This is the peer reviewed version of the following article: Jin, M, Kearney, F, Li, Y, Yang, YC. Order book price impact in the Chinese soybean futures market. Int J Fin Econ. 2021; 1– 20., which has been published in final form at https://doi.org/10.1002/ijfe.2439.

Order Book Price Impact in the Chinese Soybean Futures Market

Muzhao Jin
Queen's Management School
Queen's University Belfast
185 Stranmillis Road, Belfast BT9 5EE, UK.
Email: mjin02@qub.ac.uk

Fearghal Kearney
Queen's Management School
Queen's University Belfast
185 Stranmillis Road, Belfast BT9 5EE, UK.
Email: f.kearney@qub.ac.uk

Youwei Li
Hull University Business School
University of Hull
Cottingham Road, HU6 7RX, UK.
Email: youwei.li@hull.ac.uk

Yung Chiang Yang
UCD College of Business
University College Dublin
Belfield, Dublin, 4, Ireland.
Email: yungchiang.yang@ucd.ie

December 2020

Order Book Price Impact in the Chinese Soybean Futures Market

Abstract

We study the price impact of order flow in the world's largest soybean meal futures markets. Our intraday results indicate that incoming orders can be used to explain price changes and to significantly predict future price changes. Our results are shown to be robust to various order flow measures, price aggregation approaches and data frequencies. We compare various order flow measures; finding that Order Flow Imbalance (OFI) is a more all-encompassing measure carrying greater information about price change relative to both Trade Imbalance (TI) and volume. Moreover, while both OFI and TI are shown to predict future price changes, this predictability diminishes over longer measure and price change frequency horizons.

JEL: G12, G13, G15

Keywords: Order flow imbalance; Limit order book; Price impact; Futures market

1. Introduction

This study investigates the impact of order flow information on Chinese soybean meal futures prices. Chinese soybean meal futures are of great importance for a number of reasons. First, China is the largest soybean meal futures market in the world according to the Futures Industry Association (FIA) (Acworth, 2014; 2015). Second, soybean meal futures are the largest agricultural futures market in China, in terms of trading volume. Finally, they were the first agricultural product when futures trading started in China in 1993. Trading operates in an electronic market, whereby demand and supply of the futures contracts are represented by limit orders posted in the book, with trades occurring when buy or sell orders match (Obizhaeva & Wang, 2013). Given this market structure, much information to be incorporated into the market price will be posted in that order book. Therefore, effectively gleaning information from the flow of orders is of great importance in order to understand the price formation process. In this paper, we aim to identify the most informative order flow measure.

Price impact refers to the correlation between an incoming order (to buy or to sell) and the subsequent price change (Bouchaud, 2010). Simply put, it is how a buy trade pushes the price upwards and how a sell trade pushes the price downwards. This price impact may bring with it some unintended consequences, with the second buy trade possibly costing more due to the impact of the first buy trade. As outlined in Guéant (2015), the first trade may even lead to such an inflated cost having to be incurred to place the second trade, that the proceeding transaction may fail. Therefore, monitoring and controlling price impact is an important consideration for investors especially for those with trading large volumes or those employing sequential trading strategies.

There are relatively few prior order flow studies of Chinese markets, with Shenoy and Zhang (2007), He et al. (2014), Narayan et al. (2015), Wang et al. (2016) and Lao et al. (2017), being notable exceptions. While Shenoy and Zhang (2007), Narayan et al. (2015) and Lao et al. (2017) find a positive contemporaneous relation between daily trade imbalance and price change in Chinese stock markets, the agricultural futures focus of He et al. (2014) and Wang et al. (2016) is more closely aligned to our study. He et al. (2014) study Chinese agricultural futures markets, discovering a strong and positive contemporaneous relation between price and volume for six agricultural futures contracts including soybean meal. Wang et al. (2016) on the other hand do not find a significant relation between lagged order book imbalance and agricultural futures price changes. The lack of consensus between contemporaneous and future price impact motivates our study, as we seek to provide evidence of the link between the order book and price change. Furthermore, our paper differentiates itself by using a unique intraday data set of the most actively traded futures contract, as well as carrying out a more in-depth study through the use of a greater number of sophisticated order flow measures, the relation between them, and both contemporaneous and future price impact.

We compare the explanatory power of the below three order flow measures on price impact. Firstly, the Order Flow Imbalance (OFI) measure, as derived by Cont et al. (2014). They calculate OFI as the difference between supply and demand; represented by the best bid and ask volumes in the order book. This imbalance is in turn used to study price impact, finding that both the sign and magnitude of OFI is positively correlated with price change for NYSE stocks. Despite this measure being a relatively recent innovation, it has previously been adopted to study the effect of order flow across markets. Hyun et al. (2016) is one such example, finding that OFI is positively related to price changes for the KOSPI

200 futures contract. In our paper, we provide additional empirical evidence of OFI's usefulness, in the context of the unique Chinese soybean meal futures market.

The second measure used in our study is trade imbalance (TI), as proposed by Chordia and Subrahmanyam (2004).[1] They define the difference between buyer-initiated and seller-initiated trading volumes as the trade imbalance. They find that both the sign and magnitude of TI are positively correlated with price change for NYSE stocks. This measure is widely used in empirical studies of price impact (Chordia et al. 2008; Kaniel et al. 2008; Bailey et al. 2009; Chang and Shie 2011; Huang 2011; and Rastoqi et al. 2013).

The third measure of order flow we consider is trading volume. Since the seminal paper of Crouch (1970), a large number of price impact studies focus on the relation between trading volume and price change. Crouch (1970) finds a positive correlation between trading volume and absolute price change for both the Dow Jones index and NYSE stocks. Since then a large body of theoretical and empirical studies find a positive correlation between price and volume (Clark 1973; Copeland 1976; Epps and Epps 1976; Tauchen and Pitts 1983; Kyle 1985; Harris and Gurel 1986; Admati and Pfleiderer 1988; Harris and Raviv 1993; Ho et al. 1993; Shalen 1993; Gurgul et al. 2005; Huang et al. 2012; and He et al. 2014).

Our analysis shows that all three measures of order flow are able to explain some degree of price change. We sample the data every 10 seconds, and utilise 15 minutes as a trading interval. We find that 97.87% of order flow imbalance coefficients across all of the 15-minute trading intervals are statistically significant. The outcome is similar if we use trade imbalance (93.56%) or trading volume (80.04%) as alternate measures of order flow.

[1] Chordia and Subrahmanyam (2004) use the term "order imbalance" instead of "trade imbalance". We relabel it as trade imbalance to make a clearer distinction from Cont et al.'s (2014) order flow imbalance.

We further verify our results by specifying a full trading day as the trading interval and sample the order flow and price impact every minute. The coefficients are significant in 97.58%, 91.48%, and 97.67% of trading intervals, respectively. Furthermore, when we compare the explanatory power among order flow imbalance, trade imbalance, and trading volume, we find that order flow imbalance subsumes information gleaned from both the trading volume and trade imbalance.

Other than OFI, TI and volume, market depth is another key variable when assessing the order book. Knez and Ready (1996) and Hasbrouck and Seppi (2001) conclude that market depth is able to explain price impact with Farmer et al. (2004) and Weber and Rosenow (2006) finding that low depth is associated with large price changes. Cont et al. (2014) derive an implied order book depth which is half of the reciprocal of the OFI price impact coefficient in their stylized model. However, they find that the implied order book depth is in fact only a rough proxy of market depth. We also compare the implied market depth with price impact and find that implied depth is lower than half of the reciprocal of the OFI price impact coefficient. This means that in Chinese soybean meal markets price is more resilient to incoming orders than indicated by market depth.

Besides using information from order flow to explain contemporaneous price impact, we also study the predictive power of order flow for future price change. Prior studies for other markets document that order flow is able to predict future returns (Chordia et al., 2002, Berkman and Lee, 2002, Chordia and Subrahmanyam, 2004, Lee et al., 2004, Shenoy and Zhang, 2007, and Narayan et al. 2015). Our finding is that for Chinese soybean meal futures order flow information is able to predict future return, but that the lifespan of this predictive power is short. We find that prior 10-second order flow information is able to predict price change up to 90 seconds in the future. These results are consistent with

Narayan et al. (2015), with trade imbalance being used to predict Chinese stock returns from one to 90 minutes ahead. Furthermore, we discover that the higher the level of trading activity by a liquidity provider, the higher the contemporaneous price impact and the shorter the lifespan of order flow impact on future price change. These results have important implications for both academic and market practitioners, alike.

In summary, this paper contributes to the literature in several aspects. First, distinct from most previous studies which focus on stock markets, we provide new evidence of order flow price impact in commodity futures markets. This serves as an out-of-sample test of existing findings and theories. Second, we identify the most informative order flow measure by taking advantage of our novel high frequency data set and comprehensively studying order flow price impact in the Chinese soybean meal futures market. Third, the agricultural futures market in China has become more and more important to the world in recent years. Furthermore, soybean meal futures is the largest and most important market, with our study contributing to a better understanding of order flow price impact, not only in this important market particularly, but also more generally in emerging markets.

The rest of the paper is organized as follows. We highlight the importance of the soybean meal futures market in China and the data used in the study in Section 2. Section 3 is devoted to the empirical methodology including a brief outline of the stylized model in Cont et al. (2014). We present our main findings in Section 4, with Section 5 concluding.

2. Data Description and Institutional Features

2.1 Soybean meal market in China

We now introduce the Chinese soybean meal market before describing the data used in our study. The Dalian Commodity Exchange (DCE) was first established on February 28, 1993.

However, no contracts were listed until November 18, 1993. Soybean meal became the first batch of standardized futures contracts traded on the exchange. Following the merger of China's futures exchanges, all contract listings were suspended or removed from DCE. Soybean meal was the first product to be re-listed with DCE, with trading reinstated on July 17, 2000. As one of the most actively traded futures contracts in the Chinese market, soybean meal futures are of great interest to both market participants and academic readers.

According to the FIA rankings (Acworth, 2014; 2015) presented in Table A1, the trading volume of Chinese soybean meal futures constitutes the largest soybean meal futures market in the world and is also ranked as second among all agricultural futures and options contracts in 2014. In 2015, trading volume of soybean meal futures increased by 41.2% as it became the largest agricultural futures contract. Furthermore, seven out of ten of the world’s most active agricultural futures were traded in China in 2015. Therefore, understanding the agricultural futures market in China is an important topic in a global context. This is particularly true for the soybean meal contract, as it is not only the largest agriculture contract traded in China but also the first futures contract traded after the Chinese government revamped the original futures market (CSRC, 2009). Therefore, Chinese soybean meal futures command a very prominent position in a global context.

