THE COMPONENTS OF THE BID-ASK SPREAD: EVIDENCE FROM THE CORN FUTURES MARKET

BY

QUANBIAO SHANG

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Agricultural and Applied Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Adviser:

Assistant Professor Mindy Malloy

ABSTRACT

Using the Best Bid Offer data from the CME, this thesis decomposes the Bid-Ask Spread (BAS) in the Chicago Board of Trade (CBOT) corn futures market into its three components, which are adverse selection, inventory, and order processing costs. Approximately 34.8% of the BAS is attributable to the order processing cost, and the order processing cost is relatively stable across different months, days, or trading hours. Liquidity providers' inventory cost is the highest cost component at 53.1%. And the adverse selection cost is 12.1%, which is the smallest BAS component. However, the adverse selection component can be higher when corn prices are more volatile in 2008 and 2011 than other less volatile years of 2009 and 2010. In general, the monthly pattern of the adverse selection cost seems to be different from year to year; the intraday pattern of the adverse selection cost appears to be U-shaped. In contrast, the inventory cost pattern is a strong inverted U-shape. The intraday order processing cost is relatively stable throughout each trading day. The market conditions are relatively different between USDA announcement and no-announcement days, especially during market opening and closing hours. In the first trading hour on USDA announcement days, the adverse selection cost is higher but the inventory cost is lower than on no-announcement trading days. Overall, this thesis shows that the BAS in the CBOT electronically traded corn futures market is relatively low and stable, but the magnitude of each BAS component varies.

To My Family and All Faculty Members from the UIUC ACE Department

ACKNOWLEDGEMENT

First and foremost, I would like to thank Professor Mindy Mallory for guiding me into the field of futures market microstructure and consistently giving me valuable advice and generous support. Her guidance and encouragement helps me to build a strong foundation for my future research career. Secondly, many sincere thanks to Professor Philip Garcia, who spends enormous amount of time in helping me to improve my research. I admire his expertise in related fields and also desire to learn much more from him in the coming years. In addition, I wish to present my gratitude to my committee members, Professor Scott Irwin and Professor Maria Teresa Serra Devesa, they have been giving me valuable suggestions and insights regarding my Master's Thesis. Last but not the least, this Master's Thesis is made possible by the Office of Futures and Options Research (OFOR) offering me the BBO dataset, I want to dedicate my acknowledgement of gratitude towards the OFOR.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The Bid-ask Spread (BAS) is the most common measure of the futures market liquidity. The futures market literature contains numerous studies of the BAS in agricultural commodity futures markets (e.g, Bryant and Haigh 2004 and Wang, Garcia, and Irwin 2013). While estimating commodity futures market BAS is important, understanding the components of BAS is more essential in market microstructure research. Because liquidity costs that comes from asymmetric information rather than inventory and order processing costs may require different market surveillance, and policy-making to ensure market efficiency. The objective of this thesis is to decompose the BAS in the Chicago Board of Trade (CBOT) corn futures market into three components: adverse selection, inventory, and order processing cost components. Minimizing the liquidity cost is beneficial to liquidity traders, who are often uninformed traders. From regulators' point of view, it is important that market prices reflect true asset values quickly and accurately. In addition, it is critical to quantify the adverse selection cost component because severe market asymmetric information may induce uninformed traders to leave the market. Under such circumstances, the market function is undermined since no price can clear the market (O'Hara 1995).

I analyze the pattern of BAS components on daily, monthly, and intraday time horizons, to shed light on what gives rise and fall to different components of liquidity costs in the CBOT corn futures market. In addition, the thesis examines how major United States Department of Agriculture (USDA) report announcement days differ from no-announcement days. I examine whether the asymmetric information component of liquidity cost is atypically high on USDA announcement days. This provides insights on market efficiency and information flows on announcement days. There is abundant literature on the decomposition of the BAS in equity markets (e.g, Huang and Stoll 1997, Madhavan, Richardson, and Roomans 1997). The research on the BAS of futures markets has been relatively scarce. This is the first study to decompose the BAS in any CBOT agricultural commodity futures market.

1.1. Background Information

Grain merchandisers and large end users make substantial use of commodity futures contracts as temporary substitutes for merchandising contracts that they will make later (Working 1970). Major participants in the futures market are long hedgers, who temporarily take long positions in the futures market to hedge price fluctuations for their later cash transactions; short hedgers, who are usually long in cash positions and temporarily short in futures markets; and speculators. Each of the participant may be ‘providers’ or ‘demanders’ of liquidity at different times. A limit order provides liquidity since it is a posted order to buy/sell at a specific price. A market order demands liquidity since it is an order to sell/buy against the best posted buy/sell limit order.

