Cost of Immediacy during Large Price Movements: Evidence from Corn Futures Market

by

Xinyue He, Teresa Serra and Philip Garcia

Suggested citation format:

He, X., T. Serra, and P. Garcia. 2018. "Cost of Immediacy during Large Price Movements: Evidence from Corn Futures Market." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Minneapolis, MN. [http://www.farmdoc.illinois.edu/nccc134].

Cost of Immediacy during Large Price Movements: Evidence from Corn Futures Market

Xinyue He, Teresa Serra and Philip Garcia*

Paper presented at the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management
Minneapolis, Minnesota, April 16-17, 2018

* Xinyue He is a doctoral student, Teresa Serra is an Associate Professor and Philip Garcia is the T.A. Hieronymus Distinguished Chair in Futures Markets, all in the Department of Agricultural and Consumer Economics, University of Illinois, 335 Mumford Hall, Urbana, IL 61801, USA
E-mail: xinyueh4@illinois.edu

Cost of Immediacy during Large Price Movements: Evidence from

Corn Futures Market

Recent years have witnessed growing presence of intra-day large price movements in corn futures market. This paper focuses on the behavior of bid-ask spread, a gauge for the cost of immediacy, during various large price movements featuring dramatic price decline/increase in a short time period in corn futures market, from 2014 to 2017. We specify a vector autoregressive model (VAR) to model the dynamics in the top of the book and use impulse response functions (IRFs) to examine the dynamic behavior of the spread. Our results reveal a resilient spread which is expected to narrow substantially within 5 – 20 seconds and completely revert back to normal state within 15 - 40 seconds once being shocked to widen. Along with the small average magnitude of bid-ask spread, our results suggest that corn futures market does not appear to experience significant liquidity deterioration over highly volatile periods, and that traders and hedgers who demand immediate execution can expect to do so at a reasonably and consistently low cost throughout the large price movement horizon.

Key words: corn, futures market, bid-ask spread, liquidity cost, price movements, resiliency

1 Introduction

Bid-ask spread is a transactional property of markets relevant to practitioners. It measures the magnitude of the price concession paid for immediate order execution [Working and Larson, 1967, Stoll, 2000]. It is also closely related to the concept of liquidity in that a liquid market is characterized by small bid-ask spread [Kyle, 1985]. In electronic markets currently employed by agricultural commodity trading, participants supply or demand liquidity by submitting limit orders or market orders, respectively. Studies on agricultural commodity markets reveal generally reduced execution costs represented by small bid-ask spreads in electronic markets [Shah and Brorsen, 2011, Wang et al., 2014]. Nevertheless, the positive relationship found between bid-ask spread and price volatility [Bryant and Haigh*, 2004, Frank and Garcia, 2011, Shah and Brorsen, 2011, Wang et al., 2014] raises potential concerns about the liquidity costs, in recent years when markets have witnessed increased volatility due to the growing presence of intra-day large price movements in major agricultural commodity markets. In 2014, Chicago Mercantile Exchange (CME) raised the price limit for grains futures in response to the increasing number of days when prices hit the limit. In 2015, Commodity Futures Trading Commission (CFTC) identified ’flash events’ in corn futures market characterized by a large round-trip of prices within one hour[1]. While the intra-day large price movements have received much attention from market participants as well as regulators, little evidence is provided regarding the bid-ask spread faced by hedgers and traders during these periods when their need for immediate

[1] Remarks of Chairman Timothy Massad before the Conference on the Evolving Structure of the U.S. Treasury Market, October 21th 2015, available at https://www.cftc.gov/PressRoom/SpeechesTestimony/opamassad-30.

execution may be most pronounced. Moreover, while most existing literature has focused on the magnitude of the bid-ask spread, its dynamic response to an external shock is argued to be a more essential consideration in optimal execution strategies [Obizhaeva and Wang, 2013], and therefore warrants the attention from market participants.

Prior to the transition to an electronic trading platform, the open out-cry futures markets recorded only transaction prices but not bid and ask quotes. Thus, the studies on open out-cry agricultural commodity markets have largely relied on inferring the bid-ask spread based on transaction data [Thompson et al., 1987, Roll, 1984, Hasbrouck, 2004]. The inferred bid-ask spreads, however, are found to be sensitive to the choice of estimation method and when further analyses are conducted using the estimated bid-ask spread, sometimes counterintuitive results are generated [Frank and Garcia, 2011]. Several recent studies involve the use of electronically-traded data for agricultural commodities [Shah and Brorsen, 2011, Martinez et al., 2011], but still work with transaction-price based estimator for bid-ask spread. The only study on the observed bid-ask spread in agricultural commodity markets is conducted on a daily basis and uses data from 2008 - 2010, a particularly turbulent period that has witnessed possibly unusual market behavior [Wang et al., 2014]. However, focusing on daily patterns of bid-ask spread may overlook important aspects of liquidity costs, since over 40 % (50 %) of orders (in volume) placed by liquidity providers are executed within 1 minute in corn (soybean) futures market from 2012 - 2014 [Haynes and Roberts, 2015]. While studies on agricultural commodity markets have revealed quick adjustments by the market in terms of price discovery and volatility using intra-day data [Lehecka et al., 2014, Adjemian and Irwin, 2016, Joseph and Garcia, 2018], little is known about the adjustment process of the quoted bid-ask spread in agricultural commodity markets and its interaction with other market dynamics.

Using the Market Depth Data from Chicago Mercantile Exchange (CME) for corn futures, we obtain not only actual bid-ask spread at a higher granularity, but also market depth which documents the number of contracts available at the best quoted price level, the number of transactions and the time between market events, through the reconstruction of the limit order book. This allows us to simultaneously model an extensive set of market dynamics, taking into consideration the potential co-movement of spread and market depth as documented in Frank and Garcia [2011] who study livestock futures markets and use volume per transaction as a proxy for the depth.

To our knowledge, this is the first study to explore the intra-day dynamic behavior of bid-ask spread in corn futures market. Bid-ask spread is an important component of transaction costs. Anyone who wishes to trade immediately has to sell at a lower price and buy at a higher price than could be otherwise achieved by posting a limit order. Therefore, the bid-ask spread measures the costs of immediacy faced by traders, and a persistently wider spread imposes larger costs and yields lower profits. Since it is most relevant for traders to understand the cost of immediacy when the immediacy is most needed, we focus on the intra-day periods characterized by large price movements and high volatility from Jan 4, 2014 to May 31, 2017. We then construct a Vector Autoregressive Model (VAR) to model the top of the book as well as transactions dynamics and generate impulse response functions (IRFs) to examine the dynamic behavior of the spread in face of different market orders acting as liquidity shocks. Despite large and rapid movements in price, the spread is found to be not only small on average, but also resilient in face of different liquidity shocks. The spread reverts to its steady state quickly when unexpectedly widened. While the adjustment speed of the spread varies according to when the large price movement is observed, our findings suggest that a higher-than-average liquidity cost decreases substantially in 5 - 20 seconds, and the impact completely dies out in no longer than 40 seconds. As a result, market participants who want to trade immediately should be able to do so at reduced costs, even if the market is experiencing large price movements. Our study provides the existing literature on liquidity costs in agricultural commodity markets with evidence of a resilient bid-ask spread over extreme periods and provides useful information to market participants.

2 Relevant Literature

Studies on liquidity costs in agricultural markets have focused mostly on the bid-ask spread, a price concession paid by traders who demand immediate execution of their orders. Numerous studies on the old open out-cry markets for agricultural commodities have recognized the presence of scalpers, or equivalently market makers, who earn a return for providing liquidity. Thompson et al. [1987] argue that scalping is efficient because of the competition among scalpers in providing liquidity to off-floor traders, which minimizes the transaction costs in futures markets. They also find that liquidity in a market is the primary determinant of the size of bid-ask spreads, leading to smaller costs of trading in more heavily traded markets. Brorsen [1989] estimated the liquidity costs in the corn futures market using data on transactions and revealed a liquidity cost nearly equal to the tick size. With the introduction of the electronic trading platform to agricultural commodity markets, studies involving electronic data on transactions and quotes have emerged. Bryant and Haigh* [2004] find larger bid-ask spreads in cocoa and coffee futures after the trading was automated and argue that the anonymity in electronic trading increases the information asymmetry and results in wider spread. In contrast to their findings, Shah and Brorsen [2011] reveal considerably lower liquidity costs in the electronic wheat market compared to the open-outcry market. Frank and Garcia [2011] also identify significant competitive effect of electronic trading on liquidity costs in livestock markets. Using the observed bid-ask spread, Wang et al. [2014] recognize that electronic trading leads to low and stable liquidity costs in corn futures market.

In addition to the overall assessment of the size of bid-ask spreads in agricultural commodity markets, it is also of broad interest to researchers and practitioners to understand when and how the spread varies. Stoll [1978] was among the first to discuss determinants of the bid-ask spread and proposed that it is positively correlated with the risk during the period when dealers are holding a position. Higher risk induces higher return demanded for liquidity provision, and therefore leads to wider bid-ask spread. This implies a positive relationship between the price volatility, which adds risk to the value of the inventory, and the size of the bid-ask spread. Studies conducted on agricultural commodity markets largely support this view. Shah and Brorsen [2011] document a positive relationship between volatility and bid-ask spread in wheat futures markets. Wang et al. [2014] use a structural model to reflect the relationship between bid-ask spread and volatility, and report similar results. Our analysis builds on this literature by elucidating how the bid-ask spread behaves during extreme market events.

3 Data

3.1 Limit Order Book Data

The data used in this study are the Market Depth Dataset (MDP) for corn futures from Chicago Mercantile Exchange (CME). We choose corn futures because it is one of the most actively traded agricultural commodities and thus provides a rich amount of observations. This dataset contains incremental messages that can be used to reconstruct the limit order book and transactions at millisecond resolution. The data spans from Jan 4th, 2014 to May 31st, 2017. Contracts of corn futures are traded with five maturities per year: March, May, July, September and December. For each day, we choose the most traded contract which is considered to contain leading information in price discovery [Hu et al., 2017] and also receives the most attention from market participants.

Corn futures are traded on the Chicago Board of Trade (CBOT) using a centralized electronic trading platform named CME GLOBEX. Corn futures contracts are traded in two sessions, 8:30 a.m. to 1:20 p.m. CT (morning session) and 7:00 p.m. to 7:45 a.m. CT (evening session). We use data from the morning session from Monday to Friday due to the low volume traded overnight.

To construct the limit order book, the incremental update messages are used to determine the status of the limit order book after each event (order submission, cancellation or transaction). The messages contain information on the time at a millisecond granularity, the type of the update, the contract name on which the update happens, the position in the order book on which the update occurs and the remaining quantity and prices after the update.