According to the development report by Shanghai Futures Exchange (Yang and Xu, 2016), only 3% of market participants were institutional investors during 2008-2014, with the remaining 97% being individual investors. Furthermore, institutional investors accounted for less than 10% of trading volume during 2008-2014. Therefore, we can see that the Chinese soybean meal market is dominated by individual investors. In addition, the turnover rate (ratio of trading volume to open interest) of individual investors in the Chinese futures market is much higher than the turnover rate of institutional investors. For instance, the average turnover rate of individual investors is five times that of institutional investors. This is an indication that Chinese futures markets positions are driven predominantly by speculative individual investors.

2.2 Data and sample selection

Our data set for soybean meal contracts contains snapshot tick records at an interval of one every quarter second. It includes time, price, trading volume, open interest, bid volume, ask volume, average bid price, average ask price, market depth, and order queue. The data excludes the call auction periods before market open and after market close. The trading hours for the contracts are 9:00-11:30 and 13:30-15:00, Monday to Friday. There is however, a short 15 minutes break in trading from 10:15 to 10:30 every day. The sample period for the soybean meal data is June 2, 2008 to June 28, 2013, totalling 1,105 trading days.

There are eight futures contracts traded on the soybean meal futures market at any one time, with each contract maturing in a different month. The months of contract maturity are January, March, May, July, August, September, November and December. There are two common methods for constructing the required single time series of price data from contracts of different maturities. The first method is to splice near-to-maturity contracts conditional on liquidity to represent the price series. This method is outlined in Booth et al. (1999) and is based on the rationale that the expiring contracts have more information contained in their price. It is employed in recent empirical research such as Covrig et al. (2004), Shastri et al. (2008), Cabrera et al. (2009) and Chen et al. (2010). The second commonly employed method utilizes only recently issued or on-the-run contracts instead of expiring contracts. Fricke et al. (2011) present this “auto scroll” process that uses on-the-run contracts with the highest trading volume to combine prices of multiple contracts into a single price series.

Trading in Chinese soybean meal futures is unusual in that it is concentrated in only three contracts that mature in January, May, and September. The contracts with maturity in these three months contribute to approximately 99% of the trading volume. This phenomenon is primarily due to the seasonality of agricultural products. In China, September is the last month before the autumn harvest, and January is the main time to sell agricultural products before the Chinese New Year. However, despite May being of no great significance to agricultural production, it is the middle contract between January and September, so investors focus on it to fill leftover market demand. This focused trading activity creates a distinctive seasonality characteristic. While the contracts of Chinese soybean meal futures are dominated by January, May and September contracts, these three contracts do not have a fixed active period of trading before the next contract takes over. Therefore, to construct our single price series, we adopt the "auto scroll" process outlined above, conditional on trading volume for each of these three dominated contracts.

2.3 Determine the sampling interval

We sample the price and order flow information at fixed time intervals to construct a continuous time-series of these variables. We firstly, follow Cont et al. (2014) in using 10-second frequency data to construct each 15-minute subsample. However, we also use alternative continuous times-series with different intervals to serve as robustness checks for data frequency sensitivity.

When selecting alternative time intervals, we jointly consider allowing both enough information to arrive to the market and also using a short enough time interval to maintain a large sample size. Therefore, we analyse our data with time intervals of 10-second, 20-second, 30-second, 1-minute, 2-minute, 3-minute, 5-minute, 10-minute, 15-minute and 20-minute as shown in Table 1. We use trading volume to compute non-trading probabilities within these intervals.

[Insert Table 1 Around Here]

Based on the above table, we find that all intervals look reasonable, with the non-trading probability of the 10-second interval being only 1.52%, and less than 1% for all others. Based on this, we specify 1-minute to compute the price change and OFI as it is the most granular frequency with a non-trading probability of less than 0.5%. Alongside this, we employ daily subsamples to estimate our OLS regressions.

2.4 Variable constructions

2.4.1 A stylized model of the order book

Before we introduce the variables used in our study, we first summarize the model proposed by Cont et al. (2014), in terms of how they construct OFI. In an order driven market, the order book is altered through three different actions or events, namely, market order (M), limit order (L) and order cancellation (C). Due to the relatively recent advancement of reliable database technology, these three events are now accurately recorded. If we assume that limit order arrivals and cancellations only occur at the best bid/ask, the depth (quantity available) at each price level for both bid and ask is constant and equal to D. For this reason the mid-price change can only be influenced by the three events identified above (M, L, C) and the market depth (D):

$$\Delta P^b = \frac{\delta(L^b - C^b - M^s)}{D} \ , \tag{1}$$

where ΔP^b is the price change on the bid side, δ is the difference between each limit order price level, L^b and C^b are bid side limit orders and cancellations, respectively, M^s is the market order on the ask side, D is market depth. $L^b - C^b - M^s$ represents the net order flow change on the bid side. When the net order flow change reaches D, it means all bid side quantity is executed using the best bid price, with the second-best limit order price becoming the new best bid price, and so on.

Similarly, for the ask side, the price change can be defined as:

$$\Delta P^s = -\frac{\delta(L^s - C^s - M^b)}{D} \ . \tag{2}$$

Therefore, the mid-price change can be shown as the average of price change on the bid and ask sides:

$$\Delta P = \frac{1}{2} \ \frac{\delta(L^b - C^b - M^s)}{D} - \frac{1}{2} \ \frac{\delta(L^s - C^s - M^b)}{D} \ . \tag{3}$$

For simplicity Cont et al. (2014) normalize the mid-price by the difference between each limit order price level $P = \frac{P^b + P^s}{2\delta}$, therefore, the mid-price change can be rewritten as:

$$\Delta P = \frac{OFI}{2D} \ , \tag{4}$$

$$OFI = L^b - C^b - M^s - L^s + C^s + M^b \ . \tag{5}$$

In our study, the data set comprises four times a second snapshot tick records. Unlike transaction data, the data we use does not include each order book event. Therefore, we use the method of Cont et al. (2014) on the change of limit order price and quantity at the best bid and ask, to identify order flow imbalance.

2.4.2 Order flow imbalance

The limit order book only includes four variables, the bid price P^B, the size of the bid queue q^B, the ask price P^A, and the size of the ask queue q^A. The bid side represents the demand for the contract with the ask price and size representing the supply of the contract. P^B is the price investors are willing to pay for a contract and q^B is the number of contracts they are willing to buy. While the ask price and size represent the supply side of the contract. Using time ordered variables $(P_n^B, q_n^B, P_n^A, q_n^A)$, we can compare two time points *n* and *n-1*, with only one of the following events occurring:

$P_n^B > P_{n-1}^B$ or $q_n^B > q_{n-1}^B$ means an increase in demand. As P^B is the best (highest) bid price, $P_n^B > P_{n-1}^B$ means investors are willing to buy the contract at a higher price than before. Meanwhile, if P^B remains the same, and $q_n^B > q_{n-1}^B$, it means investors are willing to buy a greater number of contracts. Likewise, $P_n^B < P_{n-1}^B$ or $q_n^B < q_{n-1}^B$ means a decrease in demand.[2]

Therefore, the change on the bid side in a given period can be defined as:

$$\Delta\text{B:}\ \ I\{P_n^B \geq P_{n-1}^B\}q_n^B - I\{P_n^B \leq P_{n-1}^B\}q_{n-1}^B\ , \tag{6}$$

with the change on the ask side represented as:

$$\Delta\text{A:}\ \ I\{P_n^A \leq P_{n-1}^A\}q_n^A - I\{P_n^A \geq P_{n-1}^A\}q_{n-1}^A\ . \tag{7}$$

[2] On the other hand, $P_n^A < P_{n-1}^A$ or $q_n^A < q_{n-1}^A$ means an increase in supply. As P^A is the best (lowest) ask price, $P_n^A < P_{n-1}^A$ means investors are willing to sell the contract at a lower price than before. Meanwhile, if P^A remains the same and $q_n^A < q_{n-1}^A$, it means investors are willing to sell a greater number of contracts. Likewise, $P_n^A > P_{n-1}^A$ or $q_n^A > q_{n-1}^A$ means a decrease in supply.

Utilising the above notation, the contribution of the n^{th} event, e_n, can be represented using the following formula:

$$\begin{aligned} e_n = & (I\{P_n^B \geq P_{n-1}^B\}q_n^B - I\{P_n^B \leq P_{n-1}^B\}q_{n-1}^B) \\ & - (I\{P_n^A \leq P_{n-1}^A\}q_n^A - I\{P_n^A \geq P_{n-1}^A\}q_{n-1}^A), \end{aligned} \tag{8}$$

and the order flow imbalance over the time interval $[t_{k-1}, t_k]$ can be defined as:

$$\mathrm{OFI} = \sum_{n=N(t_{k-1})+1}^{N(t_k)} e_n \ . \tag{9}$$

2.4.3 Market depth

In the above stylized model, price changes based only on *OFI* and market depth. Market depth is assumed constant, however in practice, market depth at each price level for both bid and ask is not constant. Therefore, a closer approximation of actual market depth can be achieved by averaging the bid/ask quantity right before or right after a price change:[3]

$$\begin{aligned} D_i & \\ = & \frac{1}{2}\left[\frac{\sum_{n=N(T_{i-1})+1}^{N(T_i)}(q_n^B I\{p_n^B < p_{n-1}^B\} + q_{n-1}^B I\{p_n^B > p_{n-1}^B\})}{\sum_{n=N(T_{i-1})+1}^{N(T_i)} I\{p_n^B \neq p_{n-1}^B\}}\right. \\ & \left. + \frac{\sum_{n=N(T_{i-1})+1}^{N(T_i)}(q_n^A I\{p_n^A > p_{n-1}^A\} + q_{n-1}^A I\{p_n^A < p_{n-1}^A\})}{\sum_{n=N(T_{i-1})+1}^{N(T_i)} I\{p_n^A \neq p_{n-1}^A\}}\right] . \end{aligned} \tag{10}$$

2.4.4 Trade imbalance

[3] In this paper, we specify time periods T_i and T_{i-1} to differ by 15 minutes, when t_k and t_{k-1} differ by 10 seconds, and time periods T_i and T_{i-1} to differ by a day, when t_k and t_{k-1} differ by 1 minute.