Working (1967) was the first one to study agricultural futures market microstructure. Floor traders derive income from hedgers through temporarily absorbing hedging orders that are otherwise not absorbed immediately. Buy orders are often executed at the liquidity provider’s ask price while sell orders are usually executed at the liquidity provider’s bid price. Transaction prices bounced between the bid and ask prices, which causes the transaction prices to be negatively autocorrelated. The negative price autocorrelation is a result of scalpers attempting to earn profits from filling market orders. In the stock markets, Demsetz (1968) suggests that liquidity providers are willing to stand ready to buy/sell at posted prices. Liquidity providers tend to hold relatively neutral net positions in order to manage their risks from price volatility.[1] To cover their cost of standing ready, they are willing to sell at a price that is slightly higher than the previous equilibrium price and buy at a price that is slightly lower than the previous equilibrium price. This results in a premium paid for immediacy, which has been called "the price concession" (O'Hara 1995).

1.2. The Components of the Bid-Ask Spread

The BAS has three components, which reflect adverse selection, order processing, and inventory cost. Informed traders are those who have private and exclusive information about the true fundamental values of underlying commodities. Based on the model from Glosten and Milgrom (1985), liquidity providers are assumed to be uninformed traders. Informed traders buy/sell if the true asset value is expected to be higher/lower than current levels. In a futures market, a zero-sum market, liquidity providers experience losses when trading against informed traders. Therefore, adverse selection exists in market activities mainly due to information asymmetry between liquidity providers and informed traders. This is verified by Stoll (1976), who finds that liquidity providers' inventories decline prior to price increases, and increase prior to prices decreases. Such behavior supports the idea that liquidity providers lose money to informed traders. As a result, liquidity providers attempt to recoup losses by widening the BAS. Working (1967) also finds that larger hedging orders sometimes cause more long-lived market volatility than economists previously expected. Professional scalpers tend to lose money in the absorption of large hedging orders.

[1] A neutral position means that the liquidity provider holds relative balanced short and long contracts for their inventory, so the net position is close to zero.

The order processing cost is a second component of the BAS. This cost refers to labor, computer programming, and other types of fixed and physical costs when liquidity providers process and execute orders. Many past BAS components studies of stock markets such as Huang and Stoll (1997), Stoll (1989), George et al. (1991) find that the order processing cost is the largest BAS component among three, generally 50% or more of the BAS is attributable to the order processing cost. In addition, Lin, Sanger, and Booth (1995) find that the order processing cost has a negative relationship with the order size. Nonetheless, the order processing cost becomes high for the largest orders in the top one percentile. This can be due to the fact that exceptionally larger orders are computationally and technologically difficult to process.

The third cost for liquidity providers is the inventory cost. When supplying liquidity, liquidity providers try to balance their inventory levels to control risks. So they avoid holding too many long or short positions. When prices fluctuate, unbalanced positions significantly increase risks for liquidity providers. However, it is very infrequent that liquidity providers can perfectly balance their long and short positions because buyers and sellers often do not arrive simultaneously. Therefore, liquidity providers often have to adjust their inventory levels to supply liquidity to those who are willing to pay the immediacy (Garman 1976). Amihud and Mendelson (1980) study inventory behaviors from a different perspective, they perceive that liquidity providers operate to maintain their inventory balance at a specific level. BAS increases as a liquidity provider's inventory position deviates from its preferred level, because higher inventory cost must cover liquidity providers' losses from commodity price fluctuations. Overall, the inventory cost arises mainly due to liquidity providers' uncertainty of the return of their inventory and when the next transaction will occur.

1.3. Thesis Contribution and Structure

Although numerous studies have studied the BAS components in the stock market, such studies do not exist in agricultural commodity futures market. This thesis is the first to study the BAS components in the agricultural commodity futures market. The major component of the market transaction cost is the inventory cost, which on average accounts for 53.1% of the BAS in the CBOT corn futures market. The mean adverse selection proportion of the BAS is 12.1%, the least among all three components. Overall, around 30 to 35% of the BAS is attributable to the order processing cost, and the order processing cost component is relatively stable from day to day according to the daily-based analysis. The results also show that the adverse selection component appears to be higher when the corn price is higher and more volatile in 2008 and 2011 than 2009 and 2010. However, the monthly patterns for adverse selection cost and inventory cost vary from year to year.

