Study on Stock Index Futures’ Mean Reversion Effect and Arbitrage in China Based on High-Frequency Data

Wei Zhuo¹, Xiujuan Zhao², Zhou Zhou¹, Shouyang Wang³

¹School of Management, Graduate University of Chinese Academy of Sciences, Beijing, China; ²School of Economics and Management, Beijing University of Posts and Telecommunication, Beijing, China; ³Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China.
Email: xjzhao@iss.ac.cn

Received January 5th, 2012; revised February 19th, 2012; accepted February 29th, 2012

ABSTRACT

Based on 1 minute high frequency data, this paper constructs no-arbitrage band for CSI300 index futures, and empirically studies the futures-spot arbitrage. Furthermore, the mean reversion and its time effect are analyzed by ADF model for the first time in Chinese index futures market and a logit model is used to investigate the related factors of arbitrage opportunities. We find that CSI300 index futures have cash-and-carry arbitrage opportunities; the existence of mean reversion effect of index futures when arbitrage is absent; cash-and-carry arbitrage has a significant impact for mean reversion of the mispricing, while the reverse has insignificant effect. As for the time effect of mean reversion, it indicates that the time which arbitrage effect reversion most is mispricing shows 14 minutes later. The probability of arbitrage opportunities is positive correlated to first-order lagged of volume, and negative correlated to the volume; the time to expiration, the highest price and lowest price differences are also positive related to arbitrage opportunities.

Keywords: CSI 300 Index Future; High-Frequency Data; Futures-Spot Arbitrage; Mean Reversion Effect; Mispricing Ratio

1. Introduction

Since Chinese first index futures launched on China Financial Futures Exchange (CFFE), the trading volume jumped to the biggest among all futures. Besides price discovery and hedging, the function of index futures includes arbitrage, which makes a stable relationship between spot and futures market. Generally, arbitrage on index futures can divide into three types: futures-spot arbitrage, inter-delivery arbitrage, and cross-market arbitrage. Future-spot arbitrage opportunity arises when index futures deviates its theoretical value which is calculated by cost of carry theory. Because of transaction cost, index futures normally would stay in a band which arbitrage trade cannot make profit and we called “no-arbitrage band”. When futures’ price stays out of the band, index futures are mispricing. Then, arbitrageur will obtain profit by cash-and-carry arbitrage, which means they short index futures and long spot when futures price is above the upper limit of no-arbitrage band, and wind the position of both on expiration day. If index futures exceeds lower limit, they could do the opposite actions of above which named reverse cash-and-carry-arbitrage.

The research about index futures’ arbitrage has been popular. Cornell [1] proposed the cost of carry theory, and compared actual and model price of NYSE composite index futures and S & P500 index futures, which is the milestone for arbitrage research. Figlewski [2] analyzed several S & P500 index futures contracts statistically, and found there were many arbitrage opportunities indeed. Besides, Harris [3], Brennan [4], Chung [5], and Richie [6] made some similar conclusions. Fremault [7] discovered three effects of arbitrage trading for hedgers, speculators and arbitrageurs in futures and spot market. Sutcliffe [8] claimed that it is important to separate arbitrage opportunities from mispricing. McMillan [9] investigated mispricing of ISE30 index futures in Turkey, and found there was only cash-and-carry arbitrage opportunities when spot cannot be shorten.

Since profits cannot be obtained unless index futures return to no-arbitrage band, then the research has been involved into mean reversion effect of index futures. MacKinlay and Ramaswamy [10] discovered mean reversion effect existed in S & P500 index futures. Additionally, Lim [11] and Yadav [12] found Nikkei225 and FTSE-100 had this effect. Neal [13] employed logit regression model using one minute data to analyze arbitrage trades on the NYSE for first contract in 1989, who found that there mispricing reversal is significant positive related to absolute mispricing amount.
Then, some researches focused on explanation of mean reversion effect of index futures. Traditionally, reversion effect is viewed as a consequence of arbitrage in market. Garbade and Silber [14] used vector autoregression and found continuum of traders induces continuity in mean reversion. Chan [15], Kawaller [16], Stoll [17] have estimated similar model and obtain the same conclusion. Kawaller [18] and Tse [19] considered the effect was resulted from different arbitrageurs, and the reversion is a smooth auto-regression process. Yadav and Pope [20] suggested transaction was the trigger of reversion. Kempf [21] made the same conclusion by discussing the influence of short restriction and clear the position before expiration day. Fung [22] observed S & P500 index futures had price reversal phenomenon which due to the investors’ overreaction for information. Kumar [23], Holden [24-25] had some similar conclusion with arbitrage trade reason.