The MDP contains updates happening in both the outright market and the spread market. The CME Group supports implied functionality which is the ability to combine spread and outright markets into a consolidated book with the objective to increase liquidity. The outright limit order book contains ten depths on both bid and ask side for corn futures. The implied limit order book contains two depths on both bid and ask sides. The reconstruction methods for outright and implied books are the same. A more detailed procedure of merging outright and implied order book can be found in Arzandeh and Frank [2017]. For the purpose of accounting for overall available liquidity faced by market participants, the consolidated limit order book is used in our analysis. Figure 1 provides the daily settlement price for the mostly traded corn futures contracts over the data period. The dashed lines correspond to the day of rolling to the first deferred contract.

3.2 Identification of Intra-day Large Price Movements

Large price movements occurring during different intra-day time intervals are not equally likely to be anticipated by the market, and therefore might induce different market reactions. Brooks et al. [2003] identify that the adjustment speed of price is significantly slower after an unanticipated event of information arrival compared to a scheduled one. Graham et al. [2006] find that the information processing is affected by whether the timing of the information is known in advance. Distinguishing between anticipated and unanticipated events is thus warranted.

Therefore, to generate meaningful results, we identify the presence of large price movements separately during four different trading periods, namely, mid-day with or without USDA announcements as well as market opening and market closing. This is because large price movements happening in the middle of the day are less likely to be anticipated by the market relative to those observed during market opening and closing, with the exception of those caused by USDA announcements whose timing is fully anticipated by the public[2]. In the following subsections we present a detailed procedure of how the large price movements are identified during different intra-day time periods.

Mid-day Large Price Movement on Non-announcement Days

Inspired by the CFTC's characterization of 'hourly flash events' in corn futures market[3], we apply a rolling window of 60 minutes and a threshold of at least 200 basis points (bps) change to search for mid-day large price movements on non-announcement days. For a corn contract with a price of 400 cents per bushel, for instance, 200 bps movement corresponds to a change of 8 cents per bushel in price, or $400 in total value of the contract. With the minimum price fluctuation for corn futures being 0.25 cents per bushel (1 tick), the 200 basis points movement results in a price movement of 32 ticks.

To account for intra-day seasonality as discussed before, for each trading day, the first 15 and the last 15 minutes of the morning session are excluded and separately considered in

[2] With the absence of identifiable public information, the price volatility is found to be significantly higher during the market opening and closing period in the morning trading session [Lehecka et al., 2014], likely due to the clustering of private information that arrives overnight. This reflects intra-day seasonality and thus is largely anticipated by the participants.

[3] CFTC defines an hourly flash event to be the one where price moves at least 200 basis points within an hour but recovers to within 75 basis points of initial price within the same hour.

the following subsections. We use mid-quote at millisecond granularity as the price. The procedure is summarized as follows:

(a) Roll the time window starting from each mid-quote observation and identify if there is a price movement larger than 200 basis points in this time window.

(b) If a price movement larger than 200bps is found, record the time it reaches the minimum (maximum), and repeat (a) after this point.

(c) If no large price movement is found, repeat (a) starting from the next mid-quote observation.

This procedure identifies a total of 33 large price movements in the sample period, including 16 price decreases and 17 price increases. The days observed with mid-day large price movements constitute about 4 % of the total trading days in our data. The intervals from the starting of the price decrease (increase) to the minimum (maximum) point are then chosen as the samples for further analysis. The average magnitude of price movements for these identified events are 217 basis points.

Mid-day Large Price Movements after USDA Announcements

The USDA announcements such as WASDE, Crop Production, Grain Stocks, Perspective Planting and Acreage reports, are major reports affecting corn futures, and are well-known to cause large price movements [Isengildina-Massa et al., 2008, Lehecka et al., 2014]. During our data period, these reports are released at 11:00 a.m. CDT. Since the market adjusts quickly to the incoming information [Lehecka et al., 2014], the post-announcement period is chosen to be the 15 minutes after the USDA report release, or equivalently 11:00 a.m. - 11:15 a.m. We drop 2015/08/12 because the corn futures price hit limit following the release of the USDA reports, resulting in potentially abnormal market behaviors[4]. Throughout our sample period, there are in total 51 announcement days and 20 have witnessed price movements over 200 basis points within 15 minutes after the announcement. We thus choose these 20 announcement days for further analysis. The average magnitude of price movements in this case is 319 basis points and the days chosen represent 2.5% of the total trading days.

Large Price Movements during Opening and Closing

The opening period is chosen to be the first 15 minutes of the trading session on each trading day, or equivalently 8:30 a.m. - 8:45 a.m. The closing period is chosen to be the last 15 minutes before the morning trading session ends[5]. Since the 15-minute price

[4] On 2015/08/12 at 11:00 a.m., USDA released WASDE and Crop Production. Corn price plunged and hit price limit at around 11:08:52, with no selling happening for 3 seconds and best ask volume rising to over 1700, about 28 times higher than the previous 8 minutes.

[5] On July 6th, 2015, CME changed the closing time from 1:15 p.m. CDT to 1:20 p.m. CDT. Therefore, closing period is chosen to be 1:00 p.m. - 1:15 p.m. before the change and 1:05 p.m. - 1:20 p.m. after the change.

movement during market opening and closing is less likely to reach 200 basis points [6], we choose the 20 opening (closing) periods experiencing the maximum price movements. We calculate the magnitude of price movement during opening (closing) as the difference between the highest (lowest) price and the opening (closing) price. The average magnitude of the price movements in these selected days is 167 basis points during the opening period and 116 basis points during the closing period, representing the 2.5% of days in our sample.

3.3 Summary of Identified Large Price Movements

This subsection presents the summary of the identified large price movements with a focus on the bid-ask spread. Figure 2 shows the monthly distribution of the selected large price movements.

Several months are observed to experience large price movements more than others: April when the corn is planted in its primary growing regions across the United States, June and July when the growth of the corn is in the most critical stage and October when the harvest is in process. Specifically, the number of mid-day large price movements evidently spikes in July, two months prior to the harvest season. This is consistent with the previous findings where corn futures trading in July tend to witness higher volatility than other periods throughout the year [Anderson, 1985, Kenyon et al., 1987]. In terms of price movements during opening and closing, April, June, July and October are found to have greater numbers than other months. September is the only month that contains large price movements only induced by the release of USDA reports in our data.

Figure 3 records the second-by-second bid-ask spread averaging over the selected periods. The bid-ask spread is first aggregated to 1-second level using the mean within this second for each large price movement, and then averaged across those events. During the mid-day large price movements on non-announcement days, the mean bid-ask spread fluctuates around one tick, its minimum size, showing no significant sign of liquidity deterioration in terms of the tightness of the market. As shown in Table 1, the average bid-ask spread is $0.0026 per bushel during the mid-day large price movements, a magnitude close to one tick.

On USDA announcement days, the bid-ask spread immediately jumps to a significantly higher level following the announcement, reaching over $0.0055 per bushel. The spread then decreases sharply, stabilizing in less than 200 seconds. The average bid-ask spread in this case is $ 0.0029 per bushel as shown in Table 1, slightly higher than other periods examined. Furthermore, Table 1 shows that the best bid and ask volume as well as the rest of the book depth is significantly lower than other periods. The combination of higher

[6] There are only 2 opening periods and 1 closing period observed with over 200 bps movement from Jan 4th, 2014 to May 31st, 2017.

spread and lower depth indicates relatively worse liquidity conditions than in non-announcement days.

The bid-ask spread during opening and closing does not exhibit distinctive patterns, except that the spread is found to be relatively higher when market just opens. The average spread is \$ 0.0027 per bushel during opening and \$ 0.0026 per bushel during closing.

To conclude, the magnitude of average bid-ask spread observed during different large price movements turns out to be relatively small, close to its minimum value during most of the time, indicating an overall low cost of immediate execution faced by market participants during these periods.

In the following section, we model the limit order book and transactions dynamics through a vector autoregressive (VAR) specification and examine the behavior of bid-ask spread during large price movements, providing information on the time profile of the spread in response to liquidity shocks.

4 Method

4.1 Model Specification

In this section, we aim at exploring the dynamic behavior of bid-ask spread in the corn futures market, a purely order-driven market. This can be achieved by considering a vector autoregressive (VAR) specification that models the top of the book prices and quantities as a system, similar to Hasbrouck [1995] and Engle and Patton [2004]. In corn futures market, liquidity suppliers submit limit ask or bid orders specifying a price and a quantity. The dynamics of the spread are thus primarily determined by the differential behavior between best bid and ask prices. Incoming trades convey information on the underlying value that moves the best quoted prices [Hasbrouck, 1991], but existing theories suggest conflicting implications in terms of their impact on bid-ask spread. Glosten and Milgrom [1985] find that market orders cause the spread to fall, as the uncertainty regarding the information is partially resolved by the (direction of the) trade. In contrast, Easley and O'hara [1992] argue that trades lead to higher uncertainty for market makers regarding whether an information event has just occurred and thus results in wider spread.

The spread is also argued to interact with the depth quoted by liquidity suppliers. Kyle [1985] predicts that times of greater information asymmetry lead to both wider spread and lower depth. Empirical evidence has confirmed a negative relationship between the spread and depth both in hybrid and order-driven markets [Lee et al. 1993, Ahn and Cheung 1999], indicating that limit order traders actively manage both price and quantity dimensions of market liquidity. Our model is constructed in event time so as to avoid the loss of information in aggregating data and the ambiguity in choosing the sampling frequency. The depth also partially captures the on-going realized volatility since it is shown to vary with it [Engle and Lange, 2001]. At last, following Lo and Hall [2015], we also include the duration between each event in the model to facilitate the conversion from event time to calendar time.

The aforementioned consideration leads to a set of endogenous variables as follows:

$$x_t = \{p_t^a, p_t^b, v_t^a, v_t^b, d_t, buy_t, sell_t\}',$$

where t corresponds to an event time index indicating changes in the above variables resulting from either an incoming market order, a limit order submission or cancellation. We express the best bid and ask prices, p_t^a and p_t^b in logarithms, since taking logarithms not only helps reduce the influence of extreme values and mitigate heteroscedasticity, but also resolves the problem of bid and ask only moving in discontinuous fashion by tick size. v_t^a and v_t^b represent the (displayed) number of contracts in logarithm associated with the best ask and bid [7] respectively and is considered the best quoted depth at event t. d_t indicates the clock time elapsed between each event measured in seconds and serves the purpose of converting the event time to the clock time, allowing us to present the results in a more intuitive way. Finally, buy_t and $sell_t$ are indicators for buyer-initiated trade and seller-initiated trade, respectively, which take the value of 1 when a buyer (seller) -initiated trade occurs, and 0 otherwise. A detailed description of the variables can be found in Table 2.