Based on intraday analysis, unlike the study from Huang and Stoll (1997), the inventory cost is the predominant component of the BAS in the CBOT corn futures market. The intraday inventory cost seems to have a strong inverted U-shaped pattern. The second largest BAS component is the order processing cost; it is stably around 30 to 35% and does not vary much through different trading hours. Approximately 10-15% of the BAS is attributable to the adverse selection cost, the intraday pattern of the adverse selection cost appears to be a U-shaped pattern. In addition, the intraday empirical analysis shows that the size of the components changes on USDA announcement days in the first day-time trading hour. In the first trading hour, the adverse selection costs are higher on announcement days while the inventory costs are lower on announcement days.

The thesis is organized as follows. Chapter 2 reviews recent literature on the BAS components studies in stock markets and commodity futures markets. It also discusses current studies about intraday and seasonal patterns on BAS. Three BAS decomposition models are introduced as well. Chapter 3 covers the conceptual framework and the estimation procedure used for the Huang and Stoll (1997) model. Chapter 4 describes the data which consist of all Best Bids and Offers (BBO) and trades for the CBOT corn futures market from January 2008 to October 2011. It also recounts the process that I take to compose the data. The last part exhibits the data summary statistics and briefly explains recent years' corn price behavior. Chapter 5 provides the empirical results of three-way BAS decomposition for each trading day. Daily and monthly adverse selection and inventory cost are provided, and the adverse selection and inventory components on USDA announcement days are provided. Chapter 6 provides the intraday analysis of BAS components and market liquidity measures, and differences between USDA announcement and no-announcement days. Chapter 7 is for the discussion, suggestions for future work, and conclusions.

2. LITERATURE REVIEW

2.1. Introduction

This section reviews related research related to the thesis. The next section discusses general studies about the BAS, it includes the patterns of the BAS and announcement effects on the market BAS. In section 2.3, I briefly introduce three other BAS decomposition models other than the Huang and Stoll (1997) model.

2.2. Literature on BAS

Literature on the determinants of the BAS is abundant, although most of the studies focus on the stock markets. Many researchers have studied the interrelationships among BAS, price volatility, and volume. In order to investigate the market transaction cost, Demsetz (1968), Stoll (1976), and Glosten and Milgrom (1985) developed earliest order processing, inventory cost, and information asymmetry models respectively. From different perspectives, they all identify that BAS is positively correlated to price volatility and negatively affected by trading volume. Using Ordinary Least Squares estimations, Demsetz (1968) finds that in the NYSE, transaction costs have strong inverse relationships with trading volume. In more liquid markets, the frequency of transactions is greater than less active markets, and the inventory risks are lower for liquidity providers. Therefore, the BAS tends to be lower. From an asymmetric information context, Copeland and Galai (1983) study the liquidity provider's problem by considering the BAS as a straddle option. When a liquidity provider gives out-of-money call and put options to a potential trader, liquidity traders are willing to exercise the out-of-the-money options with a loss, while informed traders will only trade for a gain with their private information. They show that the BAS is a positive function of price level and return variance, and a negative function of market activity and depth.[2] With two financial indices and two metals futures contracts from the CME, Wang et al. (2000) build a simultaneous three-equation system to explain the relations among those three variables. Using a Generalized Method of Moments (GMM) procedure, they reveal that trading volume has a negative relationship with the BAS, but price volatility has a positive relationship with trading volume. The BAS tends to be higher during more price volatile periods.

In recent years, studies on the agricultural commodity futures markets BAS have appeared. With a modified Bayesian method, Frank and Garcia (2010) conclude that volume and volatility appear to be the most influential factors on the effective BAS in livestock markets. Among all the variables that they have examined, trading volume consistently and negatively correlate with the effective BAS.[3] However, this study did not use the tick data to test the interrelationships among those variables. Wang, Garcia, and Irwin (2013) perform a study of the BAS behavior in the corn futures market with the BBO dataset from the CBOT corn futures market. They find that the BAS reacts positively to volatility and negatively to volume, which is consistent with most of previous studies. Their research also provides evidence that price trend, market depth, and other factors that may affect the BAS in addition to volume, and price volatility.

With the corn futures market tick data, Wang, Garcia, and Irwin (2013) show that the BAS of corn futures contracts display a U-shaped pattern through the life of a contract. More distant futures contracts have higher BAS than nearby ones, and this result is in line with Bryant and Haigh (2004). Trading volume tend to be exceedingly small when contracts are initially started.