However, the liquidity theory challenges the traditional explanation, which argues the illiquidity in stock market cause the reversion of futures. One of the earliest illiquidity explanations was Kraus and Stoll [26], who found that large trade would bring about reversal price, which is consistent with Biais [27]. Alphonse [29] accounted that large trade would bring about reversal price, which is sensitive to threshold’s value, and impulse response resulted from different arbitrageurs, and the reversion is a conseq uence of arbitrage in market.

St is futures and spot price at t time. Cst and Cft is the spot pricing ratio, we can short index futures, D t,2 = 1; otherwise, negative mispricing ratio makes D t,3 = 1. In addition, p lagged mispricing changes also included in the regression equation to correct for autocorrelation in mispricing ratio. As for D t,i, if there is no mispricing and arbitrage opportunity at t time, D t,1 = 1. If there are positive mispricing ratio, we can short index futures, D t,2 = 1; otherwise, negative mispricing ratio makes D t,3 = 1. In addition, p lagged mispricing changes also included in the regression equation to correct for autocorrelation in mispricing changes. In order to investigate the higher order difference, we also estimated functions as following.

\[
\Delta \text{MR}_t = c + \sum_{i=1}^{3} \alpha_i D_{t,i} \Delta \text{MR}_{t-1} + \sum_{j=1}^{p} \beta_j \left( \Delta \text{MR}_{t-j} \right)_t + \epsilon_t
\]

where, \(\Delta \text{MR}_t\) is the first order difference of mispricing ratio. As for D t,i, if there is no mispricing and arbitrage opportunity at t time, D t,1 = 1. If there are positive mispricing ratio, we can short index futures, D t,2 = 1; otherwise, negative mispricing ratio makes D t,3 = 1. In addition, p lagged mispricing changes also included in the regression equation to correct for autocorrelation in mispricing changes.

Finally, we estimated a logit model to test for some significance variables that may influence arbitrage opportunity including the expiration days, trading volumes, positions and spread between highest and lowest price in one minute of index futures.

Based on 1 minute high frequency data, this paper studies the arbitrage and mean reversion effect for CSI300 index futures. Since there are four contacts on the same time, we construct the continuous contract with the most active. The CSI300 index futures market open at 9:15, and close at 15:15, which is earlier and later 15 minutes than its underlying index, respectively. Therefore, the data we collected in 9:30 to 11:30 and 13:00 to 15:00 for the period from 23rd August 2010 to 18th February 2011.

As for transaction part, the commission for spot Cst is RMB30 for each deal, and price multiplier for each point

\[F_{\text{TT}}\] is theoretical price for index futures at T time. \(F_i\) and \(S_i\) is futures and spot price at t time. \(C_u\) and \(C_d\) is the spot commission and futures transactions. \(C_u\) and \(C_d\) is impact price for longing and shorting spot, while \(C_u\) and \(C_d\) is impact price for long and shorting futures, respectively, and \(r\) is riskless interest rate.