We first check the stability of our system by conducting Augmented Dickey-Fuller (ADF) tests on the endogenous variables in vector x_t for each large price movement period. As expected, the best bid and ask prices are found to be integrated of order one in 90 and 91 out of 93 cases, respectively, as shown in Table 3. Other endogenous variables as well as the log spread calculated as the difference between log quote prices ($s_t = p_t^a - p_t^b$), reject the null of unit root in all sample periods. This leads to use a vector error correction model (VECM) in which only best bid and ask prices are cointegrated with spread acting as the error-correction term. The other stationary variables just correspond to other cointegrating vectors constituted by themselves. As such, the model can be formally written as:

Equation 1

$$\Delta x_t = \mu + \alpha\beta' x_{t-1} + \sum_{i=1}^{p-1} \Psi_i \Delta x_{t-i} + \epsilon_t$$

with the cointegrating matrix β being specified as:

[7] The data do not include the iceberg orders which are not displayed to other market participants. Moreover, since there is only a marginal proportion (less than 5%) of transaction being executed beyond the posted best ask and best bid during our sample period. (See Table 4)

$$\beta = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

where the constant term μ captures any potential non-zero mean remaining in both the equilibrium relationships and the differenced series (See Hendry and Juselius [2001], Section 6.1).

The VECM we considered above embeds a reduced form VAR for the log spread, log mid-quote return as well as the level of depths, duration and trade indicators, analogous to Engle and Patton [2004] and Lo and Hall [2015]. This representation is more appealing than the original VECM given that the log spread, which is of most interest to us, as well as the remaining variables, are stationary. It also allows the interactions between the spread and other market dynamics to be directly observed. To obtain the reduced form VAR, we apply a rotation matrix defined as below:

$$R = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0.5 & 0.5 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

Multiplying both sides of Equation 1 by R yields:

Equation 2

$$R\Delta x_t = R\alpha\beta' x_{t-1} + \sum_{i=1}^{p-1} R\Psi_i R^{-1} R\Delta x_{t-i} + R\epsilon_t$$

where $Rx_t = \{\Delta s_t, \Delta q_t, \Delta v_t^a, \Delta v_t^b, \Delta d_t, \Delta buy_t, \Delta sell_t\}$. Here, $\Delta s_t = \Delta p_t^a - \Delta p_t^b$ and $\Delta q_t = 0.5(\Delta p_t^a + \Delta p_t^b)$ are changes in log spread and changes in the midpoint of log bid and ask prices (log mid-quote return), respectively. By construction, the changes in bid and ask prices are fully reflected in these two variables (Δs_t and Δq_t), and the dynamics of log spread is explicitly modeled. Further manipulation of equation 2 leads to the implied VAR model:

Equation 3

$$y_t = c + \sum_{i=1}^{p} A_i y_{t-i} + u_t$$

where $y_t = \{s_t, \Delta q_t, v_t^a, v_t^b, d_t, buy_t, sell_t\}$ and the transformation implies that the second column of A_p is zero. Derived from the original VECM, this re-specified VAR is a stable system that maintains the efficient parameterization and can be estimated equation by equation using OLS with the second column of A_p removed [Lo and Hall, 2015].

Instead of concatenating data, estimation of equation 3 is conducted for each large price movement observed during different intra-day periods to avoid potential structural breaks. To determine the lag number p, we use Ljung and Box [1978] to test for serial correlation in the residual series. The results indicate that inclusion of 14 lags is sufficient to remove the serial correlation in the residuals for post-USDA announcement large price movements when the quoting activities are most intensive. Similar procedure leads us to include 12 lags during mid-day large price movements, and 8 lags for opening and closing price movements. We anticipate the heteroskedasticity in the residuals due to the presence of large price movements, and therefore use White-consistent estimator of variance-covariance matrix for inference.

4.2 Dynamic Behavior of Bid-ask Spread

The dynamic behavior of the spread can be obtained from IRFs that measure the adjustment path of the expectation of y_t in response to the occurrence of shocks in the multivariate system:

Equation 4

$$f(h;\sigma_y) = E_t(y_{t+h}|y_t + \sigma_y, y_{t-1}, \cdots) - E_t(y_{t+h}|y_t, y_{t-1}, \cdots)$$

where h indicates the number of future steps and σ_y indicates the changes in y_t induced by the shocks. We compute the spread IRFs to different shocks defined below.

4.2.1 *Unexpected market orders as liquidity shocks*

In this section, we consider the liquidity shocks in the form of different types of unexpected market orders that result in the deviation from the system's steady state. Emerging from the error term of the system, the unexpected market order induces changes in y_t that cannot be explained by historical information, and thus could alter the subsequent market expectation of the spread through the dynamic interactions estimated in equation 3. Let σ_y be the shock vector representing the change in y_t induced by the unexpected market order. Computation of the IRF in equation 4 requires defining the size of the shock (σ_y) and the state of the system before the shock occurs. All variables are initialized to their long-run equilibrium. Following Lo and Hall [2015], the long-run equilibrium is determined by the unconditional mean over all sample periods and the trade indicator is set to zero to represent a tranquil period.

We design the unexpected market order to be similar to what we see in the market to accommodate representative situations. Due to the linearity of VAR model, the absolute magnitude of the shock vector is of less importance as it does not alter the shape of IRFs. Table 4 demonstrates the summary of different types of market orders during the sample period by type of large price movement. For each type, trades executed within the best bid and ask are predominant. Such trades, on average, consume approximately 10% of the depth at best quoted prices. A small portion of transactions result in direct changes of prevailing best bid or ask price by executing against all available contracts, while a marginal portion walks down (up) the limit order book and is partially filled by the limit orders posted on second level bid or ask.

Analogous to these three situations, we consider three representative shocks summarized as follows:

1. Normal Market Orders (Scenario NMO): This scenario arises when an unexpected market order is executed within the best bid or ask level and consumes 10% of the depth at best bid or ask.
2. Aggressive Market Order (Scenario AMO): This scenario arises when an unexpected market order is executed against all the depth available of the prevailing best bid or ask. This scenario is characterized by instantaneous changes in mid-quote and bid-ask spread as one of the best quotes moves one tick away from the top of the book. The exhausted level is replenished with the available depth at the previous second-best quote.
3. Aggressive Market Order Walking through the L1 Depth (Scenario AMOW): This scenario arises when the unexpected market order is executed against all limit orders of best bid or ask as well as 50 % of those of the second-best quotes, i.e., 'walking through the first level of depth'. This results in not only the immediate change in mid-quote and spread, but also the depth at the previous second-best quote that automatically becomes the current best one.

Table 5 summarizes the shock vectors of each scenario introduced above. For each scenario, shocks caused by market buy order as well as market sell order are included.

4.2.2 *Impulse Response Functions*

This section presents the calculation of impulse response functions. We do not use orthogonalized impulse response function here since we are not interested in the casual relationships between variables, but rather the overall movements of the whole set of market dynamics in the system. Moreover, the contemporaneous relationships among the variables are captured by the construction of the shock vectors [Hautsch and Huang, 2012]. Since the system in equation 3 is stationary, then by Wold representation theorem, it can be equivalently expressed in a VMA(∞) form as follows:

$$y_t = v + \Phi_0 u_t + \Phi_1 u_{t-1} + \Phi_2 u_{t-2} + \cdots$$

where $\Phi_0 = I_K$ and for $i > 0$, $\Phi_i$can be calculated recursively using the equation below [Hamilton, 1994]:

Equation 5

$$\Phi_i = A_1\Phi_{i-1} + A_2\Phi_{i-2} + \cdots A_p\Phi_{i-p}$$

where $\{A_j\}_{j=1}^{p}$ are the coefficients of the VAR model in equation 3.

It is then straightforward to see that matrix Φ_h corresponds to

$$\frac{\partial y_{t+h}}{\partial u_t'} = \Phi_h,$$

thus $\Phi_h(i,j)$, the element in row i column j in matrix Φ_h, represents the consequences of one unit increase in j-th innovation at time t on the i-th endogenous variable at time $t\ + h$. Then the impulse-response function in equation 4 can be calculated using

$$\hat{f}\big(h;\sigma_y\big) = \widehat{\Phi_h}\sigma_y$$

where $\widehat{\Phi_h}$ is calculated by replacing the$\{A_j\}_{j=1}^{p}$ in equation 5 by its OLS estimator $\{\widehat{A_j}\}_{j=1}^{p}$. In particular, to obtain the impulse response function of the spread, we only need to focus on the first element of $\hat{f}\big(h;\sigma_y\big)$ for any h.

Taking into consideration the potential heteroscedasticity, we generate the confidence interval for the impulse response functions using the wild bootstrapping method proposed by Wu [1986]. The bootstrapping procedure is repeated for 1,000 times for each estimation.

4.2.3 *Converting Event time to Calendar Time*

Due to the tick data used in estimation, the impulse response function illustrates the dynamic adjustment path of the bid-ask spread to different shocks in event time. However, it is more informative to present the evolution of bid-ask spread in calendar time for a given shock. Because the high-frequency data used in our sample are irregularly spaced, additional steps are needed to convert the event time to calendar time.

Since inter-event duration is endogenous in our model, it is possible to calculate the conditional expected duration at each update t. As suggested by Manganelli [2005], one can feed into the system the unconditional mean of original series of duration and compute the expected timing of future events by progressively cumulating these durations. Denote the unconditional mean of duration as $\bar{d}$, then the timing of s-th event time is obtained by:

$$E\big(d_{t+s}\big|y_t + \sigma_y, y_{t-1}, \cdots\big) = \bar{d} + \sum_{i=1}^{s}[\Phi_i]_k.\sigma_y$$

where $[\Phi_i]_{k\cdot}$ represents the k-th row corresponding to the position of the variable duration in the model. Obviously, the expectation of the timing is conditional on the initial shock.

5 Empirical Results

5.1 Estimation results of VAR model

This section presents the estimation results of the system of equations in 3. The estimated VAR system contains 7 equations that capture the extensive interactions among spread, mid-quote return, best quoted depth, duration and their lagged values. The estimation procedure is conducted for each large price movement period to avoid potential structural breaks generated by concatenating the data. This results in 33 separate estimations for mid-day large price movements, 20 for post-announcement large price movements and 20 for large price movements in opening and closing respectively. For the sake of brevity, we report parameter estimates as follows. For each variable, we define a coefficient group which includes all coefficients related to the variable's lag structure. We conduct a joint Wald-test on each coefficient group using the White-consistent estimator of the standard errors to determine the statistical significance of this coefficient group. Finally, we group the results by the four different types of large price movements (mid-day, post-announcement, opening and closing), reporting the mean sum of each coefficient group, as well as the proportion of the coefficient groups being positively or negatively significant for each event-type. This allows us to summarize our results in a relatively concise fashion, and also maintain the possibilities for comparison across different types of large price movements. The summary of results can be found in Table 7.