Besides *OFI*, there is another popular way to assess Order Imbalance, namely Chordia and Subrahmanyam (2004)'s supply and demand measure. Their approach is based on the method of Lee and Ready (1991) in that they infer trade direction using buyer-initiated market orders minus seller-initiated market orders. They define the relative number of buyer-initiated and seller-initiated trades in the market as order imbalance, also known as trade imbalance (*TI*):

$$TI_k = \sum_{n=N(t_{k-1})+1}^{N(t_k)} b_n - \sum_{n=N(t_{k-1})+1}^{N(t_k)} s_n = \sum_{n=N(t_{k-1})+1}^{N(t_k)} E_n \ , \tag{11}$$

$$\begin{aligned} E_n = {} & (I\{P_n^T > P_{n-1}^M\}VOL_n - I\{P_n^T < P_{n-1}^M\}VOL_n) \\ & + (I\{(P_n^T = P_{n-1}^M) \cup (P_n^T > P_{n-1}^T)\}VOL_n \\ & - I\{(P_n^T = P_{n-1}^M) \cup (P_n^T < P_{n-1}^T)\}VOL_n) \ , \end{aligned} \tag{12}$$

where b_n is the number of buy-initiated trades at the nth-quote, s_n is the number of sell-initiated trade at the nth-quote, P^T is transaction price, P^M is mid-price.

2.4.5 Trading volume

There are a number of studies that focus on the relation between trading volume and price change. We use total trading volume that includes both buy side and sell side volumes.

$$VOL_k = \sum_{n=N(t_{k-1})+1}^{N(t_k)} b_n + \sum_{n=N(t_{k-1})+1}^{N(t_k)} s_n \ , \tag{13}$$

where b_n is the number of buy-initiated trades at the n^{th} quote and s_n is the number of sell-initiated trades at the n^{th} quote.

2.5 Summary statistics

After defining the variables used in our paper, we present the descriptive statistics of the order book and main variables in Table 2.

[Insert Table 2 Around Here]

Panel A of Table 2 presents summary statistics of the order book which include average price, daily trading volume, price change of best bid (ask) price, average bid (ask) depth and average spread. From this table, we find that the average price is CNY 3,162.37, with daily trading volume standing at 998,376.92. The average absolute price change of both the best bid and best ask is 1.08. This finding is consistent with the stylized model outlined in Section 2.4.1, that the difference between both bid and ask is constant for each price level, with both of them approximatively equal to the minimum price fluctuation of 1. The average bid depth is 411.89 and average ask depth is 418.30. These two numbers being broadly similar is also consistent with the above stylized model, which assumes that depth for both bid and ask is constant. We hypothesise that the main reason the soybean meal market is conforming to the above stylised models is because of the high liquidity levels it demonstrates as the largest agricultural futures contract in the Chinese market.

Panel B of Table 2 presents the summary statistics and distribution of main variables used in the analysis. The means and median of the mid-price and transaction price changes are zero for the 10-second interval, indicating that the mid-price and transaction price remain mostly unchanged during the short interval. The means and median of *OFI* and *TI* are around zero but they have large variation. The summary statistics for the longer 15-minute interval have larger variation compared to those of the 10-second interval. This is not surprising as we expect price change, order flow imbalance, and trade imbalance to be larger when measured over a longer interval.

In order to compare order flow imbalance, trade imbalance, and price change, we normalize these variables and plot them on the same diagram.

[Insert Figure 1 Around Here]

Figure 1 compares the average 15-minute absolute mid-price change ΔP, Order Flow Imbalance (*OFI*), and Trade Imbalance (*TI*) for Chinese soybean meal futures. From this, we can observe, that these three variables at a 15-minute frequency show a similar U-shape, in that they increase during the opening time period and subsequently decrease, finally reversing this trend near market close.

We also compare price change $|\Delta P|$, absolute Order Flow Imbalance $|OFI|$, absolute Trade Imbalance $|TI|$ and trading volume for Chinese Soybean Meal futures. We observe in Figure 2 that these four variables at a 15-minute frequency show a similar U-shape, in that they increase during the market opening period, subsequently decrease and finally reverse as the market closes.

[Insert Figure 2 Around Here]

These figures indicate that all three incoming orders exhibit an intraday shape consistent with price change.

3. Methodology

3.1 Relation between OFI and price change

The stylized model presented in Section 2 is able to explain the relation between *OFI* and price change in the economic sense. However, the above assumptions may not be valid in reality, for instance, the market depth may not be constant. Therefore, we use the following regression model to empirically test the relation between *OFI* and price change:

$$\Delta \mathrm{P} = \alpha_i + \beta_i \ OFI_k + \epsilon_k \ , \tag{14}$$

where β is the price impact coefficient.

3.2 Relation between price impact coefficient and market depth

Furthermore, based on the stylized model of Cont et al. (2014) that the price impact coefficient is inversely related to market depth, the implied order book depth can be defined as $\frac{1}{2\beta}$, where β is the price impact coefficient of *OFI* from Eq. (14). In our study, we analyse the relation between the *OFI* price impact coefficient and market depth in the Chinese soybean meal market:

$$\beta_i = \frac{c}{D_i^{\lambda}} + v_i \ , \tag{15}$$

where c and λ are constants, and v_i is a noise term. D is the market depth. If the implied order book depth is indeed $\frac{1}{2\beta}$, then $c = \frac{1}{2}$ and λ = 1.

3.3 Comparison of order flow imbalance and trade imbalance

Trade imbalance as proposed by Chordia and Subrahmanyam (2004) is an alternative to order flow imbalance. While Cont et al.'s (2014) OFI captures the changing of ex-ante order flow, Chordia and Subrahmanyam's (2004) TI measures the realised trade imbalance.

In order to compare and contrast the explanatory power of order flow imbalance (*OFI*) and trade imbalance (*TI*), we use the following three regressions:

$$\Delta \mathrm{P} = \alpha_i + \beta_i \, OFI_k + \epsilon_k \ , \tag{16a}$$

$$\Delta \mathrm{P} = \mathrm{\alpha_i} + \mathrm{\beta_i} \, \mathrm{TI_k} + \mathrm{\epsilon_k} \ , \tag{16b}$$

$$\Delta \mathrm{P} = \mathrm{\alpha_i} + \mathrm{\beta_{i,1}} \, \mathrm{OFI_k} + \mathrm{\beta_{i,2}} \, \mathrm{TI_k} + \mathrm{\epsilon_k} \ . \tag{16c}$$

3.4 Comparison of traded volume and order flow imbalance

Many of the early studies on price impact (e.g., Karpoff, 1987) focus on the relation between trading volume and price change. Cont et al. (2014), in contrast study the relation between volume and price volatility, finding an exponential relation between volume and price magnitude. To compare the explanatory power of *OFI* and volume, we take the absolute values of the price change and *OFI* because these values are signed while volume is unsigned. Subsequently, we run the following regressions.

$$|\Delta \mathrm{P}| = \alpha_i + \beta_i \ |OFI_k| + \epsilon_k \ , \tag{17a}$$

$$|\Delta \mathrm{P}| = \alpha_i + \beta_i VOL_k^{H_i} + \epsilon_k, \tag{17b}$$

$$|\Delta \mathrm{P}| = \alpha_i + \beta_{i,1} \ |OFI_k| + \beta_{i,2} VOL_k^{H_i} + \epsilon_k \ . \tag{17c}$$

3.5 Future price impact

In most price impact studies, the predictability of order imbalance for future price change has received the greatest levels of attention (Chordia et al. 2002; Berkman and Lee 2002; Chordia and Subrahmanyam 2004; Lee et al. 2004; Shenoy and Zhang 2007; Cont et al. 2014; Narayan et al. 2015; and Hyun et al. 2016). We extend prior studies on the predictive power of order flow by establishing the lifespan of order flow predictive power for future price change.

To this end, we use the following three first order lagged regressions:

$$\Delta \mathrm{P} = \alpha_{\mathrm{i}} + \beta_{\mathrm{i}}\mathrm{lag}(OFI_k) + \epsilon_k \ , \tag{18a}$$

$$\Delta \mathrm{P} = \alpha_{\mathrm{i}} + \beta_{\mathrm{i}}\mathrm{lag}(\mathrm{TI_k}) + \epsilon_{\mathrm{k}} \ , \tag{18b}$$

$$\Delta \mathrm{P} = \alpha_{\mathrm{i}} + \beta_{\mathrm{i},1}\ \mathrm{lag}(\mathrm{OFI_k}) + \beta_{\mathrm{i},2}\ \mathrm{lag}(\mathrm{TI_k}) + \epsilon_{\mathrm{k}} \ . \tag{18c}$$

If we establish that *OFI* is useful in forecasting future price change, we can then use the following regression to study how long it affects price change for:

$$\Delta \mathrm{P} = \alpha_i + \beta_i\ lag_{pre}(OFI_k) + \epsilon_k \ , \tag{19}$$

where pre is the pre-specified lag order of OFI_k between 1 and 12. As we specify 10-second price changes and incoming orders each lag is 10 seconds apart. Finally, we also use Granger causality between any two of *OFI*, *TI* and price change to find which measure leads the others.

4. Empirical Findings

4.1 Price impact

[Insert Table 3 Around Here]

In this section, we first investigate the relation between price change and *OFI* with Table 3 reporting the regression results. Panel A relates mid-price change and *OFI*. We find that both 15-minute and 1-day specifications provide strong evidence that *OFI* can be used to explain price change with an average significant β coefficient of 0.002 using both intraday and daily frequencies. Furthermore, the last three columns show the percentage of subsamples where the coefficients were deemed to be significant, finding over 97% of significant *OFI* coefficients in both cases.[4] The findings indicate that *OFI* has a strong

[4] The percentage of significance coefficient is calculated as the number of regressions that produce significant

explanatory power for price changes with *OFI* being positively correlated with price change. This result is consistent with both Cont et al. (2014) and Hyun et al. (2016). Furthermore, we also use transaction prices to verify Cont et al. (2014)'s claim that their adoption makes no material difference. The results are shown in the Panel B regression, with the results being qualitatively similar to the use of mid-price.

We pose the question; do past price changes influence current price change? In order to answer this, we use the Akaike's information criterion (AIC) test to determine the number of lagged dependent variables to be included. We run the regression for every long interval (either 15-minute or 1-day), with AIC suggesting an optimal number of lags for each long period. Instead of incorporating a different number of lags into each model, we adopt a fixed number of lags across all regression models. In the 15-minute long interval regression, AIC indicates the inclusion of one lag for most models, followed by zero and two lags. Therefore, we include one lag of the dependent variable and re-examine the relation between price change and *OFI*. The results of these models that include one lag of price change are shown in Panel C and D of Table 3. The results are qualitatively similar to that of our baseline, with the *OFI* coefficient becoming even more significant after including lag price change in the model. We have also conducted the analysis using two lagged dependent variables, and the result remains unchanged.