---

[2] Market Depth is the daily average number of ask or daily average number of bid limit orders to trade (Wang, Garcia, and Irwin 2013).

[3] The effective BAS takes into consideration of the transactions that occur *inside* the best bid and ask. For example, when a transaction occurs at the quote midpoint, the effective BAS is smaller than the quoted BAS since liquidity providers do not receive the full quoted BAS as their market making compensation under such circumstances. The effective spread is often regarded as a more accurate measure of market liquidity cost (Petersen and Fialkowsi 1994).

After this early stage, trading volume and volatility start to grow and remain at high levels. The BAS declines after the early stage then becomes relatively steady until its expiration month. During the steady stage, the BAS is empirically shown to be slightly more than a tick in the corn futures market.[4] When maturity approaches, price volatility and BAS sharply increase because traders try to offset their positions to avoid delivery. Under such circumstances, the immediacy costs for traders become much higher, so the cost of immediacy increases substantially before contracts maturity. The BAS is the lowest in December contracts, and then increases through March, May, July, with a slight decrease in September contracts. Although they find significant relationships among volume, volatility, and the BAS for all contracts, the empirical results for the September contracts are insignificant. Further, they discover that the trading volume of December corn futures contracts start to significantly increase as early as five months in advance of its expiration, while contracts in other months are less active than December contracts. This phenomenon is explained by Smith (2005), December contracts are actively used for hedging since it is the first contract with new crop delivery in each crop year. In contrast, September contracts are usually lightly traded due to the uncertainty of new/old crop delivery.

Other than BAS patterns in the futures market discussed above, BAS also appears to have strong intraday patterns. Wang, Garcia, and Irwin (2012) conclude that the BAS from the corn futures market consistently exhibit intraday L-shape, which contrast from similar studies on open outcry markets. As one of the earliest studies on agricultural commodity markets BAS, Bryant and Haigh (2004) find a weak "reverse-J" pattern of BAS in coffee and cocoa futures markets. Their study is not on the U.S. futures market and they do not discover apparent seasonal patterns

[4] The tick size is the minimum increment in which price can change. The current tick size in the CBOT corn futures market is 0.25 cents/bushel.

of the BAS. The “reverse-J” pattern BAS from Bryant and Haigh (2004) indicates that the BAS increases before market closure, while the “L-shape” BAS from Wang, Garcia, and Irwin (2012) shows that the BAS declines after market opening and remains low throughout each trading day.

The BAS research has been more developed in the stock markets. McInish and Wood (1992) point out in their studies that in the NYSE stock market, the daily minute-by-minute BAS exhibits reverse J-shape pattern. At the beginning of each trading day, the diversity of information may cause traders to have different views of market prices since traders may gather different private information before market openings; accordingly, the price volatility is generally high at the market opening. With a more up to date NYSE stock markets dataset, Tannous, Wang, and Wilson (2013) find a similar intraday BAS behavior. After dramatic declines following market openings each day, the information asymmetry component increases from 12:30 PM-2:30 PM trading intervals. Then the information asymmetry component of the BAS sharply decreases and reaches its lowest level before market closes on each trading day. McInish and Wood (1992) also find that the BAS of each weekday is not significantly different from one another, so no stable weekly BAS patterns were found.

There are numerous studies that investigate the announcement effects on financial and futures markets. Kim and Verrecchia (1994) develop a theoretical model to show that more asymmetric information occurs on announcement days. Although the increased BAS implies less market liquidity conditions at earnings announcement periods, the trading volume may still increase despite less liquid market conditions. Krinsky and Lee (1996) study the earnings announcement effects on each component of the BAS in the NYSE stock markets. They calculate the BAS components of the event, pre-disclosure, and benchmark periods with the Stoll (1990) model, and then make comparisons between the BAS components during events periods with the other two periods. Their empirical results show that the adverse selection cost significantly increases during announcement periods, while the inventory cost and order processing cost decline. Green (2004) examines how the U.S. bond markets react to government announcements by decomposing the BAS with the Madhavan, Richardson, and Roomans (1997) (MRR) model. It suggests the release of public information increases the level of information asymmetry in the bond markets, which lasts only around 15 minutes.