Overall, the estimated coefficients are largely consistent with our expectations and the findings of existing literature. The average adjusted R^2 is high for log spread and log depths equations, above 0.9 in general, while lower for log mid-quote return and duration equations. The joint Wald-test identifies multiple significant coefficient groups in each equation and also the existence of several two-way Granger causality relationships that further justify the VAR specification of the system. Moreover, the magnitude of coefficients and the significance varies substantially across different types of large price movements, indicating that the market indeed behaves differently during large price movements observed in different intra-day time intervals.

Specifically, the estimated coefficients in the equation of log spread (s_t) show that the spread exhibits strong persistence: the coefficients associated with the lagged spread are positive and significant during all large price movements. The results also reveal a significantly negative correlation between the spread and the past value of the depth at best bid or ask. At the same time, equations of the best quoted depth (v_t^a and v_t^b) show that the depth is negatively affected by the lags of spread, indicating a two-way Granger causality. This is consistent with the findings of Lee et al. [1993] for NYSE and Ahn and Cheung

[1999] for SEHK that higher spread is accompanied by lower depth. In corn futures market, liquidity providers also adjust both the spread and depth during large price movements to manage their risks. The relationship, however, is weaker during closing periods relative to other large price movements. This is possibly due to both the perceived information risk by liquidity providers during a large price movement that could warrant a wider spread, and the typical requirement to close positions before market closing that might cause them to post a greater depth or reduce the proportion of iceberg orders [Esser and Mönch, 2007]. Interestingly, the coefficients associated with the trade indicators are positive but only significant in a limited proportion of large price movements, with the post-USDA announcement price movements having the greatest magnitude as well as the highest ratio of significance. This implies that liquidity providers infer higher degree of information asymmetry from trades arriving after the announcement and tend to widen the spread subsequently, but still to a very limited extent. Shang et al. [2016] also find that although the adverse selection component of bid-ask spread in corn futures market increases after USDA announcement, the total variation of the spread is minimal.

Results of equation mid-quote return (Δq_t) show positive autocorrelation of mid-quote return except post-announcement large price movements. As expected, this short-run predictability of return is due to the sample featuring the directional large price movements and is not necessarily a sign of market inefficiency. The coefficients of trade indicators x_t^b and x_t^s lead to significant quote revision activities consistent with standard microstructure theories. The uninformed liquidity providers infer from the direction of the trade the current underlying value of the commodity and submit limit orders accordingly, which leads to a higher mid-quote return after a buyer-initiated trade and lower mid-quote return after a seller-initiated one. This effect is most pronounced during post-announcement large price movements in terms of the magnitude of the coefficients associated with x_t^b and x_t^s which are about three times larger than mid-day and opening price movements, and nearly ten times larger than closing. This corresponds to higher price impact of trades during times of higher information asymmetry. Moreover, significant two-way associations between the return and the depth at best quoted prices are observed. Higher depth at best bid helps predict a higher subsequent return, while higher depth at best ask helps predict a lower return. This is consistent with Cao et al. [2009] that the imbalance between limit buy orders and limit sell orders contributes to price discovery. Furthermore, the past mid-quote returns are shown to relate to the current depth at best quoted prices (shown in equation v_t^a and v_t^b) where a higher return would attract more limit sell orders and a lower return would attract more limit buy orders. This direction of Granger causality, however, is much less significant during post-announcement large price movements.

By including duration in the system, we are also able to examine the relationships between the frequency of the limit order book updates and the observed market conditions. In equation duration (d_t), the significant negative relationship between duration and the spread is most pronounced, indicating that book events tend to cluster after a wider spread is observed. This effect is found to be greatest during mid-day large price movements. The duration is also negatively affected by the arrival of trade and positively affected by the depth. Therefore, the frequency of book updating increases either after an incoming market order or a depletion of the depth.

5.2 Impulse responses of the spread

In this section, we present the impulse response functions that capture the dynamics of spread during large price movements. The shocks are constructed to represent six types of unexpected market orders that cause a deviation of the spread from its steady state and are examined individually for each large price movement. The 95% confidence intervals are generated by Wild bootstrapping for each shock. As our model is constructed in event time, the impulse response is originally generated in event time, too. By recursively accumulating the expected duration as described in section 4.2.3, we calculate the time in seconds corresponding to each event time for each impulse response function.

Due to space constraints, we generate average impulse response functions in calendar time for each type of large price movement: mid-day, post-announcement, opening and closing. This is done by taking average of the response of spread as well as the calendar time corresponding to each step. Therefore, the resulted impulse response function could be viewed as the average adjustment of the spread after the shock and the average time taken for the adjustment. Figures 4 to 7 display the average impulse response of spread for the four types of large price movements.

Mid-day large price movements

Figure 4 presents the average impulse responses of spread across 33 mid-day large price movements. The points on the horizontal line indicate the average number of events occurred before the impact of the shock is no longer significantly different from zero. Mid-day large price movements are those where the price moves over 200 basis points within an hour during the middle of the trading session and are less likely to be anticipated by the market. As noted, occurrence of such large price movements in a time interval as short as one hour turns out to be rare for corn futures market (they are observed in only 4 % of the total trading days in our sample), and thus might be conjectured to be accompanied by heightened pressure for market liquidity as what multiple other markets have experienced[8]. The impulse response functions, however, indicate a resilient spread that reverts back to its normal level in a quick fashion.

Specifically, in the scenarios where the unexpected buy or sell NMO order arrives (two top panels in figure 4), the spread has a tendency to widen subsequently, consistently with the information-based microstructure theory where the unexpected trade raises the information

[8] 'Flash event' in U.S. treasury futures market on October 15 2014 led to a significant strained liquidity according to the joint staff report ('The U.S. Treasury Market on October 15, 2014'). 'Flash crash' on May 6, 2010 involving multiple U.S. financial markets resulted in liquidity crisis and trading halt.

asymmetry perceived by liquidity providers and leads to wider spreads. Such an increased liquidity cost is found to have a limited magnitude and little persistence. The change on the vertical axis indicates that after the shock, the extra cost of immediacy is only approximately 0.001 % of the current price, corresponding to 0.004 cents per bushel if the transaction happens at 400 cents per bushel. This impact fades substantially over the first 5 seconds after the NMO, and the effect completely dies out in less than 10 seconds.

The scenarios characterized by different aggressive market orders share a similar adjustment path and time profile, as shown in the lower panels in Figure 5. At time 0, the spread is unexpectedly widened by 1 tick as a result of the liquidity shock and is shown to drop sharply immediately. In 10 to 20 seconds, the spread has virtually returned to its steady state with hardly any visible difference remaining, implying that traders can expect to demand immediacy at a barely higher-than-average cost even after large and aggressive transactions being observed. On average, the adjustment process of the spread takes less than 35 seconds before incoming new liquidity completely offsets the initial deviation. This is consistent with Biais et al. [1995] who found evidence that investors quickly place limit orders within the best quotes if the spread is large to gain price and time priority. The behavior of liquidity providers in corn futures market does not seem to steer away from this principle even in the presence of mid-day large price movements. As a result, it is expected that frequent traders or hedgers, whenever demanding immediate execution, are very likely to incur a reasonably and consistently low cost throughout the price movement period.

Post-announcement price movements

USDA announcements usually induce large responses from the agricultural commodity markets. For corn futures, USDA announcements contain important information on yield and stocks, for example, and are found to be highly valued by market participants. In average, the price movement following the release of USDA reports turns out to be the largest among all trading periods, and its timing is well-known to the public. Figure 5 presents the adjustment path of spread estimated from 20 announcement days with the magnitude of price movement being at least 200 basis points. In sharp contrast to the particularly volatile post-announcement period, the spread behaves resiliently.

In the upper panel, the unexpected NMO induces a positive impact on the spread, about 2 times the magnitude in mid-day large price movements. This implies higher information asymmetry perceived by the liquidity providers, as the same market order would be considered more likely to contain superior information. This is consistent with Kim and Verrecchia [1994] who point that the disclosure of public information allows some traders to make better forecast of the underlying value of the commodity than other traders, and that liquidity providers are unable to distinguish between these two groups of traders. The impact of a normal market order lasts significantly longer in event time than mid-day large price movements, an indication of deterioration of liquidity following the announcement.

However, it dissipates quickly in calendar time due to the active quoting activities clustering after the report release, and completely dies out in as short as 8-10 seconds.

The aggressive market orders induce similar adjustment paths for the spread, as shown in the lower panels in the figure. Immediately after being widened, the spread decreases dramatically in only 5 seconds, indicating a quick attempt of liquidity providers to fill in the gap in only 5 seconds. This higher liquidity cost faced by traders, again, is short-lived as the spread completely returns to its normal level in less than 15 seconds. The results are consistent with Brooks [1994] who found that although the adverse selection component increases after the announcement, its duration is often short, and the bid-ask spread tends to revert back to normal level quickly. With such a resilient spread, hardly any difference is there in terms of when to trade, even if the USDA announcement surprises the market and induces a large price response.

Opening and Closing price movements

Unlike large price movements in the middle of the trading session, the price movements in the opening and closing are commonly observed and contribute to the intra-day seasonality of the return variability. The behavior of the spread during market opening and closing is of particular interest to practitioners as it helps them to make more informed decisions in day-to-day trading. Figures 6 and 7 show the average impulse response functions of the spread during opening and closing large price movements, respectively.

Shown by the upper panels of these two figures, the impact of an unexpected NMO on the spread exhibits different magnitude, depending on whether it occurs at open or at close, an indication of different degree of information asymmetry perceived by the market. An identical liquidity shock induces a wider spread that lasts longer in both event time and calendar time during opening price movements than closing price movements, as liquidity providers tend to infer a higher probability of being adversely selected around market opening than around market closing. However, the impact of this shock is limited as for its magnitude and duration. The extra liquidity cost corresponds to additional 0.0015 % of the price during opening price disappears in less than 4 seconds, while the magnitude and duration of the extra liquidity cost is even smaller and shorter for closing price movements.

In response to aggressive market orders that unexpectedly widen the spread by 1 tick, the adjustment path shows sharp and then steady decrease in the spread in both opening and closing price movements. The widened spread reverts to its steady state within 15 seconds or equivalently, 147 events in general, for different types of aggressive market orders during opening price movements, while the same process takes nearly 30 seconds (about 225 events) to complete during closing price movements. This indicates a higher incentive for liquidity providers to act in response to a spread that is wider than expected when the market opens than when the market approaches the end of the session, probably because liquidity providers are less willing to enter positions and increase inventory towards the end of the day session. Nevertheless, the higher liquidity cost induced by the shock witnesses a substantial drop in 5 - 10 seconds during both periods, recovering by over 99 % towards its steady state. Since a persistently higher cost of immediacy is not likely to occur, traders would find limited difference in execution costs if trading at different time throughout opening and closing price movements.