4.2 Relation between price impact and market depth

coefficients at the 10% level, divided by the total number of models. For example, there are 16,417 15-minute long intervals in our sample and we produce 16,417 regressions. 16,068 of the *OFI* coefficients (β) are significantly different from zero at the 10% level, therefore, the percentage of significance is 97.87%.

In this section, we empirically test the stylized model of Cont et al. (2014), who propose that implied order book depth can be defined as $\frac{1}{2\beta}$, where β is the OFI price impact coefficient. If the implied order book depth is indeed $\frac{1}{2\beta}$, then $c = \frac{1}{2}$ and λ = 1 in Eq. (15).

[Insert Table 4 Around Here]

Table 4 reports the relation between the *OFI* price impact coefficient calculated in Section 4.1 and market depth. We have 16,417 and 1,105 estimates of β under 15-minute and 1-day time horizons, respectively. We find, using the mid-price results that the exponent number, λ is equal to 0.827 (0.885) and that the coefficient c is equal to 0.161 (0.215) using 15-minute (daily) frequency. The significant coefficients λ and c imply that the impact coefficient is inversely proportional to market depth, with similar results being observed using transaction prices. As the results do not directly confirm the Cont et al. (2014) stylized values, the inference is that Chinese soybean meal futures prices are more resilient to incoming orders than indicated by market depth alone.

4.3 Comparison of order flow imbalance and trade imbalance

As opposed to using *OFI* in isolation we now employ another measure, *TI*, to define order book imbalance.

[Insert Table 5 Around Here]

Table 5 presents a comparison of *OFI* and *TI* with Panel A showing the results calculated using mid-price. Having previously established that *OFI* has significant positive price impact, we now find that *TI* is also significantly positively related to contemporaneous price change. Comparing *OFI* and *TI*, we find that *OFI* is more informative in explaining price changes, with only 93.56% (91.48%) of 5-minute (1-day) subsamples deemed to be significant for *TI* versus over 97% for *OFI*. Furthermore, when both *OFI* and *TI* are included in the regression model at a 15-minute (1-day) time horizon, only the *OFI* variable retains strong explanatory power with 93.84% (91.93%) of the subsample's *OFI* coefficients passing the *t*-test versus only 31.41% (22.33%) of the subsample's *TI* coefficients passing. The *t*-statistics of *TI* also reduce sharply when *OFI* is included, but the significance level of *OFI* remains similar to the simple regression model. These results are obtained using transaction price but are consistent with those of mid-price.[5] These results highlight that *OFI* is superior to *TI*, which is consistent with Cont et al. (2014) who argue that high-frequency price changes are mainly driven by imbalances between supply and demand at the best bid and ask prices (i.e. Order Flow Imbalance). However, we should not ignore *TI* since it captures a different aspect of information from high-frequency transactions.

Although both *OFI* and *TI* are used to measure order imbalance, *OFI* is derived from changes in the limit order book, whereas, *TI* is derived from trading volume. In an order driven market such as for Chinese soybean futures, limit orders, market orders and cancellations all impact the order book, and hence *OFI*. However, trading volume is only affected by market orders. Therefore, for this reason, *OFI* is a more all-encompassing measure relative to *TI,* which we believe drives the superior empirical results observed here.

4.4 Comparison of traded volume and order flow imbalance

[5] We also performed regressions with one lag of the dependent variable, as in the case of Panels C and D of Table 3, and the results remained unchanged.

The original price impact studies focus on the relation between trading volume and price change (Crouch, 1970). Therefore, in order to further understand the properties of *OFI*, we also compare the explanatory power of *OFI* and trading volume.

[Insert Table 6 Around Here]

Panel A of Table 6 shows the results obtained using the absolute mid-price change. We find that in 36.59% of instances, *OFI* can be used to explain absolute price change under the 15-minute time horizon, with 34.17% of instances at a daily frequency. The explanatory power of absolute *OFI* is insignificant, due to most of its information being eliminated by taking the absolute value of the measure. However, 80.04% (97.67%) of instances of trading volume pass the *t*-test under the 15-minute (1-day) frequency. Simultaneously, using these two variables, *OFI* and trading volume, in a multiple regression framework, the explanatory power of trading volume decreases dramatically, with only 18.83% and 17.50% of the trading volume coefficients being statistically significant under 15-minute and 1-day frequencies, respectively. As indicated by the above results, we infer that these two variables are related to each other in that they explain similar price change dynamics. However, trading volume only incorporates information on trade magnitude, with *OFI* containing information on both direction and magnitude.

4.5 Forecasting prices with order imbalance

Table 7 shows regression results of how well current order book imbalance can predict subsequent price changes.

[Insert Table 7 Around Here]

Firstly, using mid-price change in Panel A, we find that 29.69% (27.00%) of first order lagged 15-minute (1-day) horizon *OFI* coefficients are said to be statistically significant. Therefore, we infer that a one period lagged measure of *OFI* possesses significant explanatory power. Alternatively, 26.69% (19.82%) of lagged *TI* coefficients are found to be significant at a 15-minute (1-day) frequency. From this we derive that both lag *OFI* and *TI* are useful in predicting future price changes. To test for sensitivity to the type of price aggregation process used we confirm the above results using transaction data, as shown in Panel B.

Specifying a multiple regression comprised of both lagged *TI* and lagged *OFI* explanatory variables, the percentage of significant subsample coefficients for lagged OFI decreases to 22.00% (21.08%) under a 15-minute (1-day) time horizon. Significant lagged *TI* coefficients drop to 20.28% and 14.80%, at 15-minute and 1-day time horizons, respectively. Meanwhile, the Newey-West *t*-statistic produced from the sample of 1105 lagged *TI* estimated coefficients under the hypothesis of a non-zero beta stands at -0.97, and is therefore not significant. Although lagged *OFI* in the multiple regression setup passes the *t*-test under a one-day time horizon, the explanatory power (based on absolute *t*-stat) is much lower than under the 15-minute time horizon. Therefore, we suppose that the longer the interval to calculate incoming order, the lower the levels of predictability for both *OFI* and *TI*. Using transaction data in Panel B leads to qualitatively similar results. Although both lagged *OFI* and *TI* pass the *t*-test, the explanatory power is still much lower for the 1-day horizon than the 15-minute horizon, which is a similar result to the above.

In comparison with Table 5, we find that the coefficients of lagged *OFI* and lagged *TI* are much lower than the coefficients obtained using contemporaneous *OFI* and *TI*. Although the coefficients of lagged *OFI* and lagged *TI* are close to 0, we find using the

Newey-West *t*-test that both the sign of lagged *OFI* and *TI* coefficients are positive when we use 10-second price changes and incoming orders but negative when we use a 1-minute time horizon. These results bring with them two main questions: Does order flow positively or negatively affect the price? For how long does order flow affect the price?

In order to answer these questions, we run a series of regressions; specifying each lag as 10 seconds and testing 1st to 12th order lagged *OFI* variables sequentially. Table 8 shows the lifecycle of the order flow price impact:

[Insert Table 8 Around Here]

In this table, using mid-price, the impact from all lags between the 1st and 9th (except the 3rd lag) are significant, with the first two lags positively impacting the price change, and negative price impacts observed beyond the 4th lag. Furthermore, when using transaction price, the duration and direction of order flow price impact is almost identical with the first two lags positively impacting price change, and lags beyond the 3rd order having a negative relation with price change. Therefore, from the above table, we conclude that 10-second *OFI* impacts price change up to 90 seconds into the future, with it positively impacting price change during the first 20 seconds and adjusting impact direction after 30 seconds. As there is not consistent significance after the 9th lag, our inference is that order flow has a temporary price impact, with *OFI* representing temporary supply and demand inequality in Chinese soybean markets.

In addition, we also divide the above table into several subsamples grouped on a yearly basis as shown in Table 9.

[Insert Table 9 Around Here]

From Table 9, we find the period of time for which 10-second OFI impacts price change in 2008 and 2010 is shorter than in other years. Meanwhile, the contemporaneous price impact coefficient in these two years is higher than in others. This means that during these two years (especially in 2008) there was a higher temporary price impact but that it took less time to adjust this temporary supply and demand inequality. We observe that order flow imbalance in these two years are the lowest amongst all sample years. We suspect it is partly due to the global financial crisis. Another explanation is based on the Shanghai Futures Exchange report (Yang & Xu, 2016), which states that the ratio of individual to institutional investors' turnover in these two years is much higher than in other years. As individual investors in financial markets are typically assumed to provide liquidity, we infer that the above phenomenon is primarily due to the high level of trading activity by the aforementioned liquidity providers during these two years.

4.6 Granger causality between order flow and price change

In this section, we use the Granger causality between any two of, *OFI*, *TI* and price change.[6] Table 10 shows the results of the tests. We calculate the Granger causality results for each day, and report the average result in the table. We use the label "Success" for the percentage of total samples that exhibit significant Granger causality at the 10% level.

[Insert Table 10 Around Here]

From this table, we find that using 10-second frequency the success of *OFI* Granger causing price change is much higher than price change Granger causing *OFI* for both mid

[6] The Granger causality test is used to determine whether a time series is useful in forecasting another. If predictions of an evolving variable Y based on its own past values and the past values of another evolving variable X are better than predictions of Y based only on its own past values, it is said that variable X Granger-causes variable Y. When time series X Granger-causes time series Y, the patterns in X are approximately repeated in Y after some time lag, or more simply X leads Y.

and transaction prices. Meanwhile, the success of *TI* Granger causing price change is also stronger than price change Granger causing *TI*. Therefore, we infer that order imbalance, as measured by both *OFI* and *TI*, leads price change. The above results are consistent with results from prior studies (Brown et al., 1997; Hyun et al., 2016) who found that order book imbalance moves faster than returns. Furthermore, we also test Granger causality between *OFI* and *TI* directly, showing that *OFI* leads *TI*.

In addition, we study Granger causality using the 1-minute data horizon. The results uncovered are similar but weaker than those in the 10-second study, especially the causality observed between *TI* and price change, where it does not indicate that *TI* leads price change. This is perhaps due to information from past order imbalance already being incorporated into price due to the increased time horizon.