Once the price exceeds no-arbitrage band, index futures is mispricing and we named mispricing ratio (MR) to value as following.

\[
\text{MR} = \begin{cases} \frac{F_i - F_u}{F_u} & (F_i > F_u) \\ 0 & (F_d < F_i < F_u) \\ \frac{F_i - F_d}{F_d} & (F_i > F_d) \end{cases}
\]

Referring Mackinlay’s [10] research which first and higher orders of autocorrelation coefficient show the persistence of mispricing, we estimated ADF functions as following.
of RMB 300, so the transaction cost for index futures \( C_{it} \) is \( 30/(30 \times F_t) \). Since the trading volume is huge for spot and the continuous contract of index futures, we consider that the impact cost for them is zero, which means \( C_{ss}, C_{sl}, C_{fs}, C_{fl} \) are all zero.

3. Empirical Analysis

3.1. No-Arbitrage Band and Mispricing Ratio

Based on these data, we calculated the no-arbitrage band and mispricing ratio of CSI300 index futures, which is drawn in Figure 1.

From Figure 1, we can see that most of time index futures stays out of upper limit of no-arbitrage band when mispricing exist. The detail of arbitrage opportunity and the time mispricing persist are shown in Table 1. There are 36.2% of the time that index futures exceeds upper limit, while only 0.2% for lower limit, which suggest there are much more cash-and-carry-arbitrage for CSI300 index futures than the reverse one. This is consistent with McMillan [7], who found there are only cash-and-carry-arbitrage chances for ISE40 index futures, and the reason is margin trading and short selling just start in China, which cannot short spot. Thus, the deficiency for investment channel and limitation for trading mechanism lead to the asymmetrical arbitrage direction. Moreover, the persistent time for positive mispricing is 12.6 minutes on average, which illustrate that index futures need 12.6 minutes to return to no-arbitrage band.

![Figure 1. No-arbitrage band and mispricing ratio of CSI300 index futures.](image)

### Table 1. Arbitrage opportunities and mispricing ratio.

<table>
<thead>
<tr>
<th>Samples</th>
<th>No-arbitrage band</th>
<th>No. of times</th>
<th>Frequency</th>
<th>Average time of mispricing</th>
<th>Average mispricing ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exceed upper limit</td>
<td>11,303</td>
<td>36.2%</td>
<td>12.6</td>
<td>0.0045</td>
<td></td>
</tr>
<tr>
<td>In no-arbitrage band</td>
<td>19,814</td>
<td>63.5%</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Exceed lower limit</td>
<td>68</td>
<td>0.2%</td>
<td>5.3</td>
<td>-0.0003</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2 depicts the relationship between mispricing and the time to expiration. From Figure 2(a), the biggest mispricing is happen on 23 days before maturity, which suggest if investors start to arbitrage at this time, they will gain profit mostly. Besides, it shows that the closer to expiration day, the less the mispricing and it appears. Figure 2(b) shows the average mispricing ratio for different days to expiration. In general, the closer to expiration day, the less mispricing ratio, which state that CSI300 index futures convergent to its underlying as closing to maturity. While when days to expiration is more than 24 days, there are huge mispricing ratio occasionally, but it is not significant on statistical.

3.2. The Result of ADF Model

Figure 3 is the coefficient of first order difference of mispricing, which examines its persistence. As the Figure 3 shown, the coefficient reaches the biggest when its 4th lagged is used, and the value is decreasing, becoming insignificant when up to 14th lag. It explains that mispricing would disappear after shown in a while, and it has mean reversion effect, which is consistent with Alphonse [29]. Therefore, once mispricing is discovered, arbitrageurs could short futures and long spot, and wind position when...

![Figure 2. (a) Relationship between mispricing and TTE; (b) Average mispricing ratio and the number of TTE.](image)
the futures return to no-arbitrage band, instead clear positions on expiration day.

The result of model 1 is shown in Table 2. The coefficient of D2 is negative significantly while D3 is insignificant, which suggest two points. First, cash-and-carry-arbitrage influences mean reversion effect strikingly. In a short time, if arbitrageurs found arbitrage opportunity, they would make the deal immediately, which drives futures going back to no-arbitrage band. Second, the coefficient of D1 is negative significantly indicates mean reversion exists even arbitrage is absent. Some hypotheses [31,32] demonstrate that the lead-lag relationship between futures and spot, which the response for information of spot is behind futures, can explain for non-arbitrage mean reversion effect. When some information shock financial market, index futures react fast, and mispricing would show up. Investors will take action to adjust their orders when they found the arbitrage, such as canceling the earlier order and submitting a new one, which will impel index futures get back to no-arbitrage band and mispricing disappear.