6 Conclusion

Bid-ask spread is an important component of transaction costs in corn futures markets. It measures the price concession paid for immediate execution and is viewed as the common gauge of liquidity costs faced by traders. As corn futures market has witnessed a growing presence of intra-day large price movements and increased volatility in recent years, much attention has been drawn to market participants and regulators regarding the liquidity costs. In this paper, we focus on large price movements observed during the middle of the day session, after the USDA announcements, as well as in the market opening and closing periods from Jan 2014 to May 2017. These intra-day large price movements may feature intensive inflow of information, anticipated or not, and varying perception of the underlying price, during which the demand of immediate execution is most pronounced, but the supply is potentially under stress.

Our findings first reveal an overall small bid-ask spread throughout different types of intra-day large price movements. The spread is on average 0.01 or 0.02 cents larger than its minimum level of 0.25 cents/bushel, with a tendency to be slightly higher during post-USDA announcement large price movements, reaching an average of 0.29 cents/bushel. This finding supplements Wang et al. [2014] who find a generally small bid-ask spread almost impervious to short-term changes in demand for spreading and trend-following trade activities in corn futures market from 2008 to 2010.

To understand how the cost of immediate execution changes over time, we then study the dynamic behavior of the spread during the large price movements. Using a vector autoregressive specification, we model the top of the book as a system where liquidity provision is driven by public limit orders that determine the prevailing best bid and ask prices, and therefore the spread. Our estimation results reveal significant dynamic interactions among an extensive set of market variables including the log spread, log mid-quote return as well as the log depths at best ask and bid, and are largely consistent with the predictions of existing microstructure theories. Using impulse response functions, we examine the dynamic behavior of the spread in face of representative market orders acting as liquidity shocks. We find that despite the large and rapid movement in price, the spread is resilient as it reverts quickly to the steady state once being widened unexpectedly. Our results suggest that a higher-than-average execution cost is expected to be short-lived, typically dropping substantially in 5 - 20 seconds and lasting no longer than 40 seconds, with the specific adjustment speed differing upon when the large price movement is observed. In contrast to other markets that have witnessed significantly deteriorated liquidity characterized by persistently increasing bid-ask spread accompanying large and unexpected price movements, the corn futures market maintains not only a small bid-ask spread on average during the large price movements, but also a resilient one that responds fast to deviations from the expectation, leading to a consistently low liquidity cost no matter when execution.

To conclude, we find little evidence of corn futures market suffering from significantly strained liquidity condition or growing difficulties to execute market orders even during periods of unexpected and rapidly developing price movements. Our analysis provides insights of how the liquidity cost behaves when the market is experiencing large price movements along with high volatility and can potentially help traders and hedgers in corn futures market to make more informed decisions.

Figures and Tables

Table 1: Mean Statistics of Market Variables during Large Price Movements

Events	#Sample Periods	Price Move(%)	#LOB updates	#Sell Trades	#Buy Trades	Ask Price	Bid Price	L1 Ask	L1 Bid	Duration (seconds)
Mid-day	33	2.17	17636	1594	1485	383.35	383.09	147	141	0.2
Opening	20	1.67	9938	978	953	386.70	386.43	113	104	0.1
Closing	20	1.16	7970	1017	921	389.11	388.85	274	305	0.12
Post Announcement	20	3.19	14597	2175	2165	388.99	388.70	61	65	0.07

Note: This table summarizes the main market variables across different type of large price movements. #Sample periods indicate the number of large price movements within each category. Price Move (%) is the average magnitude of the price movement. #LOB updates and #Trades stand for the average number of limit order book updates and transactions, respectively. Ask Price and Bid Price represent the average best ask and best bid price measured in cents per bushel. L1 Ask and L1 Bid represent the average depth at best ask and bid. Duration measures the average time elapsed between two consecutive updates in the top of the book, expressed in seconds.

Table 2: Description of the endogenous variables

Variable	Description
p_t^a	The logarithm of ask price after t-th update
p_t^b	The logarithm of bid price after t-th update
$v_t^{a,1}$	The logarithm of best ask volume after t-th update
$v_t^{b,1}$	The logarithm of best bid volume after t-th update
d_t	Duration between t-th and $(t-1)$-th updates measured in seconds
x_t^b	Indicator for buyer-initiated trade that takes the value of one if a buyer-initiated trade occurs at t-th update and zero otherwise
x_t^s	Indicator for seller-initiated trade that takes the value of one if a seller-initiated trade occurs at t-th update and zero otherwise
s_t	The logarithm of bid-ask spread calculated by $s_t = p_t^a - p_t^b$
Δq_t	Mid-quote return of log quoted prices calculated by $\Delta q_t = 0.5(\Delta p_t^a + \Delta p_t^b)$

Table 3: Augmented Dickey-Fuller test of Variables

Events	Rejection	p_t^a	p_t^b	$v_t^{a,1}$	$v_t^{b,1}$	d_t	x_t^b	x_t^s	s_t	Δq_t
Mid-day	Number	0	0	33	33	33	33	33	33	33
	Ratio (%)	0	0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Opening	Number	2	2	20	20	20	20	20	20	20
	Ratio (%)	10.0	10.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Closing	Number	0	0	20	20	20	20	20	20	20
	Ratio (%)	0	0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Post	Number	1	0	20	20	20	20	20	20	20
Announcement	Ratio (%)	5.0	0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Note: The ADF test is conducted for all variables listed above across all selected sample periods. The results show the number of rejection at 1 % significance level as well as the ratio of rejection across all samples. The number of lags used in the test is 50.

Table 4: Summary of Market Order Activities

Events	Trade within L1 depth			Trade equal to L1 depth			Trade exceeding L1 depth		
	Frequency (%)	AvgSZ	AvgL1	Frequency (%)	AvgSZ	AvgL1	Frequency (%)	AvgSZ	AvgL1
Mid-day	90.31	11	102	7.11	14	14	2.58	38	25
Opening	88.49	10	76	8.33	13	13	3.18	34	22
Closing	94.60	13	257	3.91	13	13	1.49	36	23
Post Announcement	84.09	10	54	10.97	9	9	4.93	31	14

Note: Trade within L1 summarizes the transactions executed within the best bid or ask volume; Trade equal to L1 summarizes the transactions of the same size to the prevailing best bid or ask; Trade exceeding L1 summarizes the transactions that walks up or down the book. Ratio indicates the percentage of the corresponding type of transactions, AvgSZ indicates the average volume of the corresponding type of transactions and AvgL1 indicates the average size of best bid or ask when the transaction occurs.

Table 5: Shock Vectors Used in Impulse Response Function

Panel A: Shock vectors of Market Sell Orders

Scenario	s_t	Δq_t	$v_t^{a,1}$	$v_t^{b,1}$	d_t	x_t^b	x_t^s
MO	0	0	0	-0.105	0	0	1
AMO	$s_t^a - \bar{s_t}$	$\Delta q_t^a - \bar{\Delta q_t}$	0	0	0	0	1
AMOW	$s_t^a - \bar{s_t}$	$\Delta q_t^a - \bar{\Delta q_t}$	0	-0.693	0	0	1

Panel B: Shock vectors of Market Buy Orders

Scenario	s_t	Δq_t	$v_t^{a,1}$	$v_t^{b,1}$	d_t	x_t^b	x_t^s
MO	0	0	-0.105	0	0	1	0
AMO	$s_t^b - \bar{s_t}$	$\Delta q_t^b - \bar{\Delta q_t}$	0	0	0	1	0
AMOW	$s_t^b - \bar{s_t}$	$\Delta q_t^b - \bar{\Delta q_t}$	-0.693	0	0	1	0

Note: $\bar{s_t} = log\bar{p_t}^a - log\bar{p_t}^b$ represents the equilibrium value of spread that is calculated for each type of large price movements where $\bar{p_t}^a$ and $\bar{p_t}^b$ corresponds to the unconditional mean of bid and ask prices as in Table 1. Accordingly, $s_t^a = log\bar{p_t}^a - log(\bar{p_t}^b - 0.25)$ and $s_t^b = log(\bar{p_t}^a + 0.25) - log\bar{p_t}^b$. By simple calculation, we also have $\Delta q_t^a - \bar{\Delta q_t} = -\frac{1}{2}(s_t^a - \bar{s_t})$ and $\Delta q_t^b - \bar{\Delta q_t} = \frac{1}{2}(s_t^b - \bar{s_t})$.

Table 7: VAR estimation results of large price movements

	Mid-day			Post announcement			Opening			Closing		
Variable	Mean sum of Coef	Pos Sig (%)	Neg Sig (%)	Mean sum of Coef	Pos Sig (%)	Neg Sig (%)	Mean sum of Coef	Pos Sig (%)	Neg Sig (%)	Mean sum of Coef	Pos Sig (%)	Neg Sig (%)
(a) Equation: Spread												
s_t	0.787	100	0	0.806	100	0	0.769	100	0	0.823	100	0
Δq_t	0.007	0	0	0.044	5	0	-0.032	5	5	-0.036	0	5
$v_t^{a,1}$	-0.00001	6	88	-0.00001	10	90	-0.00001	0	55	-0.00001	10	45
$v_t^{b,1}$	-0.00001	3	85	-0.00002	0	75	-0.00001	0	80	-0.00001	5	50
d_t	-0.00000	6	12	-0.00002	0	5	-0.00001	0	10	-0.00001	0	20
x_t^b	0.00001	6	0	0.00003	20	0	0.00001	20	0	0.00001	0	0
x_t^s	0.00001	12	0	0.00004	25	0	0.00002	10	0	0.00000	0	0
$adj.R^2$	0.979			0.929			0.971			0.982		
(b) Equation: Mid-quote return												
s_t	-0.002	0	12	0.006	5	0	0.012	10	5	-0.008	0	10
Δq_t	0.473	97	0	-0.044	5	5	0.260	85	0	0.499	90	0
$v_t^{a,1}$	-0.00001	0	100	-0.00002	5	95	-0.00001	0	100	-0.00001	5	90
$v_t^{b,1}$	0.00001	97	3	0.00001	80	20	0.00001	100	0	0.00001	85	5
d_t	-0.00000	0	0	0.00002	0	0	0.00000	0	10	0.00000	0	0
x_t^b	0.00003	91	0	0.0001	95	0	0.00004	90	0	0.00002	30	0
x_t^s	-0.00003	0	85	-0.0001	0	95	-0.00004	0	75	-0.00001	0	15
$adj.R^2$	0.110			0.090			0.102			0.110		
(c) Equation: Depth at best ask												
s_t	-495.510	0	88	-250.220	0	100	-465.994	0	90	-550.439	0	65
Δq_t	2,377.238	97	0	224.894	15	5	1,397.255	90	0	3,132.909	70	0
$v_t^{a,1}$	0.881	100	0	0.859	100	0	0.853	100	0	0.881	100	0
$v_t^{b,1}$	0.010	21	15	0.020	45	20	0.022	15	15	0.005	10	5
d_t	0.003	9	0	0.125	10	0	0.007	5	0	-0.066	0	10
x_t^b	0.049	9	0	0.003	10	0	0.027	5	5	0.018	0	0
x_t^s	-0.042	0	21	-0.090	0	35	-0.069	0	5	0.006	0	0
$adj.R^2$	0.979			0.947			0.970			0.982		
(d) Equation: Depth at best bid												
s_t	-455.793	0	94	-267.153	0	95	-520.150	0	85	-452.028	0	40
Δq_t	-2,285.280	0	91	-104.434	0	15	-1,302.235	0	85	-3,433.512	0	80
$v_t^{a,1}$	0.027	39	3	0.052	85	0	0.036	20	0	0.047	25	5
$v_t^{b,1}$	0.881	100	0	0.860	100	0	0.844	100	0	0.875	100	0
d_t	0.002	3	0	0.022	0	0	-0.022	5	0	-0.066	0	10
x_t^b	-0.068	0	24	-0.102	0	40	-0.080	0	15	-0.010	0	0
x_t^s	0.058	15	0	0.029	10	0	0.070	5	0	-0.028	0	0
$adj.R^2$	0.978			0.947			0.972			0.984		
(e) Equation: Duration												
s_t	-144.444	0	91	-15.098	5	90	-79.909	0	95	-40.258	5	50
Δq_t	-0.572	15	12	8.038	0	5	-14.502	0	0	-2.848	5	10
$v_t^{a,1}$	0.021	79	6	0.004	70	5	0.009	85	10	-0.006	20	10
$v_t^{b,1}$	0.043	97	0	0.022	60	5	0.038	95	0	0.044	40	0
d_t	0.544	100	0	0.510	100	0	0.454	100	0	0.574	100	0
x_t^b	-0.180	0	94	-0.065	0	100	-0.101	0	100	-0.101	0	85
x_t^s	-0.158	3	88	-0.059	0	100	-0.086	0	90	-0.075	0	80
$adj.R^2$	0.273			0.272			0.267			0.285		
N	17636			14597			9938			7970		