In addition to studies on the announcement effects in financial markets, there is also research on the agricultural commodity futures markets. Among various studies on the announcements effects on the U.S. agricultural commodity markets (e.g. Isengildina-Massa et al. 2008a, 2008b, Lehecka, Wang, and Garcia 2014), Lehecka, Wang, and Garcia (2014) use the BBO data to investigate the announcement effects on the corn futures market from July 2009 to May 2012. They discover that the price reactions are strong immediately after market openings on U.S. Department of Agricultural (USDA) CP, WASDE, GS, PP, and AR reports announcement days.[5] Specifically, the electronic corn futures market incorporates new public information quickly on USDA announcement days, and the market volatility goes back to normal level after 10 minutes of market opening. Wang, Garcia, and Irwin (2013) find that USDA announcements significantly affect BAS in nearby contracts. Especially, the BAS increase by nearly 12% on days when the GS reports are released. They suspect that higher adverse selection cost was the driving force for higher BAS on USDA announcement day.

[5] CP: crop production; WASDE: world agricultural supply and demand estimates; GS: grain stocks; PP: prospective plantings; AR: acreage report.

2.3. BAS Decomposition Models

There are three other major BAS decomposition models before Huang and Stoll (1997) came up with the three-way decomposition model. Glosten and Harris (1988) (G-H) introduce one of the earliest BAS spread decomposition models. They separate the BAS into two components --- adverse selection cost and a transitory cost component. They refer to the transitory component as inventory costs, clearing fees, and/or liquidity provider monopoly profits. Glosten and Harris (1988) use transaction data that contains transaction prices and volume, but no quote information. Therefore, they could not observe the trade indicator (Q_t) directly from their data. Instead, they used a maximum likelihood function method from Friedman and Harris (1998) to generate trade classifications conditioning on the observed volume at time t. The Glosten and Harris' study reveals that large portion of the NYSE BAS is due to adverse selection. In addition, they find that the spread is a function of trade size.

Following Huang and Stoll (1994), Lin (1993), and Stoll (1989), Lin, Sanger, and Booth (1995) decompose the BAS from 150 NYSE stocks in 1988 and explore the empirical relation between the BAS components and the trade size. They conclude that adverse selection increases uniformly with trade size, and that order processing costs has a negative relationship with trade size except for the largest trades. The model by Lin, Sanger, and Booth (1995) is a one-period horizon model and does not capture long-term inventory adjustment effects.

Madhavan, Richardson, and Roomans (1997) (MRR) assume quotes and transaction price changes are determined jointly by a set of parameters B, and $B = (\theta, \phi, \lambda, \rho)$. The adverse selection component from the BAS is appraised by θ ; ϕ, the cost of supplying liquidity; λ, the probability a transaction takes place inside the spread; and ρ, the autocorrelation of the order flow. MRR discover that the order processing component is small and the adverse selection component comprises a major proportion of the BAS. Overall, without considering the trade direction reversal probability, the MRR does not fully decompose the BAS. In their study, MRR also find that their estimation of the BAS is consistently one-third lower than expected, which is a possible outcome of transactions take place at quotes that are not representative.

3. EMPIRICAL MODEL AND ESTIMATION

3.1. Introduction

There are four subsections in this section. Section 3.2 briefly explain the theoretical framework of the Huang and Stoll (1997) (H-S) model and its advantages over other BAS decomposition models. Section 3.3 explains the underlying theory of the H-S BAS decomposition method. Lastly, section 3.4 concentrates on the GMM procedure and moment conditions that I used while estimating estimate the H-S model.

3.2. H-S Model Background

The H-S model completely decomposes BAS into three components, adverse selection, inventory, and order processing costs, and is also called the H-S three-way decomposition model. Under Huang and Stoll (1997) and Ho and Stoll (1981), liquidity providers operate in a competitive environment. After observing the trade direction and BAS, liquidity providers post quotes based on their inventory levels and beliefs of the new equilibrium values. Liquidity providers receive the BAS to recoup their inventory cost and losses against informed traders.