5. Conclusion

As one of the most actively traded agricultural futures contracts in the world, the soybean meal contract in China has become increasingly important. In this study, we first utilize *OFI*, a variable derived by Cont et al. (2014) that incorporates the size of order book events, to investigate the relation between order book events and the price of Chinese soybean meal futures contracts. We show that *OFI* has strong power in explaining contemporaneous price changes, and that the impact coefficient is inversely proportional to market depth. Furthermore, we also document that the price is more resilient to incoming orders than indicated by market depth in soybean meal market. In addition, we demonstrate that *OFI* is a more all-encompassing measure relative to *TI* and trading volume, with *TI* and trading volume carrying very little information about price changes in comparison to order flow imbalance. Furthermore, we find that *OFI* captures the transitory mismatch of supply and demand and that our adopted order flow measures are able to effectively extract and exploit that information to forecast subsequent price changes. In summary, we find that the predictive power of *OFI* is stronger than those of *TI*, so we recommend its adoption when studying the order book of soybean meal futures.

Overall, this study provides comprehensive evidence of the relation between order flow and price impact in the Chinese soybean meal markets, equipping us with a greater understanding of the efficiency of the market and its price formation process. The new evidence is important in its own right, in terms of understanding agricultural futures markets in China, but also in contributing to our understanding of emerging markets more generally. Moreover, the findings serve as an out-of-sample test of existing findings and theories of order flow price impact. Our findings will better inform the decisions of policy makers, regulators, and practitioners, and may lead to improved market information efficiency.

Bibliography

Acworth, W. 2016, 2015 Annual Survey: Global Derivatives Volume, Asia Takes the Lead. Available: http://marketvoicemag.org/?q=content/2015-annual-survey-global-derivatives-volume.

Acworth, W. 2015, 2014 FIA Annual Global Futures and Options Volume: Gains in North America and Europe Offset Declines in Asia-Pacific. Available: https://fimag.fia.org/sites/default/files/content_attachments/2014%20FIA%20Annual%20Volume%20Survey%20–%20Charts%20and%20Tables.pdf.

Admati, A.R. & Pfleiderer, P. 1988, "A theory of intraday patterns: Volume and price variability", Review of Financial Studies, vol. 1, no. 1, pp. 3-40.

Bailey, W., Cai, J., Cheung, Y.L. & Wang, F. 2009, "Stock returns, order imbalances, and commonality: Evidence on individual, institutional, and proprietary investors in China", Journal of Banking & Finance, vol. 33, no. 1, pp. 9-19.

Berkman, H. & Lee, J.B.T. 2002, "The effectiveness of price limits in an emerging market: Evidence from the Korean Stock Exchange", Pacific-Basin Finance Journal, vol. 10, no. 5, pp. 517-530.

Booth, G.G., So, R.W. & Tse, Y. 1999, "Price discovery in the German equity index derivatives markets", Journal of Futures Markets, vol. 19, no. 6, pp. 619-643.

Bouchaud, J. 2010, "Price impact", in R. Cont, ed., ‘Encyclopedia of Quantitative Finance’, John Wiley & Sons, Ltd.

Brown, P., Walsh, D. & Yuen, A. 1997, "The interaction between order imbalance and stock price", Pacific-Basin Finance Journal, vol. 5, no. 5, pp. 539-557.

Cabrera, J., Wang, T. & Yang, J. 2009, "Do futures lead price discovery in electronic foreign exchange markets?", Journal of Futures Markets, vol. 29, no. 2, pp. 137-156.

Chang, C.Y. & Shie, F.S. 2011, "The relation between relative order imbalance and intraday futures returns: An application of the quantile regression model to Taiwan", Emerging Markets Finance and Trade, vol. 47, no. 3, pp. 69-87.

Chen, Y. & Gau, Y. 2010, "News announcements and price discovery in foreign exchange spot and futures markets", Journal of Banking & Finance, vol. 34, no. 7, pp. 1628-1636.

Chordia, T., Roll, R. & Subrahmanyam, A. 2008, "Liquidity and market efficiency", Journal of Financial Economics, vol. 87, no. 2, pp. 249-268.

Chordia, T., Roll, R. & Subrahmanyam, A. 2002, "Order imbalance, liquidity, and market returns", Journal of Financial Economics, vol. 65, no. 1, pp. 111-130.

Chordia, T. & Subrahmanyam, A. 2004, "Order imbalance and individual stock returns: Theory and evidence", Journal of Financial Economics, vol. 72, no. 3, pp. 485-518.

Clark, P.K. 1973, "A subordinated stochastic process model with finite variance for speculative prices", Econometrica: Journal of the Econometric Society, pp. 135-155.

Cont, R., Kukanov, A. & Stoikov, S. 2014, "The price impact of order book events", Journal of financial econometrics, vol. 12, no. 1, pp. 47-88.

Copeland, T.E. 1976, "A Model of Asset Trading under the Assumption of Sequential Information Arrival", Journal of Finance, vol. 31, no. 4, pp. 1149-1168.

Covrig, V., Ding, D.K. & Low, B.S. 2004, "The contribution of a satellite market to price discovery: Evidence from the Singapore exchange", Journal of Futures Markets, vol. 24, no. 10, pp. 981-100

Crouch, R.L. 1970, "The volume of transactions and price changes on the New York Stock Exchange", Financial Analysts Journal, vol. 26, no. 4, pp. 104-109.

CSRC 2009, China Capital Markets Development Report 2008. Available: http://www.csrc.gov.cn/pub/csrc_en/Informations/publication/200911/P020091103520222505841.pdf [3 July 2017].

DEC 2017, DEC History. Available: http://www.dce.com.cn/DCEENMO/About_Us3/DCE_History3/index.html [3 July 2017].

Epps, T.W. & Epps, M.L. 1976, "The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis", Econometrica: Journal of the Econometric Society, pp. 305-321.

Farmer, J.D., Gillemot, L., Lillo, F., Mike, S. & Sen, A. 2004, "What really causes large price changes?", Quantitative Finance, vol. 4, no. 4, pp. 383-397.

Fricke, C. & Menkhoff, L. 2011, "Does the “Bund” dominate price discovery in Euro bond futures? Examining information shares", Journal of Banking & Finance, vol. 35, no. 5, pp. 1057-1072.

Guéant, O. 2015, "Optimal execution and block trade pricing: a general framework", Applied Mathematical Finance, vol. 22, no. 4, pp. 336-365.

Gurgul, H., Majdosz, P. & Mestel, R. 2005, "Joint dynamics of prices and trading volume on the Polish stock market", Managing Global Transitions, vol. 3, no. 2, pp. 139.

Harris, L. & Gurel, E. 1986, "Price and volume effects associated with changes in the S&P 500 list: New evidence for the existence of price pressures", The Journal of Finance, vol. 41, no. 4, pp. 815-829.

Harris, M. & Raviv, A. 1993, "Differences of opinion make a horse race", Review of Financial Studies, vol. 6, no. 3, pp. 473-506.

Hasbrouck, J. & Seppi, D.J. 2001, "Common factors in prices, order flows, and liquidity", Journal of Financial Economics, vol. 59, no. 3, pp. 383-411.

He, L., Yang, S., Xie, W. & Han, Z. 2014, "Contemporaneous and asymmetric properties in the price-volume relationships in China's agricultural futures markets", Emerging Markets Finance and Trade, vol. 50, no. sup1, pp. 148-166.

Ho, R.Y., Cheung, Y. & Cheung, D.W. 1993, "Intraday prices and trading volume relationship in an emerging Asian market-Hong Kong", Pacific-Basin Finance Journal, vol. 1, no. 2, pp. 203-214.

Huang, H. 2011, "An analysis of intraday return–order imbalance relation to stealth trading", Investment Management and Financial Innovations, vol. 8, no. 1, pp. 234-242.

Huang, P., Ni, Y. & Yu, C. 2012, "The microstructure of the price-volume relationship of the constituent stocks of the Taiwan 50 Index", Emerging Markets Finance and Trade, vol. 48, no. sup2, pp. 153-168.

Hyun, C., Park, J. & Lee, K. 2016, "The effect of limit order flows at the best quotes on price changes", Risk and Decision Analysis, vol. 6, no. 1, pp. 23-36.

Kaniel, R., Saar, G. & Titman, S. 2008, "Individual investor trading and stock returns", The Journal of Finance, vol. 63, no. 1, pp. 273-310.

Karpoff, J.M. 1987, "The relation between price changes and trading volume: A survey", Journal of Financial and Quantitative Analysis, vol. 22, no. 01, pp. 109-126.

Knez, P.J. & Ready, M.J. 1996, "Estimating the profits from trading strategies", Review of Financial Studies, vol. 9, no. 4, pp. 1121-1163.

Kyle, A.S. 1985, "Continuous auctions and insider trading", Econometrica: Journal of the Econometric Society, pp. 1315-1335.

Lao, L., Tian, S. & Zhao, Q. 2017, "Will Order Imbalances Predict Stock Returns in Extreme Market Situations?: Evidence from China", Emerging Markets Finance and Trade, forthcoming.

Lee, C. & Ready, M.J. 1991, "Inferring trade direction from intraday data", The Journal of Finance, vol. 46, no. 2, pp. 733-746.

Lee, Y., Liu, Y., Roll, R. & Subrahmanyam, A. 2004, "Order imbalances and market efficiency: Evidence from the Taiwan Stock Exchange", Journal of Financial and Quantitative Analysis, vol. 39, no. 02, pp. 327-341.

Narayan, P.K., Narayan, S. & Westerlund, J. 2015, "Do order imbalances predict Chinese stock returns? New evidence from intraday data", Pacific-Basin Finance Journal, vol. 34, pp. 136-151.

Obizhaeva, A.A. & Wang, J. 2013, "Optimal trading strategy and supply/demand dynamics", Journal of Financial Markets, vol. 16, no. 1, pp. 1-32.

Rastogi, N., Reddy, V. & Kotha, K.K. 2013, "Order imbalance and returns: evidence from India", International Journal of Managerial Finance, vol. 9, no. 2, pp. 92-109.

Shalen, C.T. 1993, "Volume, volatility, and the dispersion of beliefs", Review of Financial Studies, vol. 6, no. 2, pp. 405-434.

Shastri, K., Thirumalai, R.S. & Zutter, C.J. 2008, "Information revelation in the futures market: Evidence from single stock futures", Journal of Futures Markets, vol. 28, no. 4, pp. 335-353.

Shenoy, C. & Zhang, Y.J. 2007, "Order imbalance and stock returns: Evidence from China", The Quarterly Review of Economics and Finance, vol. 47, no. 5, pp. 637-650.

Tauchen, G.E. & Pitts, M. 1983, "The price variability-volume relationship on speculative markets", Econometrica: Journal of the Econometric Society, pp. 485-505.