Note: The table shows estimation results of equation 3 specified in section 4.1. Mean sum of Coef summarizes the the average value of the sum of coefficients associated with all lags of a certain variable across all mid-day, post announcement, opening and closing large price movements. Pos Sig and Neg Sig indicate the ratio of days where the coefficients are positive and jointly significant or negative and jointly significant. The joint significance is determined using Wald test, and White-consistent estimator is used for the testing. The adjusted R^2 and N (observations) are average values across the equations estimated.

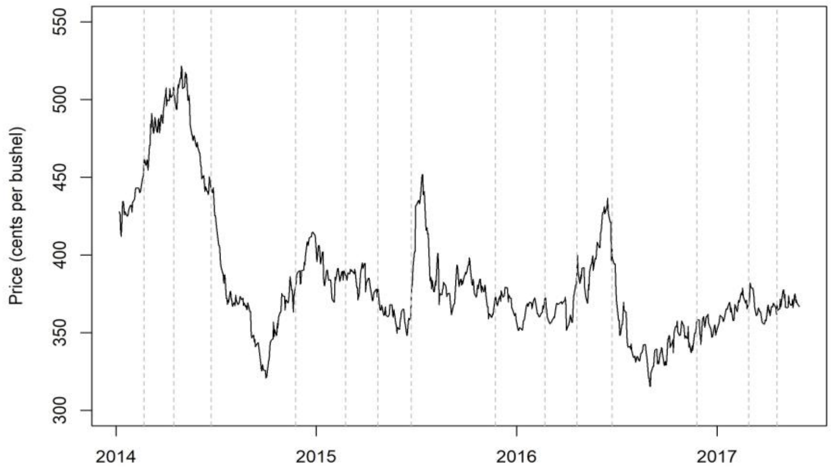

Figure 1: Settlement price of mostly traded corn futures contracts

Note: The dashed lines in Figure 1 indicates the time for rolling from the nearby contract to the deferred one.

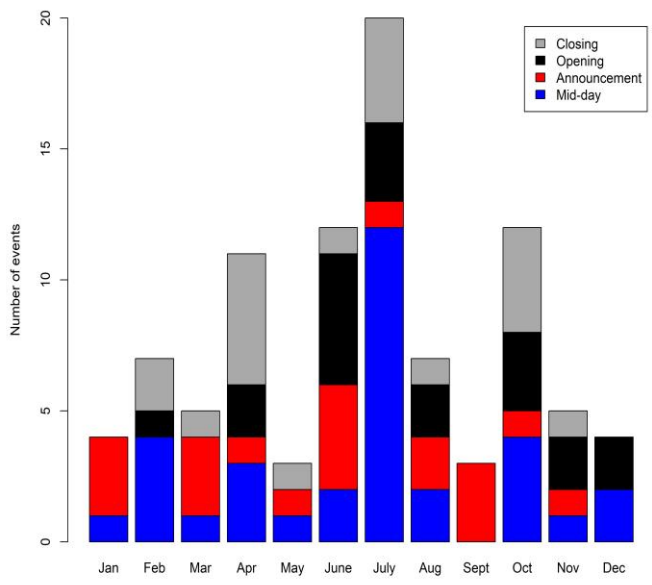

Figure 2: Monthly Distribution of Large Price Movements

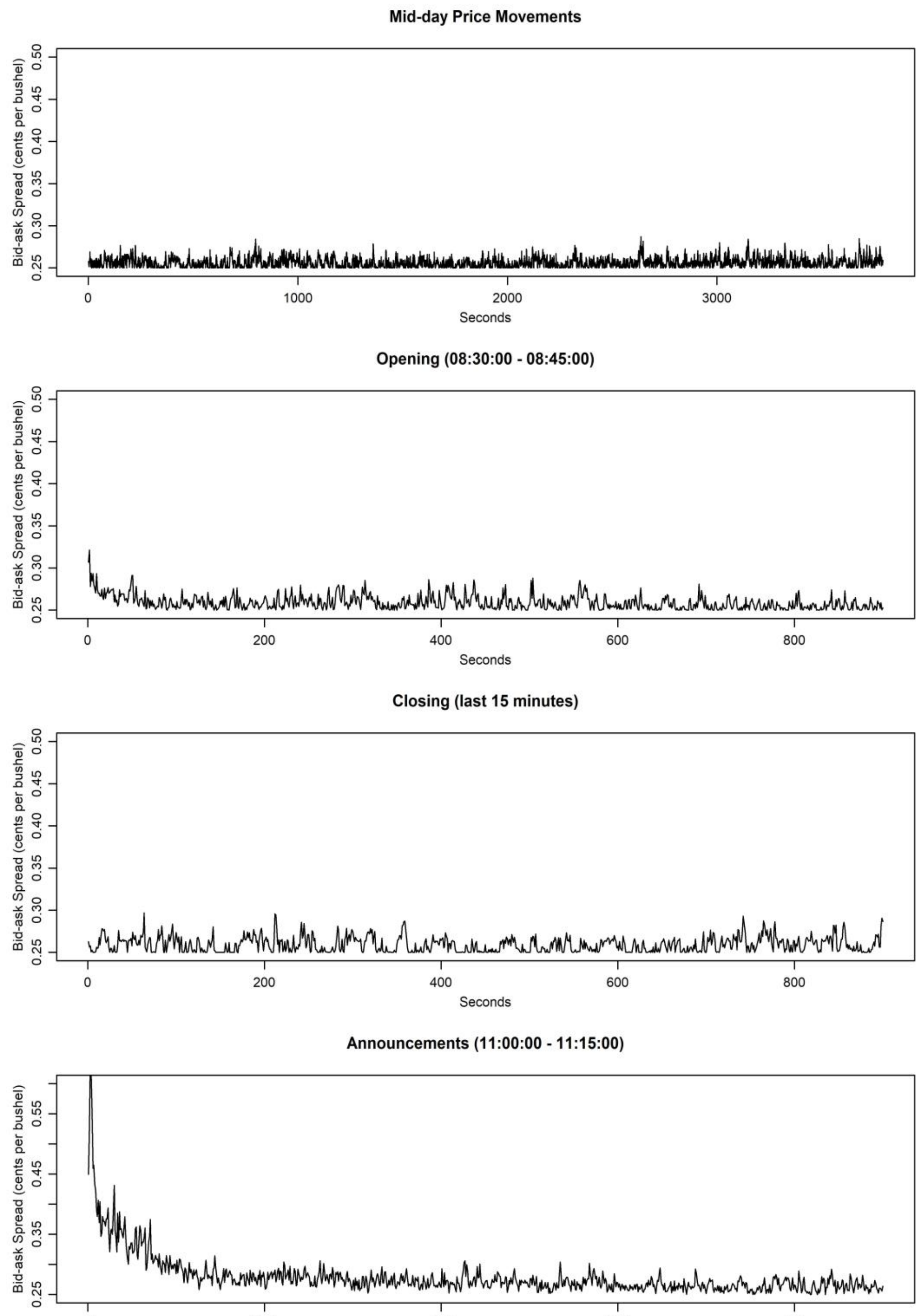

Figure 3: Second-by-second Bid-ask Spread during different sample periods

Note: The bid-ask spread is first aggregated to 1-second level using the mean within this second for each large price movement, and then averaged across those events, namely, 33 mid-day large price movements, 20 post-announcement, opening and closing price movements respectively.