Van Ness, Van Ness, and Warr (2001) examine the performance of five different BAS decomposition models. In their research, there are 12 variables that measure the information asymmetric and volatility of each sampled stock. The adverse selection components are calculated from 5 different decomposition models. They build a three equation system over log (adverse selection component/stock price), log(number of financial analyst following a stock), and the log(trading volume of a stock). While estimating the three equation system, they add 12 other control variables to each equation (in order to control for different volatility and liquidity of different stocks), which are variance of spread midpoint, standard deviation of returns, standard deviation of daily volume, analyst forecast error, dispersion of analyst forecasts, debt/total asset, market-to-book ratio, percentage of institutional ownership, and number of institutional owners. They then estimate the adverse information component of each stock by five different BAS decomposition models including the H-S model. By running the Three-Stage Least Squares Regressions of the adverse selection components over all other information and volatility measurements, they found that adverse selection models are relevant to stock volatility and the presence of informed traders. From the performance test, the H-S model outperforms all other adverse selection models. Because compared with other models, adverse selection estimations from the H-S model best correspond with their other information asymmetric measures. [6]

Overall, the allowance of natural trade direction reversals makes the H-S model the first model that decomposes the BAS into three separate components. In this thesis, I select the H-S model to decompose the BAS of the corn futures market. The next subsection describes the empirical model in more detail.

3.3. Empirical H-S Model

The H-S three-way decomposition model starts by modelling the unobserved fundamental value (V_t) of an underlying asset.

$$(1)\quad V_t = V_{t-1} + \alpha \frac{S}{2} Q_{t-1} + \eta_t$$

In equation (1), the subscript (t) denotes separate but sequential occurrences of events, S denotes a constant spread, α is the percentage of the half-spread attributable to information asymmetry,

[6] See Van Ness, Van Ness, and Warr (2001) for more details.

and Q_{t-1} denotes the trade indicator at t-1. Q_{t-1} equals -1 if it is a public sell and it is +1 for public buy. Lastly, η_t is assumed to be the serially uncorrelated public information shocks. From equation (1), V_t denotes the post-trade asset fundamental value, V_{t-1} is liquidity providers' previous belief on the asset fundamental value. Equation (1) is based on theories developed by Copeland and Galai (1983) and Glosten and Milgrom (1985), who assume that the BAS can be a purely informational phenomenon, occurring even when all liquidity providers have zero inventory and other fixed costs. Copeland and Galai (1983) assume that liquidity providers are risk neutral, and their object is maximizing profits. In addition, information about the realizations of the true asset value is conveyed by informed traders to the marketplace, while liquidity providers and liquidity traders are uninformed. When the liquidity provider has an ex ante expectation on the probability that the next trader is informed, it can set the equilibrium BAS to recover potential losses to informed traders against potential gains from liquidity traders. Private information is revealed immediately after each trade. Therefore, after a transaction occurring at t-1, liquidity providers revise their belief on the asset fundamental value. Equation (1) decomposes the change of fundamental value into two components, which are private information revealed from the last transaction by observing Q_{t-1}, and the public information component that is captured by the error term. The true immediacy cost paid by liquidity traders is the half BAS. When the expectation of the error term is zero, and the spread is constant, the absolute value of the change on the fundamental value, $|\Delta V_t| = (\alpha \frac{S}{2})$, is the fundamental value change that was from the private information. When multiplying the trade indicator, it also reflects the direct of fundamental value change. Therefore, ($\alpha \frac{S}{2} Q_{t-1}$) is the private information that is revealed from the last trade, and thus α is a measure of the asymmetric information component of the BAS.

The fundamental value V_t is hypothetical since we cannot actually observe it, but the quote midpoint, M_t is observable. Based on the inventory theory from Ho and Stoll (1981), and Huang and Stoll (1997), liquidity providers adjust the quote midpoint on the basis of their actual cumulated inventory to their equilibrium inventory level.[7] At time t, the quote midpoint M_t can be expressed as:

$$(2)\quad M_t = V_t + \beta \frac{S}{2} \sum_{1}^{t-1} Q_i$$

The second term on the right side of equation (2) measures the cumulated inventory of a liquidity provider from market open until time t-1, where Q_i is the inventory at time i. The proportion of the half-spread that is attributable to the market marker's inventory cost is β. In the first circumstance, without any inventory cost, the quote midpoint change is exactly equal to the change of asset fundamental value. In contrast, if the quote midpoint change is only driven by inventory cost, then the asset equilibrium price is unchanged, $\Delta V_t=0$. After a transaction occurs, the new quotes are set such that the liquidity provider is indifferent between a transaction at the bid or ask price. Liquidity providers lower the quote midpoint relative to the true fundamental value in order to induce a public buy order in order to neutralize their cumulated inventory after a public sell. In the third circumstance, the change of quote midpoint (ΔM_t) is a result of both inventory and adverse selection costs. For example, when $Q_{t-1} = -1$, which denotes a public sell, then the absolute value of the quote midpoint change will be greater than the absolute value of the fundamental value change for the transaction at time t ($|\Delta M_t| > |\Delta V_t|$). The difference between the quote midpoint change and the fundamental value change can be explained by the inventory cost. Therefore, the total change of quote midpoint is attributable to both adverse