Higher orders autocorrelation coefficient of mispricing changes is estimated in model 2 to discuss the time that reversion of CSI300 index futures is biggest, including 2nd to 20th, 25th, 30th, 35th, 40th, 45th, 50th, 55th, and 60th order lag. The relationship between coefficient of D2 and the orders is shown on Figure 4. The coefficient of D2 in all orders are negative suggest cash-and-carry-arbitrage influences reversion effect invariably. The absolute value of 14th order difference is up to 0.0153, and of these orders more than 14th is maintaining around 0.015 to 0.016. It indicates arbitrage influence mean revision most is happening 14 minutes after mispricing emerge. The retardance may be caused by the whole process which need about 14 minutes from mispricing is found to finish the deal, such as orders distribution and approval.

### 3.3. Variables Effect Arbitrage

In order to analyze other factors that impact arbitrage opportunities, we regress arbitrage opportunities (AO) on a series of independent variables using logit model. Independent variables include time to expiration (TTE), trading volume of stock index futures(VOLUME) and its first-order lag (VOLUME (–1)), position of stock index futures (because first-order difference of position is steady, we use D (POSITION) as independent variable); highest and lowest price spread within one minute (BASIS). First, we test the causal relationship between the above variables and the arbitrage opportunities through second-order lag Granger causality test, as shown in Table 3. Besides the first-order difference of position, the remaining variables are all the Granger causes of arbitrage opportunities.

The results of logit model are shown in Table 4. All independent variables are significant. The coefficient of VOLUME (–1) is positive, while the coefficient of VOLUME is negative, indicating that the larger volume in the previous period, the higher probability of arbitrage opportunities while the larger current volume, the less probability of arbitrage opportunities. When investors find a mispricing in the first period, they will complete the deals

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Deviation</th>
<th>Z-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>2.60E−05***</td>
<td>4.24E−06</td>
<td>6.130829</td>
</tr>
<tr>
<td>D1</td>
<td>−0.989021***</td>
<td>0.024694</td>
<td>−40.05185</td>
</tr>
<tr>
<td>D2</td>
<td>−0.004483***</td>
<td>0.000662</td>
<td>−6.776642</td>
</tr>
<tr>
<td>D3</td>
<td>−0.087151</td>
<td>0.075530</td>
<td>−1.153854</td>
</tr>
</tbody>
</table>

**R-squared 0.101078**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Deviation</th>
<th>Z-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted R-square</td>
<td>0.100701</td>
<td>0.000698</td>
<td></td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>0.000662</td>
<td>Akaike info criterion</td>
<td>−11.80139</td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>0.013596</td>
<td>Schwarz criterion</td>
<td>−11.79762</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>183012.2</td>
<td>F-statistic</td>
<td>268.1270</td>
</tr>
</tbody>
</table>

***means the variable is significant under 1% level for T test.
Study on Stock Index Futures’ Mean Reversion Effect and Arbitrage in China Based on High-Frequency Data

Table 3. The result of Grange causality test.

<table>
<thead>
<tr>
<th>Ho</th>
<th>F-statistic and p value</th>
<th>Ho</th>
<th>F-statistic and p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTE is not the Granger cause of AO</td>
<td>140.381 (0.0000)</td>
<td>AO is not the Granger cause of TTE</td>
<td>0.12378 (0.8836)</td>
</tr>
<tr>
<td>VOLUME is not the Granger cause of AO</td>
<td>3.63730 (0.0263)</td>
<td>AO is not the Granger cause of VOLUME</td>
<td>67.5285 (0.0000)</td>
</tr>
<tr>
<td>VOLUME (–1) is not the Granger cause of AO</td>
<td>3.63730 (0.0263)</td>
<td>AO is not the Granger cause of VOLUME (–1)</td>
<td>67.5285 (0.0000)</td>
</tr>
<tr>
<td>D (POSITION) is not the Granger cause of AO</td>
<td>0.70457 (0.70457)</td>
<td>AO is not the Granger cause of D (POSITION)</td>
<td>19.9306 (0.0000)</td>
</tr>
<tr>
<td>BASIS is not the Granger cause of AO</td>
<td>15.6440 (0.0000)</td>
<td>AO is not the Granger cause of BASIS</td>
<td>294.728 (0.0000)</td>
</tr>
</tbody>
</table>