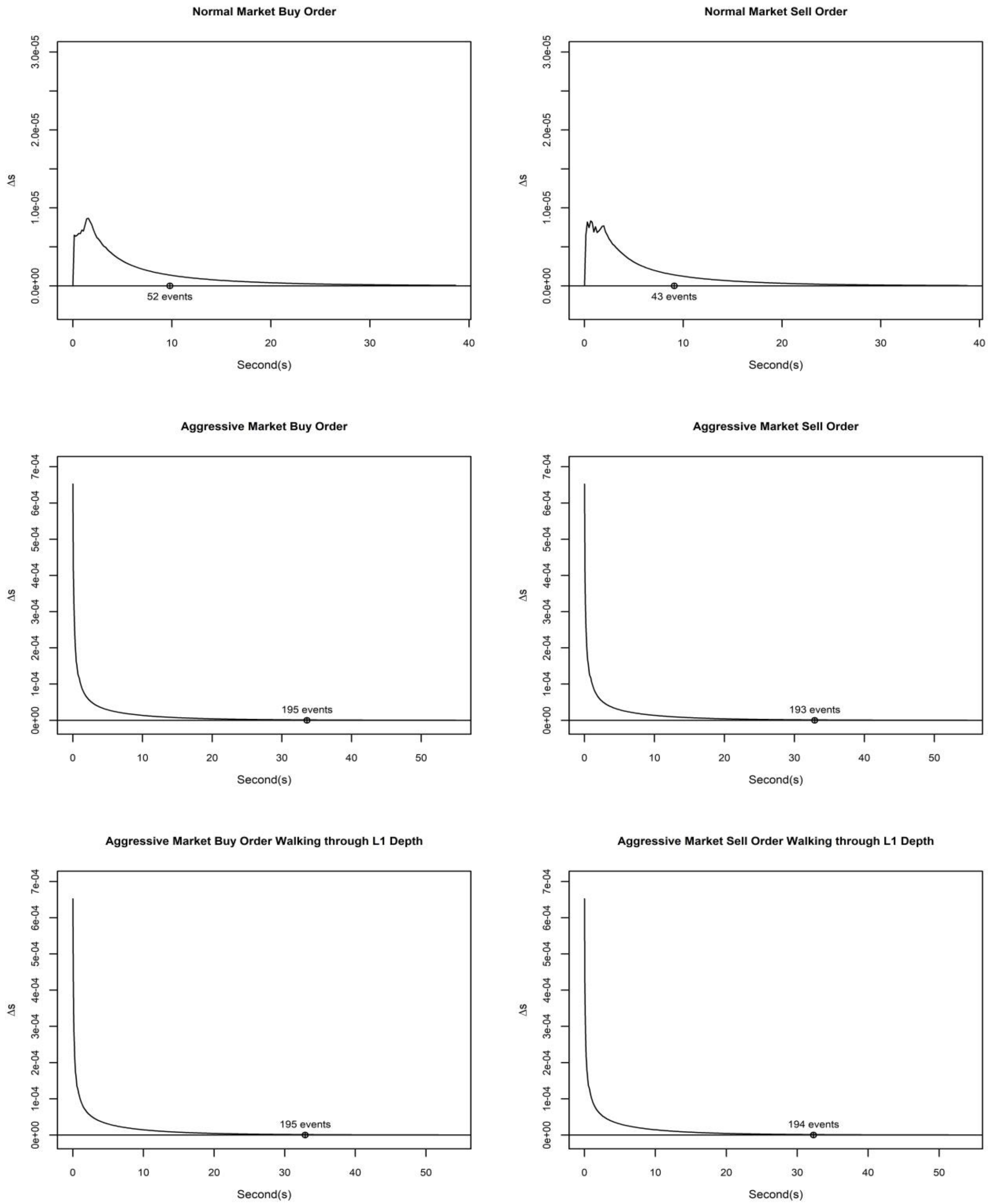

Figure 4: Average impulse responses of spread: 33 mid-day large price movements

Note: The figure above presents average impulse responses of spread across 33 mid-day large price movements under different shocks. The points on the horizontal line indicate the average number of events occurred before the impact of the shock is no longer significantly different from zero.

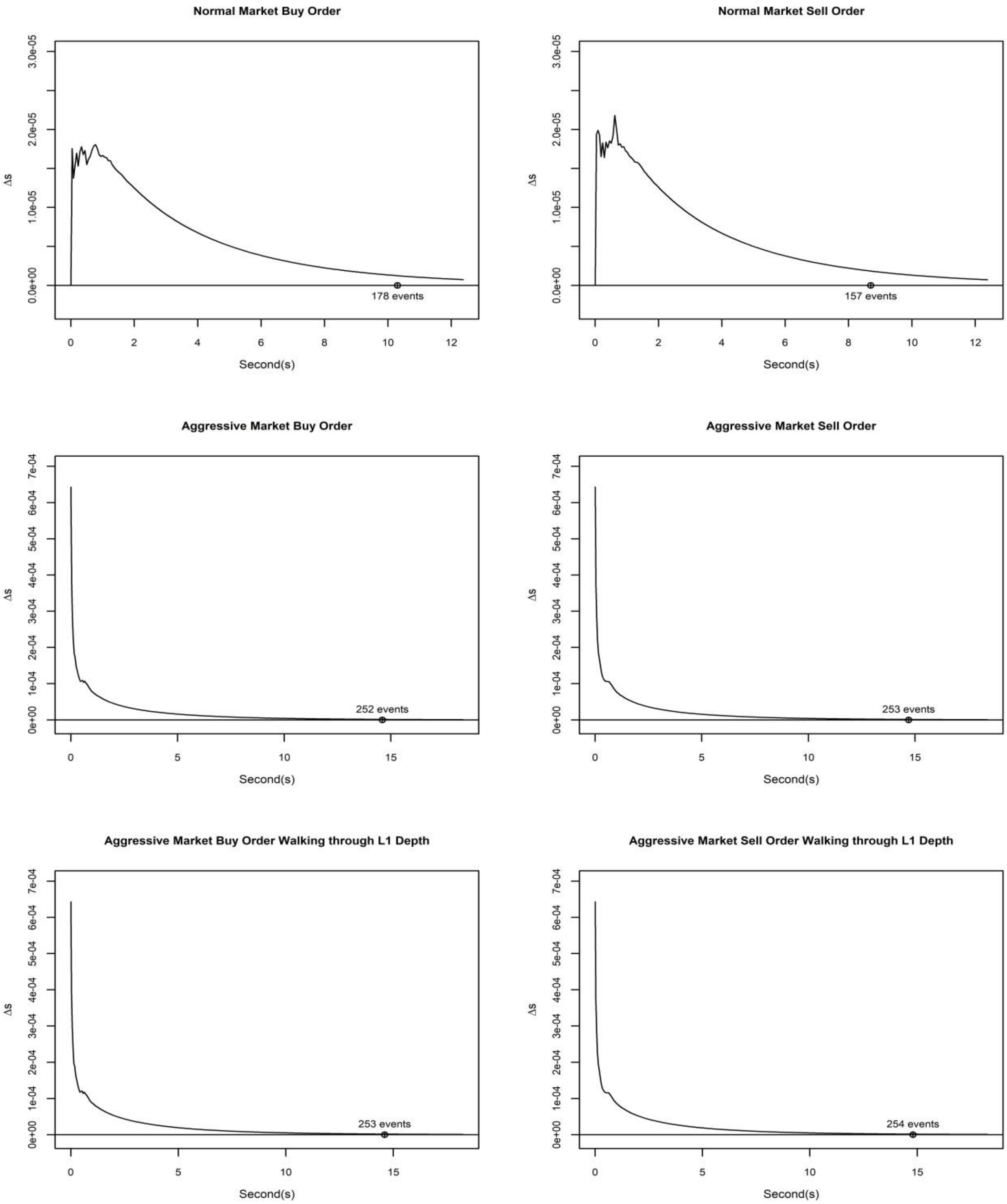

Figure 5: Average impulse responses of spread: 20 post-announcement large price movements

Note: The figure above presents average impulse responses of spread across 20 post-announcement large price movements under different shocks. The points on the horizontal line indicate the average number of events occurred before the impact of the shock is no longer significantly different from zero.

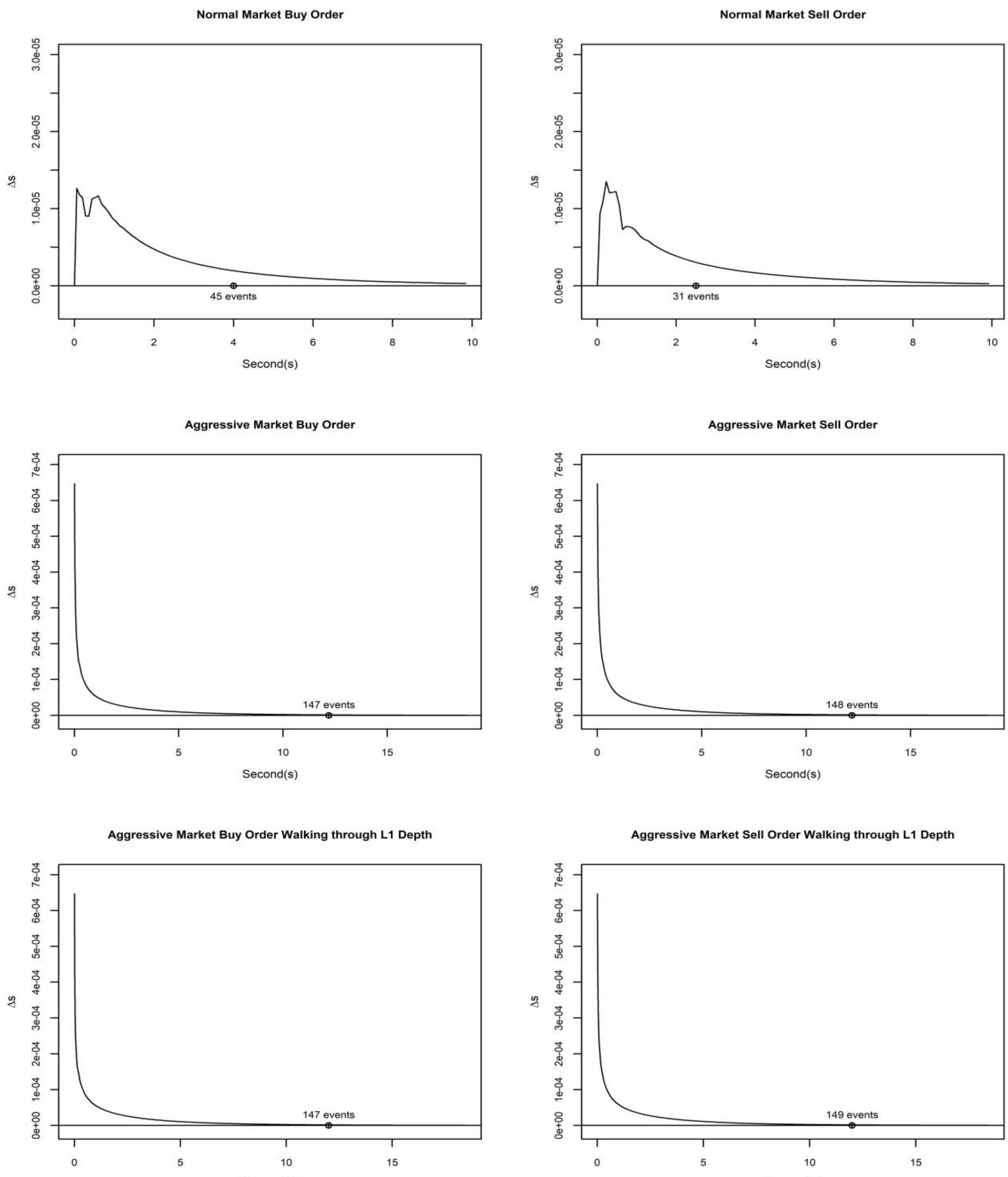

Figure 6: Average impulse responses of spread: 20 opening large price movements

Note: The figure above presents average impulse responses of spread across 20 opening large price movements under different shocks. The points on the horizontal line indicate the average number of events occurred before the impact of the shock is no longer significantly different from zero.

