \( \Pi Y_{t-1} \) must be either \( I(0) \) or zero except that \( Y_t \) is already stationary.

**B. An Error Correction Model**

Before building the error correction model, we need to do the cointegration analysis for futures and spot variable. If there is a long run cointegration relationship, ‘u’ can be an explained variable with other variables to build a typical regression model.

\[ \Delta Y_t = \text{lagged} (\Delta Y, \Delta X) - \lambda u_{t-1} + \varepsilon_t \]

*Engle-Granger two steps to ECM*

1. Test the cointegration relationship and estimate the cointegration vector.

2. Put the residual \( u \) as error term into ECM, and then estimate parameters by OLS.

http://www.ijmsbr.com
In step 2, we should use lagged residual term to omit the unit root.

Cointegration between futures and spot prices has significant implications for modeling the dynamics of individual series since deviations from the long-run equilibrium relationship may affect subsequent price movements in either market.

Departure of individual variables in the cointegration vectors from the equilibrium will be subsequently reversed back to the equilibrium -- a dynamic adjustment process called error correction mechanism (ECM).

C. Tsay test for nonlinearity

Tsay test can be used as a diagnostic tool for building linear or nonlinear time series models. This idea comes from Tukey’s (1949) nonadditivity test simply uses the aggregated quantity $\hat{Y}_t^2$, the square of the fitted value of $Y_t$ based on the entertained linear model, to obtain quadratic terms upon which the residuals can be correlated.

**Step 1:** Regress $Y_t$ on $\{Y_{t-1}, Y_{t-M}\}$ by least squares and obtain the residuals $\{\hat{e}_t\}$, for $t=M+1, \ldots, n$. The regression model will be denote by

$$Y_t = W_t \Phi + e_t$$

(3.3.1)

Where $W_t = \{Y_{t-1}, \ldots, Y_{t-M}\}$ and $\Phi = (\Phi_0, \Phi_1, \ldots, \Phi_M)^T$ with $M$ being a prespecified positive integer, $n$ the sample size, and the superscript $T$ denoting the matrix transpose.

**Step 2:** Regress the vector $Z_t$ on $\{Y_{t-1}, \ldots, Y_{t-M}\}$ and obtain the residual vector $\{\hat{X}_t\}$, for $t=M+1, \ldots, n$. Here the multivariate regression model is

$$Z_t = W_t H + X_t$$

Where $Z_t$ is an $m=(1/2)M(M+1)$ dimensional vector defined by $Z_t^T = vech(U_t^T U_t^*)$ with $U_t = (Y_{t-1}, \ldots, Y_{t-M})$ and vech denoting the half stacking vector. In other words, $Z_t^T$ is obtained from the symmetric matrix $U_t^T U_t$ by the usual column stacking operator but using only those elements on or below the main diagonal of each column.
Step 3: Regress \( \hat{e}_t \) on \( \hat{X}_t \) and let \( \hat{F}_t \) be the ratio of the mean square of regression to the mean square of error. That is, fit

\[
\hat{e}_t = \hat{X}_t \beta + \epsilon_t \quad (t=\text{M+1}, \ldots, n) \tag{3.3.2}
\]

and define

\[
\hat{F} = \left( \sum \hat{X}_t \hat{e}_t \right) \left( \sum \hat{X}_t^T \hat{X}_t \right)^{-1} \left( \sum \hat{X}_t^T \hat{e}_t \right) / \left( \sum \hat{e}_t^2 / (n-M-m-1) \right) \tag{3.3.3}
\]

where the summations are over \( t \) from \( \text{M+1} \) to \( n \) and \( \hat{e}_t \) is the least squares residual for (3.3.2).

Obviously, the regression (3.3.2) one can easily identify the significant nonlinear terms to be incorporated in the model.

IV. DATA

The contracts of futures in Hushen 300 index are traded on the China futures exchanges. We choose the closing price as the price for each day due to futures contracts are marking to market. Also daily resettlement gave us strong support for using daily data. One month maturity futures are traded frequently, so we choose 426 one month maturity observations since the first day of this contract traded on the market. We define \( F(t) \), \( S(t) \) are futures and spot prices respectively. \( f_t = \ln(F_t), \ s_t = \ln(S_t) \), and the basis is \( b(t) = f_t - s_t \). Daily returns are \( \Delta k_f = f_{t+1} - f_t \), \( \Delta k_s = s_{t+1} - s_t \), \( r(t) \) is the risk free interest rate, \( d(t) \) is the index dividend yield and \( (r-d)(T-t) \) is the cost of carry where \( T \) is the expiration date of the contract and \( t \) is the current date.

V. EMPIRICAL RESULTS

Since the basis followed non–linearity threshold process, TVECM is a relatively appropriate model to simulate spot and future price dynamics. Then, we should consider how many regimes we will take. According to Tsay (1989), we choose three regimes: \( x_{t-d} < C_1 \) \( x_{t-d} > C_2 \) \( C_1 < x_{t-d} < C_2 \). In the inner regime, the mispricing is too small, not significant for arbitrageurs to obtain profit. In the upper and lower regime, mispricing is largely move at negative or positive direction. A three regime TVECM can be instructed base on VECM as:

\[
\Delta S_t = a_0^{(r)} + \sum_{j=1}^{n} a_{1j}^{(r)} \Delta S_{t-j} + \sum_{j=1}^{n} a_{2j}^{(c)} \Delta F_{t-j} + a_{3j}^{(r)} b_t + e_{1t}^{(r)} \tag{5.a}
\]

\[
\Delta F_t = \beta_0^{(r)} + \sum_{j=1}^{n} \beta_{1j}^{(r)} \Delta S_{t-j} + \sum_{j=1}^{n} \beta_{2j}^{(c)} \Delta F_{t-j} + \beta_{3j}^{(r)} b_t + e_{2t}^{(r)} \tag{5.b}
\]
Notes: $r$ means the different regimes: $r=1$ when $x_{r-d} < C_1$, $r=2$ when $C_1 < x_{r-d} < C_2$, $r=3$ when $x_{r-d} > C_2$.