Table 4. Result of logit model.

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Coefficient</th>
<th>Standard deviation</th>
<th>Z-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>–3.724031***</td>
<td>0.045370</td>
<td>–82.08164</td>
</tr>
<tr>
<td>TTE</td>
<td>0.130408***</td>
<td>0.001792</td>
<td>72.77045</td>
</tr>
<tr>
<td>VOLUME</td>
<td>–0.001206***</td>
<td>5.14E-05</td>
<td>–23.49371</td>
</tr>
<tr>
<td>VOLUME (–1)</td>
<td>0.000346***</td>
<td>3.34E-05</td>
<td>10.36400</td>
</tr>
<tr>
<td>BASIS</td>
<td>0.386481***</td>
<td>0.010229</td>
<td>37.78471</td>
</tr>
</tbody>
</table>

rapidly, therefore the futures price return to the arbitrage-free interval driven by the larger volume in the subsequent period, then the mispricing disappears, which is consistent with the findings of Merrick. The coefficient of TTE is positive, which means the closer to expiration, the more convergence to the theoretical value of stock index, lead to fewer arbitrage opportunities. The coefficient of BASIS, which to some extent reflects stock index futures volatility, is positive. The greater the spread, the more differences among investor’s limit orders, the higher the chance of mispricing occurred at this time, that is, the more arbitrage opportunities.

4. Conclusions

Based on high frequency data of CSI300 index futures, this paper construct no-arbitrage band of continuous contract by cost of carry theory, and empirically research future-spot arbitrage. Then, we investigate mean reversion effect of CSI300 index futures’ mispricing and the effect’s time for the first time by ADF model. Furthermore, a logit model is estimated to test for some variables that may influence arbitrage opportunity. The results are following.

First, CSI300 index futures had cash-and-carry arbitrage opportunities for the most part from 23rd August 2010 to 18th February 2011. Once it exceeds upper limit of no-arbitrage band, it would cost 12.6 minutes to return to the band.

Second, cash-and-carry arbitrage influence mispricing’s mean reversion significantly, while the reverse arbitrage is insignificant. Due to deficiency of investment channels and limitation of trading mechanism, the arbitrage direction effects mean reversion of index futures asymmetrically.

Third, in absence of arbitrage, mean reversion exists still. This is because when financial market is shocked some information, index futures react fast, and mispricing would show up. Investors will adjust their orders for stock which drive index futures return to no-arbitrage band.

Fourth, arbitrage influence mean revision most is happening 14 minutes after mispricing appears. At this moment, if arbitrage orders deal and wind the position when index futures return to no-arbitrage, the profit will be obtained. This may result from the retardance from mispricing is found to finish the deal.

Finally, arbitrage opportunities are positive correlated to the first order lag of trading volume, while negative correlated to volume, which results from that large trading volume drives index futures return no-arbitrage band. Additionally, arbitrage opportunities are positive related to time to expiration, spread between highest and lowest price within one minute. This is because the closer to expiration, the more convergence to the theoretical value of stock index, lead to fewer arbitrage opportunities. The greater the spread, the more differences among investor’s limit orders, the higher the possibility of mispricing occurred at this time, which means the more arbitrage opportunities.

5. Acknowledgements

This research is supported by the National Natural Science Foundation of China (No. 70801006, No. 71173023) and the Fundamental Research Funds for the Central Universities.
REFERENCES