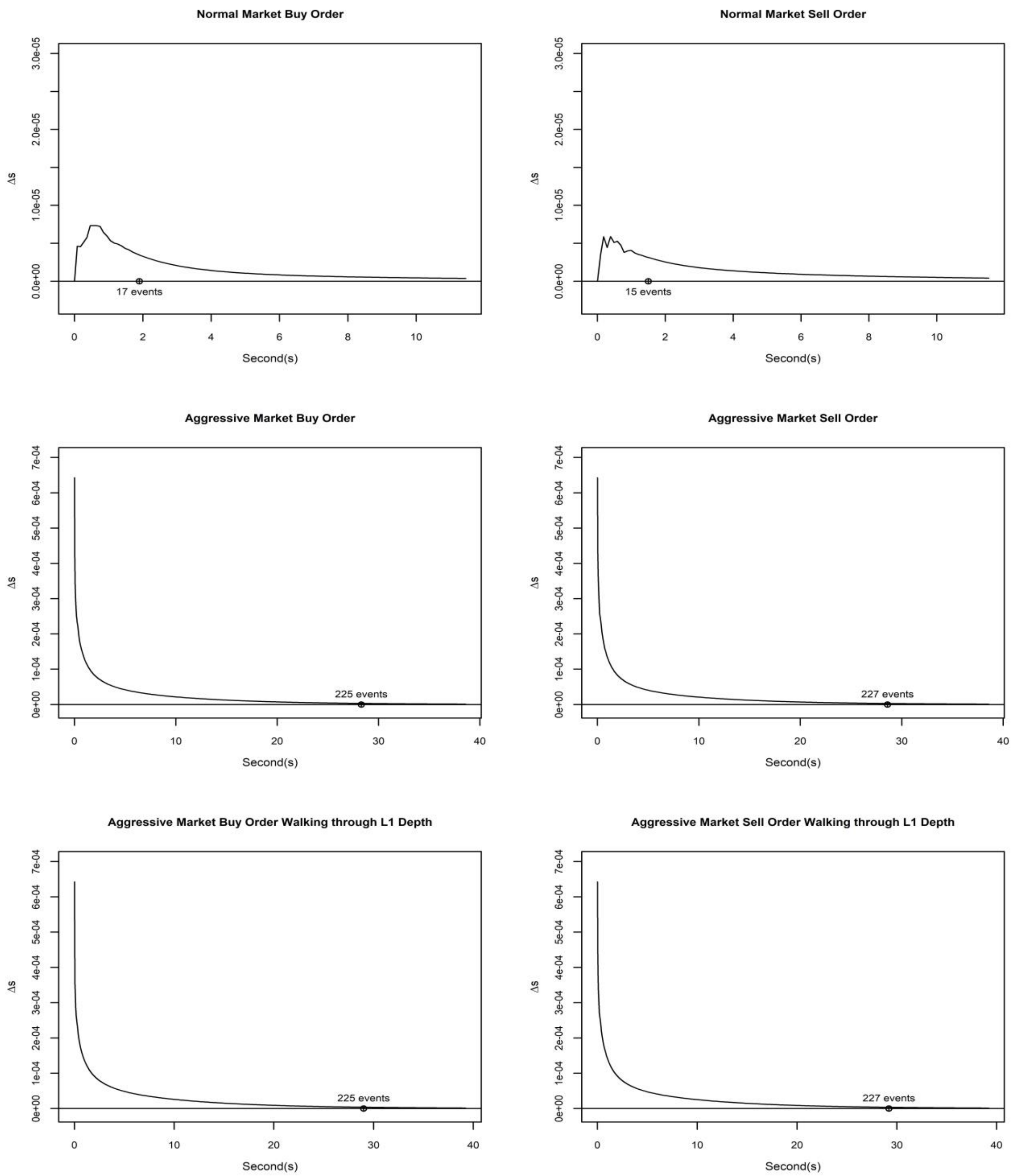

Figure 7: Average impulse responses of spread: 20 closing large price movements

Note: The figure above presents average impulse responses of spread across 20 closing large price movements under different shocks. The points on the horizontal line indicate the average number of events occurred before the impact of the shock is no longer significantly different from zero.

References

Michael K Adjemian and Scott H Irwin. Usda announcement effects in real time. *Economic Research Service*, 2016.

Hee-Joon Ahn and Yan-Leung Cheung. The intraday patterns of the spread and depth in a market without market makers: The stock exchange of hong kong. *Pacific-Basin Finance Journal*, 7(5):539–556, 1999.

Ronald W Anderson. Some determinants of the volatility of futures prices. *Journal of Futures Markets*, 5 (3):331–348, 1985.

Mehdi Arzandeh and Julieta Frank. The information content of the limit order book. 2017.

Bruno Biais, Pierre Hillion, and Chester Spatt. An empirical analysis of the limit order book and the order flow in the paris bourse. *the Journal of Finance*, 50(5):1655–1689, 1995.

Raymond M Brooks. Bid-ask spread components around anticipated announcements. *Journal of Financial Research*, 17(3):375–386, 1994.

Raymond M Brooks, Ajay Patel, and Tie Su. How the equity market responds to unanticipated events. *The Journal of Business*, 76(1):109–133, 2003.

B Wade Brorsen. Liquidity costs and scalping returns in the corn futures market. *Journal of Futures Markets*, 9(3):225–236, 1989.

Henry L Bryant and Michael S Haigh*. Bid–ask spreads in commodity futures markets. *Applied Financial Economics*, 14(13):923–936, 2004.

Charles Cao, Oliver Hansch, and Xiaoxin Wang. The information content of an open limit-order book. *Journal of futures markets*, 29(1):16–41, 2009.

David Easley and Maureen O'hara. Time and the process of security price adjustment. *The Journal of finance*, 47(2):577–605, 1992.

Robert F Engle and Andrew J Patton. Impacts of trades in an error-correction model of quote prices. *Journal of Financial Markets*, 7(1):1–25, 2004.

Angelika Esser and Burkart Mönch. The navigation of an iceberg: The optimal use of hidden orders. *Finance Research Letters*, 4(2):68–81, 2007.

Julieta Frank and Philip Garcia. Bid-ask spreads, volume, and volatility: Evidence from livestock markets. *American Journal of Agricultural Economics*, 93(1):209–225, 2011.

Lawrence R Glosten and Paul R Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics*, 14(1):71–100, 1985.

John R Graham, Jennifer L Koski, and Uri Loewenstein. Information flow and liquidity around anticipated and unanticipated dividend announcements. *The Journal of Business*, 79(5):2301–2336, 2006.

James Douglas Hamilton. *Time series analysis*, volume 2. Princeton university press Princeton, 1994.

Joel Hasbrouck. Measuring the information content of stock trades. *The Journal of Finance*, 46(1):179–207, 1991.

Joel Hasbrouck. One security, many markets: Determining the contributions to price discovery. *The journal of Finance,* 50.4 (1995): 1175-1199.

Joel Hasbrouck. Liquidity in the futures pits: Inferring market dynamics from incomplete data. *Journal of Financial and Quantitative Analysis*, 39(2):305–326, 2004.

Nikolaus Hautsch and Ruihong Huang. The market impact of a limit order. *Journal of Economic Dynamics and Control*, 36(4):501–522, 2012.

Richard Haynes and John S Roberts. Automated trading in futures markets. *Commodity Futures Trading Commission, March*, 2015.

David F Hendry and Katarina Juselius. Explaining cointegration analysis: Part ii. *The Energy Journal*, pages 75–120, 2001.

Olga Isengildina-Massa, Scott H Irwin, Darrel L Good, and Jennifer K Gomez. The impact of situation and outlook information in corn and soybean futures markets: Evidence from wasde reports. *Journal of Agricultural and Applied Economics*, 40(1):89–103, 2008.

Kishore Joseph and Philip Garcia. Intraday market effects in electronic soybean futures market during non-trading and trading hour announcements. *Applied Economics*, 50(11):1188–1202, 2018.

David Kenyon, Kenneth Kling, Jim Jordan, William Seale, and Nancy McCabe. Factors affecting agricultural futures price variance. *Journal of Futures Markets*, 7(1):73–91, 1987.

Oliver Kim and Robert E Verrecchia. Market liquidity and volume around earnings announcements. *Journal of accounting and economics*, 17(1-2):41–67, 1994.

Albert S Kyle. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335, 1985.

Charles MC Lee, Belinda Mucklow, and Mark J Ready. Spreads, depths, and the impact of earnings information: An intraday analysis. *The Review of Financial Studies*, 6(2):345–374, 1993.

Georg V Lehecka, Xiaoyang Wang, and Philip Garcia. Gone in ten minutes: Intraday evidence of announcement effects in the electronic corn futures market. *Applied Economic Perspectives and Policy*, 36(3): 504–526, 2014.

Greta M Ljung and George EP Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2): 297–303, 1978.

Danny K Lo and Anthony D Hall. Resiliency of the limit order book. *Journal of Economic Dynamics and Control*, 61:222–244, 2015.

Simone Manganelli. Duration, volume and volatility impact of trades. *Journal of Financial Markets*, 8(4): 377–399, 2005.

Valeria Martinez, Paramita Gupta, Yiuman Tse, and Jullavut Kittiakarasakun. Electronic versus open outcry trading in agricultural commodities futures markets. *Review of Financial Economics*, 20(1):28–36, 2011.

Anna A Obizhaeva and Jiang Wang. Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16(1):1–32, 2013.

Richard Roll. A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of finance*, 39(4):1127–1139, 1984.

Samarth Shah and B Wade Brorsen. Electronic vs. open outcry: side-by-side trading of kcbt wheat futures. *Journal of Agricultural and Resource Economics*, pages 48–62, 2011.

Quanbiao Shang, Mindy Mallory, and Philip Garcia. The components of the bid-ask spread: evidence from the corn futures market. *Agricultural Economics*, 2016.

Hans R Stoll. The supply of dealer services in securities markets. *The Journal of Finance*, 33(4):1133–1151, 1978.

Hans R Stoll. Presidential address: friction. *The Journal of Finance*, 55(4):1479–1514, 2000.

Sarahelen R Thompson, Mark L Waller, et al. The execution cost of trading in commodity futures markets. *Food Research Institute Studies*, 20(2):141–163, 1987.

Xiaoyang Wang, Philip Garcia, and Scott H Irwin. The behavior of bid-ask spreads in the electronically-traded corn futures market. *American Journal of Agricultural Economics*, 96(2):557–577, 2014.

Holbrook Working and Arnold B Larson. *Tests of a theory concerning floor trading on commodity exchanges*. Stanford University Palo Alto, CA, 1967.

Chien-Fu Jeff Wu. Jackknife, bootstrap and other resampling methods in regression analysis. *the Annals of Statistics*, pages 1261–1295, 1986.

Hu, Z., M. Mallory, T. Serra, and P. Garcia. 2017. “Measuring Price Discovery between Nearby and Deferred Contracts in Storable and Non-Storable Commodity Futures Markets.” Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.illinois.edu/nccc134].

Engle, Robert F., and Joe Lange. "Predicting VNET: A model of the dynamics of market depth." *Journal of financial markets*, 4.2 (2001): 113-142.