[7] H-S also assume that all trades are of a normal size of one.

selection and inventory costs. When combining the first order difference of equation (2) and equation (1):

$$(3)\ \Delta M_t = \frac{1}{2}(\alpha + \beta)(S\, Q_{t-1}) + \varepsilon_t$$

After observing the trade direction from the last trade, liquidity providers then adjust their posted quotes. Therefore, equation (3) implies that the change of quotes reflect the information revealed by the last trade and the inventory cost that liquidity providers burden from the last trade.

For equation (1) to (3), the underlying assumption is that, from time t-1 to t, the probability of trade reversal (π) is 0.5. However, the probability of trade reversal is generally greater than 0.5 because liquidity suppliers adjust quotes to equilibrate their inventory levels. H-S is the first study that takes into account of the probability of trade reversals in BAS decomposition studies. If π deviates from 0.5, then the probability of $Q_{t-1} = Q_{t-2}$ is $(1 - \pi)$, which means that probability of $Q_{t-1} = -Q_{t-2}$ is π. Therefore, given Q_{t-2}, the conditional expectation of the trade direction at time t-1 is:

$$(4)\ E(Q_{t-1} \mid Q_{t-2}) = (1 - 2\pi)Q_{t-2}$$

When allowing π to be different from 0.5, equation (1) can be modified to account for the predictable information contained in the trade at time t-2. Therefore, the modified equation (1) becomes:

$$(5)\ \Delta V_t = \alpha\ \frac{S}{2} Q_{t-1} - \alpha\ \frac{S}{2}\ (1-2\pi)\ Q_{t-2} + \eta_t$$

H-S describe the second term on the right hand side as the information in Q_{t-1} that is not a surprise. When $\pi = 0.5$, equation (5) is exactly same as equation (1). The changes in fundamental values, ΔV_t, are serially uncorrelated and unpredictable since the changes are induced by unpredictable public information shocks and trades innovations (the first two terms on right hand side of equation (5)). So the expected value of $\Delta V_t = 0$. [8] The combination of equations (5) and the first order difference of equation (2) can be shown as:

$$(6)\quad \Delta M_t = \frac{1}{2}(\alpha + \beta)(S\, Q_{t-1}) - \frac{1}{2}\alpha(1 - 2\pi)S\, Q_{t-2} + \varepsilon_t.$$

When estimating the components of the spread directly from quote data, the constant spread (S) in equation (6) is replaced with observed quote spreads. The new version of equation (6) is written as:

$$(7)\quad \Delta M_t = \frac{1}{2}(\alpha + \beta)(S_{t-1}Q_{t-1}) - \frac{1}{2}\alpha(1 - 2\pi)S_{t-2}Q_{t-2} + \varepsilon_t$$

By combining equations (4) and (7), the three-way decomposition equation system of the H-S model is as follows:

$$(8.1)\quad Q_{t-1} = (1 - 2\pi)Q_{t-2} + \delta_{t-1}$$
$$(8.2)\quad \Delta M_t = \frac{1}{2}(\alpha + \beta)(S_{t-1}Q_{t-1}) - \frac{1}{2}\alpha(1 - 2\pi)S_{t-2}Q_{t-2} + \varepsilon_t$$

The right hand side of equation (8.2) yields the estimation of α and β based on the last observed trade indicator, where α denotes the proportion of the effective BAS that is due to adverse selection, β denotes the proportion of the effective BAS this is due to inventory cost, and the observed trade indicator (Q_t) represents buyer/seller initiated trades. The equation system (8) provides estimations on α and β, anything else that is not capture by this equation is attributable to the order processing cost. In this model, $(1 - \alpha - \beta)$ is the proportion of order processing cost component of the effective BAS. In equation (8.2), the first term on the right hand side is exactly same as the right hand side term in equation (3), except the constant spread is replaced by time

[8] $E(\Delta V_t \mid V_{t-1}, Q_{t-2}) = 0$

varying spreads in (8.2). Furthermore, the first two terms on the right hand side of equation (8.2) can be rewritten as: $\frac{1}{2}\alpha(S_{t-1}Q_{t-1} - (1-2\pi)S_{t-2}Q_{t-2}) + \frac{1}{2}\beta S_{t-1}Q_{t-1}$. The term $(S_{t-1}Q_{t-1} - (1-2\pi)S_{t-2}Q_{t-2})$ is the observed trade indicator at t-1 minus the expectation of itself, which expresses the unpredictable trade innovation at t-1. This unpredictable trade innovation is ascribable to the private information that results in adverse selection cost. Therefore, after observing Q_{t-2} and S_{t-2} at time t-2, the non-surprising proportion of the information is left out of account when calculating the adverse selection component (α) in the H-S model. The second term on the right hand side in both equations (5) and (8.2) represent the proportion of information that is not a surprise.