Our main interest is to specify the lower and upper thresholds are located. These ranges of the thresholds are based on past experience and real world information. We assume the lower threshold range of $C_1$ is (-0.75, -0.05) and the upper threshold range of $C_2$ is (0.05, 0.75) based on three arguments. Firstly, Yadav and Pope (1994) present two levels of transaction costs in the period of 1986 - 1990, 0.25% and 0.75% based on the two types of arbitrageurs. One is arbitrageurs who are not subject to transaction taxes for arbitrage dealing, like market makers, those who are otherwise committed to enter the market (due to portfolio insurance or tactical price based strategies) and use the futures market only as an intermediary and those with existing arbitrage positions who seek to profitably unwind early. Another is arbitrageurs who have to pay transaction taxes in their transaction process. We can predict the transaction cost is lower now than it in older days since the quote – driven dealer system was introduced which saves cost through buy and sell orders are automatically reported and matched by the electronic order book. Secondly, Yadav and Pope stated the arbitrage band should depend on arbitrageurs with the lower transaction costs (Yadav and Pope, 1990). Lastly, In the Hushen 300 future market, we searched the transaction cost is about 0.3%, which including transaction taxes 0.1%, transaction fee 1.5% and market impact cost 0.5%.

Through the grid search method, we choose the thresholds with sum of squares residual (RSS) as the criterion (Tsay, 1998). Using 200-point grid, we selected $C_1= -0.512$ and $C_2=0.218$. In the process of test, $C_1$ and $C_2$ are considered to be wide enough to count for most active arbitrageurs in the market. Then we set $C_1$ is (-1.0, -0.05) and $C_2$ is (0.05, 1.0) to select the threshold value (-0.981, 0.463) by using 200 points grid search. However, the error term (SSR) is significant in the regression equation of TVECM with the estimated inner regime of (-0.981, 0.463). It means the estimated inner regime is not the inner regime value which arbitrage activities should not be active. Also, the true threshold values should be close to zero. This mistake may be due to some other noisy reasons that have the same effect of causing regime – different spot and future price dynamics. In our test, we only control the range of $C_1$ and $C_2$. Then, we try to narrow down the range to $C_1$ belongs to (-0.75, -0.05) and $C_2$ belong to (0.05, 0.75) until the equation satisfied the standard of no error correction effect in the inner regime.

VI. CONCLUSION

The band we calculated is narrower than one which we expected from the observable information about transaction costs given. It implies that the non-cost factors are not significant in determining the width of the band, for example, risks and trading constraints. We can see from results the lower threshold value is higher than the upper threshold value in absolute value. It suggests that transaction cost in the lower regime is higher than in the upper regime. This discovery is consistent with the statement of Mackinlay and Ramaswamy in 1988. They displayed in the lower regime condition where
futures are underpriced with spot results, arbitrage investor involves buying futures and short selling spot. It happens arbitrage incentives in this condition easily. Therefore, the higher transaction costs associated with arbitrage in the lower regime are consistent with the extra constraints and costs imposed on arbitrage by short sale restrictions in the spot market. Future – spot arbitrage opportunity is large in the market which immature and regulation is not complete. We concluded that with more efficiency in the futures market, the opportunity of arbitrage is narrowed.

VII. APPENDIX

<table>
<thead>
<tr>
<th>Index</th>
<th>Australia</th>
<th>Mainland China</th>
<th>Hoag Kong</th>
<th>India</th>
<th>Japan</th>
<th>Korea</th>
<th>Malaysia</th>
<th>Singapore</th>
<th>Taiwan</th>
<th>No. of markets offering the product type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Futures</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>8</td>
</tr>
<tr>
<td>Options</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>8</td>
</tr>
<tr>
<td>Options on futures</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>3</td>
</tr>
<tr>
<td>Stock</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>4</td>
</tr>
<tr>
<td>Options</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>6</td>
</tr>
<tr>
<td>Currency</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>6</td>
</tr>
<tr>
<td>Options</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>3</td>
</tr>
<tr>
<td>Interest rate</td>
<td>Futures</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>3</td>
</tr>
<tr>
<td>Options on futures</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>3</td>
</tr>
<tr>
<td>Bond</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>6</td>
</tr>
<tr>
<td>Options on futures</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>3</td>
</tr>
<tr>
<td>Commodities</td>
<td>Futures</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>7</td>
</tr>
<tr>
<td>Options on futures</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>7</td>
</tr>
<tr>
<td>No. of product types in each market</td>
<td>12</td>
<td>1</td>
<td>6</td>
<td>4</td>
<td>10</td>
<td>9</td>
<td>5</td>
<td>9</td>
<td>5</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 1 Derivatives traded on Asian market
REFERENCES