Wang, X., Ye, Q. & Zhao, F. 2016, "Trading activity and price behavior in Chinese agricultural futures markets", Finance Research Letters, vol.18, pp.52-59.

Weber, P. & Rosenow, B. 2006, "Large stock price changes: volume or liquidity?", Quantitative Finance, vol. 6, no. 1, pp. 7-14.

Yang, J.M. & Xu, S. Y. 2016, Institutional Investors Cultivation and Development of China Futures Markets. Available: http://www.shfe.com.cn/upload/20160816/1471335461440.pdf [11 May 2016].

Table 1
Trading Frequency

This table reports Chinese soybean meal trading frequencies. Each trading day is divided into multiple fixed time intervals. The first column is the time interval and it ranges from 10 seconds to 20 minutes. The second column presents the volume based non-trading probability in the Chinese soybean meal futures. The sample period is from June 2008 to June 2013.

Interval	Non-Trading Probability (%)
10-second	1.52%
20-second	0.72%
30-second	0.52%
1-minute	0.37%
2-minute	0.32%
3-minute	0.30%
5-minute	0.29%
10-minute	0.24%
15-minute	0.22%
20-minute	0.20%

Table 2
Summary Statistics

Panel A presents summary statistics for the limit order book of the Chinese soybean meal futures contract. All values are calculated from the filtered data outlined in Section 2. Panel B presents the summary statistics (mean, standard deviation, top and bottom five percentiles, quartiles and median) of the main variables constructed and used in the analysis. ΔP_k is the mid-price or transaction-price change over period k. OFI_k is order flow imbalance, TI_k is trade imbalance, and VOL_k^H is trading volume adjust by H, where H comes from $\log|\Delta P_k|=\log(\alpha_i)+H_i\log(VOL_k)+\log(\varepsilon_k)$. k denotes the short time interval; k can be 10 seconds or 1 minute. The sample period is June 2008 to June 2013.

Panel A: Limit order books	
Variable	Mean
Price	3,162.37
Daily trading volume	998,376.92
Average price change of best bid	1.08
Average price change of best ask	1.08
Average bid depth	411.89
Average ask depth	418.30
Average spread	1.02
Maximum spread	38

Panel B: Main variables								
Variable								
k = 10 sec	N	Mean	Std	P5	P25	Median	P75	P95
ΔP_k (Mid-Price)	1,467,720	0.000	1.721	-1.000	0.000	0.000	0.000	1.000
ΔP_k (Transaction-Price)	1,467,720	0.000	1.796	-1.000	0.000	0.000	0.000	1.000
OFI_k	1,467,718	1.206	632.264	-863	-137	0.000	136	873
TI_k	1,466,093	-3.711	995.412	-978	-140	0.000	130	964
VOL_k^H	1,462,811	7.400	93.020	0.948	1.035	1.625	3.009	15.571
k = 1 min								
ΔP_k (Mid-Price)	248,060	-0.002	4.239	-3.000	-1.000	0.000	1.000	3.000
ΔP_k (Transaction-Price)	248,060	-0.002	4.307	-3.000	-1.000	0.000	1.000	3.000
OFI_k	248,060	7.028	1,882	-2,743	-573	-1.000	564	2,778
TI_k	247,995	-23.173	2,983	-3,552	-714	-6.000	666	3,492
VOL_k^H	248,059	3,708	662,574	2.780	7.934	21.035	56.941	257.534

Table 3
Relation between Price Change and Order Flow Imbalance

This table reports the result of a regression between price change and order flow imbalance:

$\Delta P_k=\alpha_i+\beta_i OFI_k+\varepsilon_k$

where *k* is the short interval and ΔP_k are the price changes during time period *k*. In panel A, price change is calculated by mid-price. In panel B, price change is calculated using transaction price. OFI_k is the contemporaneous order flow imbalance, *i* denotes the long intervals, and we assume that the price impact coefficient β is constant over each *i*. When *i* is equal to 15 minutes and 1 day, we have 16,417 and 1,105 subsamples, respectively. Panels C and D repeat the analyses with control of one lag of dependent variable. All columns present the average across these subsamples. The *t*-statistics are computed using Newey-West standard errors. The aggregated testing columns report the percentage of samples where the coefficient passes the test at a 10% significance level.

Panel A: Mid-Price

$\Delta P_k=\alpha_i+\beta_i OFI_k+\varepsilon_k$		Average Results					Aggregated testing	
k	i	α	t(α)	β_1	$t(\beta_1)$	R^2	α≠0	β≠0
10 sec	15 min	-0.002	(-1.12)	0.002	(51.73)***	0.4648	2.75%	97.87%
1 min	1 day	-0.026	(-1.20)	0.002	(17.06)***	0.4767	4.57%	97.58%

Panel B: Transaction Price

$\Delta P_k=\alpha_i+\beta_i OFI_k+\varepsilon_k$		Average Results					Aggregated testing	
k	i	α	t(α)	β_1	$t(\beta_1)$	R^2	α≠0	β≠0
10 sec	15 min	-0.002	(-1.19)	0.001	(56.35)***	0.2606	1.00%	94.00%
1 min	1 day	-0.027	(-1.22)	0.002	(17.62)***	0.4516	4.66%	97.31%

Panel C: Mid-Price with Lag

$\Delta P_k=\alpha_i+\beta_i OFI_k+c_i lag(\Delta P_k)+\varepsilon_k$		Average Results							Aggregated testing		
k	i	α	t(α)	β_1	$t(\beta_1)$	c_1	$t(c_1)$	R^2	α≠0	β≠0	c≠0
10 sec	15 min	0.001	(1.88)*	0.002	(52.21)***	-0.137	(-133.11)***	51.65%	2.81%	98.64%	59.95%
1 min	1 day	-0.001	(-0.39)	0.002	(17.15)***	-0.063	(-28.01)***	71.48%	8.17%	100.00%	55.21%

Panel D: Transaction Price with Lag

$\Delta P_k=\alpha_i+\beta_i OFI_k+c_i lag(\Delta P_k)+\varepsilon_k$		Average Results							Aggregated testing		
k	i	α	t(α)	β_1	$t(\beta_1)$	c_1	$t(c_1)$	R^2	α≠0	β≠0	c≠0
10 sec	15 min	0.000	(1.48)	0.001	(58.59)***	0.043	(32.12)***	29.15%	0.49%	95.15%	23.03%
1 min	1 day	-0.001	(-0.28)	0.002	(17.72)***	-0.067	(-27.64)***	65.04%	4.94%	99.82%	54.13%

Table 4
Relation between Price Impact and Market Depth

This table reports the relation between the price impact coefficient and market depth:
$\log\beta=\alpha-\lambda\log AD+\varepsilon_k$
$\beta=\alpha+c/AD^{\lambda}+\varepsilon_k$
Where β_i is the price impact coefficient for the *i*-th subsample and AD_i is the average market depth for that subsample. When *i* is equal to 15 minutes and 1 day, we have 16,417 and 1,105 estimates of β, respectively. The second regression uses estimates λ which are obtained from the first regression. The *t*-statistics are computed using Newey-West standard errors.

Panel A: Mid-Price											
		$\log\beta=\alpha-\lambda\log AD+\varepsilon_k$					$\beta=\alpha+c/AD^{\lambda}+\varepsilon_k$				
k	i	α	t(α)	λ	t(λ)	R^2	α	t(α)	c	t(c)	R^2
10 sec	15 min	-1.872	(-36.76)***	0.827	(99.01)***	49.16%	0.000	(0.16)	0.161	(45.20)***	46.96%
1 min	1 day	-1.509	(-7.26)***	0.885	(26.02)***	83.70%	0.000	(0.84)	0.215	(14.50)***	59.12%
Panel B: Transaction Price											
		$\log\beta=\alpha-\lambda\log AD+\varepsilon_k$					$\beta=\alpha+c/AD^{\lambda}+\varepsilon_k$				
k	i	α	t(α)	λ	t(λ)	R^2	α	t(α)	c	t(c)	R^2
10 sec	15 min	-2.049	(-35.40)***	0.814	(85.35)***	53.39%	0.000	(1.75)*	0.133	(41.96)***	41.58%
1 min	1 day	-1.509	(-21.73)***	0.885	(75.42)***	83.70%	0.000	(9.91)***	0.164	(40.12)***	59.12%

Table 5
Comparison of Order Flow Imbalance and Trade Imbalance

This table reports regression results relating order flow imbalance and trade imbalance to price change:

$\Delta P_k=\alpha_i+\beta_i OFI_k+\varepsilon_k$

$\Delta P_k=\alpha_i+\beta_i TI_k+\varepsilon_k$

$\Delta P_k=\alpha_i+\beta_i OFI_k+\beta_i TI_k+\varepsilon_k$

Where k is the short time interval and ΔP_k is the mid-price change over period k. OFI_k is the contemporaneous order flow imbalance, TI_k is the contemporaneous trade imbalance, i denotes the long time interval, and we assume that the price impact coefficient β is constant over each i. When i is equal to 15-minute and 1 day, we have 16,417 and 1,105 subsamples, respectively. All columns present the average across these subsamples. The *t*-statistics are computed using Newey-West standard errors. The aggregated testing columns report the percentage of samples where the coefficient passes the test at a 10% significance level.

Panel A: Mid-Price																
		Order flow imbalance				Trade imbalance				Both covariates						
k	i	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	β_2	$t(\beta_2)$	$\beta\neq0$	R^2
10 sec	15 min	0.002	(51.73)***	97.87%	46.48%	0.001	(35.65)***	93.56%	24.96%	0.001	(42.94)***	93.84%	0.000	(16.29)***	31.41%	49.24%
1 min	1 day	0.002	(17.06)***	97.58%	47.67%	0.001	(15.73)***	91.48%	30.62%	0.001	(17.29)***	91.93%	0.000	(4.80)***	22.33%	50.08%
Panel B: Transaction Price																
		Order flow imbalance				Trade imbalance				Both covariates						
k	i	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	β_2	$t(\beta_2)$	$\beta\neq0$	R^2
10 sec	15 min	0.001	(56.35)***	94.00%	26.06%	0.001	(29.39)***	88.48%	1.61%	0.001	(66.96)***	80.87%	0.001	(15.51)***	28.80%	28.41%
1 min	1 day	0.002	(17.62)***	97.31%	45.16%	0.001	(15.38)***	91.93%	28.39%	0.001	(18.35)***	91.84%	0.000	(4.39)***	21.70%	47.36%

Table 6
Comparison of Traded Volume and Order Flow Imbalance

This table reports the regression results relating traded volume and order flow imbalance to price change:

$|\Delta P_k|=\alpha_i+\beta_i|OFI_k|+\varepsilon_k$

$|\Delta P_k|=\alpha_i+\beta_i VOL_K{}^H+\varepsilon_k$

$|\Delta P_k|=\alpha_i+\beta_{i1}|OFI_k|+\beta_{i2}VOL_K{}^H+\varepsilon_k$

Where k is the short interval and ΔP_k is the mid-price change during time period k. OFI_k is the contemporaneous order flow imbalance, $VOL_K{}^H$ is the contemporaneous trading volume adjusted by H, where H comes from $\log|\Delta P_k|=\log(\alpha_i)+H_i\log(VOL_K)+\log(\varepsilon_k)$. i denotes the long interval and we assume that the price impact coefficient β is constant over each i. When i is equal to 15 minutes, we have 16,417 subsamples and when i is equal to a day, we have 1,105 subsamples. All columns present the average across these subsamples. The t-statistics are computed using Newey-West standard errors. The aggregated testing columns reports the percentage of samples where the coefficient passes the test at a 10% significance level.

Panel A: Mid-Price																	
		Order flow imbalance				Trading volume				Both covariates							
k	i	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	β_2	$t(\beta_2)$	$\beta\neq0$	R^2
10 sec	15 min	0.001	(1.58)	36.59%	32.30%	5.72E+06	(1.00)	80.04%	17.84%	0.002	(1.82)*	21.87%	6.25E+06	(1.00)	18.83%	36.20%
1 min	1 day	0.003	(1.97)**	34.17%	34.22%	0.142	(2.27)**	97.67%	28.46%	-0.001	(-0.57)	15.98%	-0.473	(-0.69)	17.50%	41.29%
Panel B: Transaction Price																
		Order flow imbalance				Trading volume				Both covariates						
k	i	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	β_2	$t(\beta_2)$	$\beta\neq0$	R^2
10 sec	15 min	0.000	(0.05)	21.99%	12.18%	2.73E+04	(1.00)	55.03%	6.99%	0.004	(1.31)	14.86%	2.79E+04	(1.00)	8.67%	14.40%
1 min	1 day	0.003	(1.68)*	31.48%	29.16%	-0.599	(-0.61)	96.59%	22.16%	0.002	(1.50)	15.43%	-1.194	(-0.93)	11.39%	34.39%

Table 7
Relation between Price Change and First Order Lagged Order Flow Imbalance

This table reports the relation between price change and lag order flow imbalance:

$\Delta P_k=\alpha_i+\beta_i lag(OFI_k)+\varepsilon_k$

$\Delta P_k=\alpha_i+\beta_i lag(TI_k)+\varepsilon_k$

$\Delta P_k=\alpha_i+\beta_{i1} lag(OFI_k)+\beta_{i2} lag(TI_k)+\varepsilon_k$

Where k is the short interval, and ΔP_k is the mid-price change during time period k. OFI_k is the contemporaneous order flow imbalance, TI_k is the contemporaneous trade imbalance. i denotes the long interval, and we assume that the price impact coefficient β is constant over each i. When i is equal to 15 minutes and 1 day, we have 16,417 and 1,105 subsamples, respectively. All columns present the average across these subsamples. The *t*-statistics are computed using Newey-West standard errors. The aggregated testing columns report the percentage of samples where the coefficient passes the test at a 10% significance level.

Panel A: Mid-Price																	
		Order flow imbalance				Trade imbalance				Both covariates							
k	i	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	β_2	$t(\beta_2)$	$\beta\neq0$	R^2
10 sec	15 min	0.000	(17.77)***	29.69%	2.38%	0.000	(11.52)***	26.69%	2.50%	0.000	(9.44)***	22.00%	0.000	(6.36)***	20.28%	4.28%
1 min	1 day	0.000	(-3.75)***	27.00%	1.26%	0.000	(-4.82)***	19.82%	1.00%	0.000	(-2.02)**	21.08%	0.000	(-0.97)	14.80%	2.07%
Panel B: Transaction Price																
		Order flow imbalance				Trade imbalance				Both covariates						
k	i	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	R^2	β_1	$t(\beta_1)$	$\beta\neq0$	β_2	$t(\beta_2)$	$\beta\neq0$	R^2
10 sec	15 min	0.000	(28.09)***	38.86%	2.77%	0.000	(3.83)***	21.18%	1.59%	0.001	(25.47)***	38.23%	0.000	(-9.87)***	21.67%	4.37%
1 min	1 day	0.000	(-8.25)***	37.22%	1.61%	0.000	(-10.07)***	31.93%	1.38%	0.000	(-1.83)*	22.06%	0.000	(-5.11)***	15.61%	2.43%

Table 8
The Lifespan of Order Flow Price Impact

This table reports the time duration of order flow imbalance on price change:
$\Delta P_k = \alpha_i + \beta_i lag_{pre}(OFI_k) + \varepsilon_k$
Where k is the short time interval and ΔP_k is the price change during time period k. OFI_k is the contemporaneous order flow imbalance and i is the long time interval. pre is the lag order of OFI_k, with each lag being 10 seconds apart. In total there are 16,417 subsamples. All columns present the average across these subsamples. The t-statistics are computed using Newey-West standard errors. The aggregated testing columns report the percentage of samples where the coefficient passes the test at a 10% significance level.

lag	Mid-Price			Transaction Price		
	β_1	$t(\beta_1)$	$\beta \neq 0$	β_1	$t(\beta_1)$	$\beta \neq 0$
0	0.00162	(51.73)***	97.87%	0.00146	(56.35)***	94.00%
1	0.00020	(17.77)***	29.69%	0.00036	(28.09)***	38.86%
2	0.00005	(6.13)***	16.93%	0.00005	(6.41)***	16.23%
3	0.00000	(0.46)	14.84%	-0.00002	(-3.21)***	15.96%
4	-0.00003	(-3.85)***	14.75%	-0.00006	(-6.70)***	16.84%
5	-0.00004	(-5.00)***	14.73%	-0.00007	(-7.47)***	16.33%
6	-0.00005	(-7.76)***	15.21%	-0.00007	(-9.22)***	16.33%
7	-0.00004	(-5.61)***	14.54%	-0.00008	(-8.12)***	16.10%
8	-0.00003	(-4.00)***	14.29%	-0.00005	(-5.59)***	15.08%
9	-0.00003	(-3.63)***	13.47%	-0.00005	(-4.99)***	15.03%
10	-0.00002	(-1.86)*	14.13%	-0.00002	(-1.56)	15.04%
11	-0.00001	(-0.89)	13.82%	-0.00001	(-1.03)	14.93%
12	-0.00002	(-2.16)**	13.75%	-0.00002	(-2.38)**	15.13%

Table 9
The Lifespan of Order Flow Price Impact By Year

This table reports the time duration of order flow imbalance on price change:

$\Delta P_k=\alpha_i+\beta_i lag_{pre}(OFI_k)+\varepsilon_k$

Where k is the short time interval and ΔP_k is the price change during time period k. OFI_k is the contemporaneous order flow imbalance and i is the long time interval. pre is the lag order of OFI_k, with each lag being 10 seconds apart. In total there are 16,417 subsamples. All columns present the average across these subsamples. The *t*-statistics are computed using Newey-West standard errors. The aggregated testing columns report the percentage of samples where the coefficient passes the test at a 10% significance level. Mid-price is used to calculate price change in panel A, with transaction price used in panel B.

Panel A: Mid-Price

lag	2008			2009			2010			2011			2012			2013		
	β_1	$t(\beta_1)$	$\beta\neq0$	β_1	$t(\beta_1)$	$\beta\neq0$	β_1	$t(\beta_1)$	$\beta\neq0$	β_1	$t(\beta_1)$	$\beta\neq0$	β_1	$t(\beta_1)$	$\beta\neq0$	β_1	$t(\beta_1)$	$\beta\neq0$
0	0.00280	(40.01)***	98.48%	0.00161	(42.59)***	98.73%	0.00229	(17.44)***	94.99%	0.00177	(66.49)***	97.95%	0.00078	(71.82)***	98.92%	0.00072	(58.88)***	99.29%
1	0.00037	(4.36)***	31.45%	0.00017	(19.26)***	24.84%	0.00023	(5.83)***	25.92%	0.00027	(32.38)***	32.38%	0.00011	(32.02)***	33.50%	0.00008	(23.16)***	29.98%
2	-0.00001	(-0.18)	16.76%	0.00002	(2.38)**	17.33%	0.00008	(3.31)***	15.36%	0.00009	(14.31)***	18.33%	0.00003	(10.34)***	17.17%	0.00003	(8.32)***	16.20%
3	-0.00004	(-1.45)	14.69%	-0.00002	(-4.53)***	14.96%	0.00003	(1.22)	14.89%	0.00003	(4.61)***	14.22%	-0.00001	(-2.97)***	15.42%	0.00000	(0.08)	14.78%
4	-0.00005	(-1.06)	16.90%	-0.00004	(-7.99)***	16.13%	-0.00002	(-0.66)	13.77%	-0.00002	(-4.10)***	12.81%	-0.00002	(-7.75)***	15.81%	-0.00001	(-5.26)***	14.55%
5	0.00000	(-0.02)	15.03%	-0.00005	(-9.31)***	15.42%	-0.00005	(-2.61)***	13.48%	-0.00006	(-3.42)***	13.56%	-0.00003	(-10.91)***	16.08%	-0.00002	(-9.04)***	15.49%
6	-0.00016	(-3.65)***	15.93%	-0.00007	(-12.45)***	16.65%	-0.00006	(-2.90)***	14.30%	-0.00002	(-1.56)	13.81%	-0.00003	(-11.31)***	16.08%	-0.00003	(-9.47)***	15.08%
7	-0.00010	(-2.74)***	15.52%	-0.00005	(-8.98)***	15.14%	-0.00002	(-0.90)	13.59%	-0.00004	(-4.32)***	13.31%	-0.00002	(-8.94)***	15.94%	-0.00002	(-8.57)***	14.31%
8	-0.00007	(-1.31)	15.17%	-0.00005	(-8.69)***	14.75%	-0.00002	(-0.60)	13.65%	-0.00004	(-7.02)***	13.78%	-0.00002	(-9.85)***	14.69%	-0.00002	(-8.96)***	14.37%
9	-0.00004	(-0.53)	15.17%	-0.00003	(-4.97)***	14.68%	-0.00003	(-1.69)*	11.88%	-0.00005	(-7.70)***	12.89%	-0.00002	(-3.89)***	13.75%	-0.00002	(-7.52)***	13.90%
10	0.00007	(0.66)	15.17%	-0.00002	(-3.15)***	14.04%	-0.00003	(-1.68)**	13.65%	-0.00004	(-7.55)***	14.22%	-0.00002	(-6.56)***	14.39%	-0.00001	(-4.82)***	13.66%
11	0.00010	(0.81)	13.86%	0.00000	(-0.03)	13.94%	-0.00003	(-1.29)	12.24%	-0.00005	(-7.09)***	13.36%	-0.00001	(-2.77)***	15.33%	-0.00001	(-5.56)***	14.61%
12	0.00002	(0.45)	15.24%	-0.00001	(-2.36)**	14.57%	0.00000	(0.01)	12.74%	-0.00006	(-2.09)**	12.81%	-0.00002	(-5.31)***	14.47%	-0.00001	(-4.63)***	13.66%

Panel B: Transaction Price

lag	2008			2009			2010			2011			2012			2013		
	β1	t(β1)	β≠0	β1	t(β1)	β≠0	β1	t(β1)	β≠0	β1	t(β1)	β≠0	β1	t(β1)	β≠0	β1	t(β1)	β≠0
0	0.00265	(41.01)***	98.28%	0.00150	(42.32)***	97.46%	0.00191	(18.20)***	87.41%	0.00160	(62.31)***	91.05%	0.00075	(78.99)***	96.47%	0.00068	(56.94)***	98.82%
1	0.00069	(7.94)***	50.55%	0.00038	(31.88)***	43.01%	0.00044	(9.82)***	32.56%	0.00041	(35.24)***	36.48%	0.00017	(40.22)***	39.47%	0.00014	(31.87)***	38.32%
2	-0.00003	(-0.54)	16.34%	0.00000	(0.42)	16.09%	0.00016	(5.51)***	15.92%	0.00008	(9.13)***	16.86%	0.00002	(4.36)***	16.53%	0.00002	(4.61)***	14.96%
3	-0.00009	(-2.95)***	17.17%	-0.00006	(-9.15)***	16.02%	0.00002	(0.67)	15.72%	-0.00001	(-1.49)	15.55%	-0.00003	(-6.94)***	16.22%	-0.00002	(-4.88)***	15.61%
4	-0.00011	(-2.13)**	18.69%	-0.00009	(-13.27)***	17.33%	-0.00004	(-1.28)	17.34%	-0.00006	(-6.94)***	15.16%	-0.00004	(-11.76)***	17.22%	-0.00003	(-8.74)***	16.26%
5	-0.00006	(-0.85)	17.86%	-0.00008	(-12.93)***	15.42%	-0.00006	(-2.30)**	15.54%	-0.00010	(-5.87)***	16.33%	-0.00004	(-11.20)***	17.33%	-0.00004	(-10.29)***	16.09%
6	-0.00019	(-4.23)***	20.14%	-0.00009	(-13.16)***	16.06%	-0.00007	(-2.81)***	15.69%	-0.00005	(-3.24)***	15.69%	-0.00004	(-11.61)***	15.94%	-0.00004	(-11.36)***	17.09%
7	-0.00012	(-3.25)***	15.72%	-0.00006	(-9.13)***	15.95%	-0.00014	(-3.60)***	15.51%	-0.00006	(-5.38)***	16.19%	-0.00004	(-10.18)***	16.75%	-0.00003	(-7.92)***	16.26%
8	-0.00008	(-1.51)	15.17%	-0.00007	(-10.39)***	15.46%	-0.00007	(-1.82)*	14.54%	-0.00005	(-6.25)***	14.50%	-0.00003	(-7.11)***	15.53%	-0.00002	(-6.01)***	15.85%
9	-0.00004	(-0.49)	14.28%	-0.00003	(-4.28)***	15.63%	-0.00007	(-2.90)***	14.80%	-0.00007	(-8.79)***	14.94%	-0.00002	(-4.58)***	15.25%	-0.00002	(-6.93)***	14.84%
10	0.00006	(0.58)	14.00%	-0.00001	(-2.43)**	14.68%	-0.00002	(-0.60)	14.30%	-0.00005	(-5.18)***	16.30%	-0.00002	(-5.87)***	15.08%	-0.00002	(-4.88)***	15.32%
11	0.00010	(0.81)	14.41%	-0.00001	(-1.58)	16.02%	-0.00003	(-1.05)	14.33%	-0.00005	(-5.89)***	14.80%	-0.00001	(-2.03)**	15.42%	-0.00001	(-3.76)***	14.02%
12	0.00002	(0.51)	14.55%	-0.00001	(-2.06)***	15.24%	0.00001	(0.27)	15.28%	-0.00008	(-2.70)***	15.36%	-0.00002	(-4.67)***	14.92%	-0.00001	(-4.18)***	15.14%

Table 10
Granger Causality

This table reports the Granger causality between price change and order imbalance (both *OFI* and *TI*) in the Chinese soybean meal futures market for different time intervals. We calculate the Granger causality for each day, and average the results. We have 16,417 and 1,105 subsamples under 10 second and 1-minute frequencies, respectively. Mid-price is used to calculate price change in panel A, with transaction price used in panel B. The Granger causality between OFI and TI is also reported in panel C.

Panel A: Mid Price and Order imbalance				
	10 sec		1 min	
H_0	χ^2	Success	χ^2	Success
Price change Granger Cause *OFI*	2.46	21.95%	4.23	29.66%
OFI Granger Cause Price change	9.69	71.66%	5.49	42.00%
Price change Granger Cause *TI*	3.69	31.37%	4.77	32.55%
TI Granger Cause Price change	6.59	52.83%	4.51	35.28%
Panel B: Transaction price and Order imbalance				
	10 sec		1 min	
H_0	χ^2	Success	χ^2	Success
Price change Granger Cause *OFI*	1.84	16.34%	4.05	30.32%
OFI Granger Cause Price change	10.07	81.75%	6.5	48.20%
Price change Granger Cause *TI*	3.15	28.07%	4.71	35.58%
TI Granger Cause Price change	4.15	43.34%	4.31	33.27%
Panel C: OFI and TI				
	10 sec		1 min	
H_0	χ^2	Success	χ^2	Success
TI Granger Cause OFI	3.84	28.68%	3.41	26.59%
OFI Granger Cause TI	7.07	54.97%	4.95	43.20%

Figure 1: Average 15-minute price change, OFI and TI.

This figure reports the average 15-minute mid-price change *ΔP*, Order Flow Imbalance *OFI* and Trade Imbalance for Chinese Soybean Meal futures.

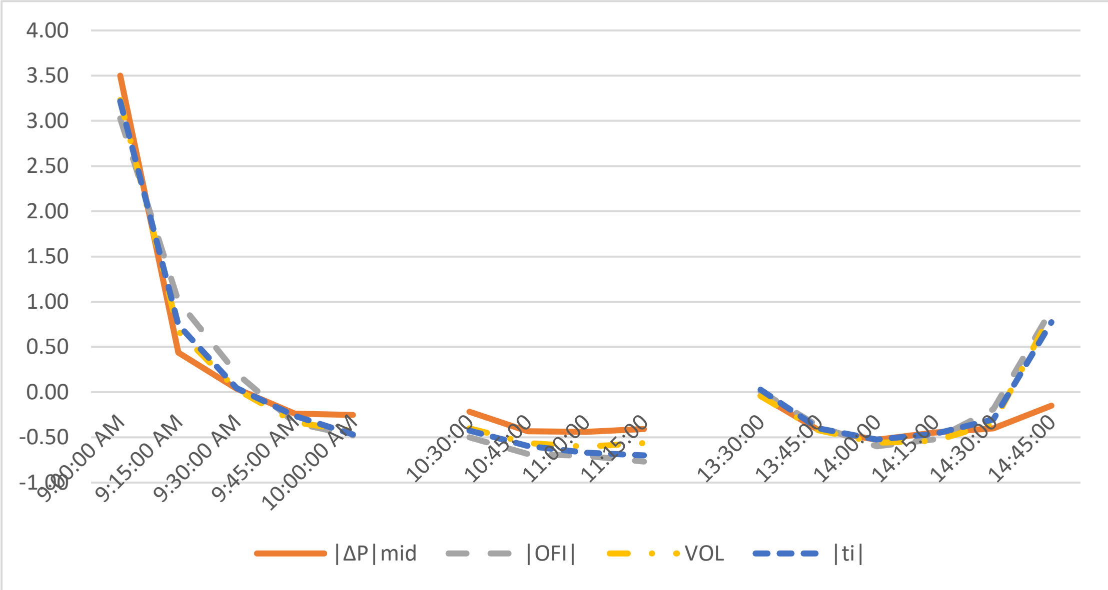

Figure 2: Average 15-minute absolute price change, absolute OFI and trading volume.

This figure reports the average 15-minute absolute price change $|\Delta P|$, absolute Order Flow Imbalance $|OFI|$, absolute Trade Imbalance $|TI|$ and trading volume for Chinese Soybean Meal futures.

Table A1

Global Futures and Options Volume Contract Rankings

This table reports the World ranking of Agricultural Futures & Option Contracts in 2014 and 2015. The trading volume of the Chinese soybean meal futures contract is ranked second among the Top 10 Agricultural Futures & Options Contracts in 2014 and first in 2015. The trading volume and change are also shown in this table.

Top 10 Agricultural Futures & Option Contracts						
Rank		Contract	Contract Size	Volume		Change
2015	2014			2015	2014	(%)
1	2	Soy Meal Futures, DCE	10 tons	289,496,780	204,988,746	41.2%
2	1	Rapeseed Meal Futures, ZCE	10 tons	261,487,209	303,515,966	-13.8%
3	3	White Sugar Futures, ZCE	10 tons	187,323,456	97,726,662	91.7%
4	5	Palm Oil Futures, DCE	10 tons	111,515,010	79,996,388	39.4%
5	7	Soy Oil Futures, DCE	10 tons	92,504,264	64,082,631	44.4%
6	6	Corn Futures, CBOT	5,000 bushels	83,094,271	69,437,304	19.7%
7	4	Rubber Futures, SHFE	10 tons	83,067,547	88,631,586	-6.3%
8	8	Soybean Futures, CBOT	5,000 bushels	54,095,051	49,169,361	10.0%
9	20+	Corn Futures, DCE	10 tons	42,090,235	9,329,939	351.1%
10	12	Sugar#11 Futures, ICE Futures U.S.	112,000 pounds	34,394,482	29,396,597	17.0%