Huang and Stoll examine 20 most actively traded stocks from the U.S. stock markets in 1992. The major findings are that the order processing cost is the dominant component of the BAS, the adverse selection and inventory holding cost components are small but significant. The average order processing component of all their study samples is 61.8%, and the mean proportion for adverse selection cost is only 9.6%. Findings from Huang and Stoll (1997) contradict both the Madhavan, Richardson, and Roomans (1997) and Glosten and Harris (1988) studies, who conclude that adverse selection cost is predominant among the BAS components.

In spite of its advantages, the H-S model is often criticized for the out-range estimations. Since α is the proportion of the effective BAS that is due to adverse selection, its theoretical range should be between zero and one. However, the adverse selection coefficient (α) sometimes falls outside of the interval [0,1] (i.e. Huang and Stoll 1997, Serednyakov 2006, Henker and Wang 2006). Although negative α coefficients are empirically possible, theoretically it is not sensible. Negative adverse selection cost indicates that liquidity providers are posting quotes are better informed than investors, which contradicts with the theories of the adverse selection models. There have been several studies that provide revised H-S models. Serednyakov (2006) revised the model to allow for trade direction positive serial covariance, and most of the (α) coefficients are between 0 and 1 in their empirical results. However, positive serial covariance contradicts the underlying theory of the H-S model, where negative serial covariance is required. Huang and Stoll (1997) think that the practice of a trader splitting a large order into many smaller orders is the cause of positive serial covariance. Their proposed solution is to merge all split orders and then estimate their model, which is to treat a cluster of split orders to be only one transaction. In section 4.4, I will discuss in more detail about how sequential orders are bunched.

3.4. Estimation Procedure of the H-S Model

In this thesis, I apply the GMM method to estimate the H-S model, the same as the study of Huang and Stoll (1997). Hansen (1982) proves that the GMM estimators are asymptotically normally distributed. The advantage of opting the GMM procedure is that it imposes weak distributional assumptions on the errors terms, δ_t and ε_t. The GMM chooses parameter values for β such that the moment conditions, which is denoted by $g_T(\beta)$, closely approximate the underlying population moments. The estimated covariance matrix of $\hat{\beta}$ equals $V_{\hat{\beta}} = [D_0 S_0^{-1} D_0]^{-1}$, where $D_0 = E[\frac{\partial g_T(\beta)}{\partial \beta}]$ and $S_0 = \sum_{l=-\infty}^{l=+\infty} E\,[f_t f_{t-1}]$, and f_t is the vector of arguments that defines the expectation of moment conditions. For the three-way decomposition form of the H-S model, the corresponding moment conditions are:

$$g_T(\beta) = \begin{Bmatrix} \delta_{t-1} Q_{t-2} \\ \varepsilon_t S_{t-1} \\ \varepsilon_t S_{t-2} \\ \varepsilon_t Q_{t-1} \\ \varepsilon_t Q_{t-2} \end{Bmatrix} = 0$$

Where $\beta = (\alpha\beta\pi)'$, which are the parameters of interests in the H-S model. For the estimation process, I use the error term estimator from Newey and West (1994). In addition to the Newey and West (1987), the Newey and West (1994) research develops an automatic lag selection technique in covariance matrix estimation. The Newey-west error term estimation is a good extension of the Huber-White robust variance estimator, which only handles the presence of heteroscedasticity. With unknown forms of disturbance autocorrelation and heteroscedasticity, the Newey-west error estimation method is a good extension that produces consistent estimators under more complicated circumstances. Furthermore, as both Newey and West (1994) and Andrews and Monahan (1992) demonstrate, conducting the pre-whitening with a first-order vector auto-regression prior to application of the Newey and West (1994) procedure improves the size of the test statistics, so that the estimations are both more consistent and efficient. Moreover, the choice of kernel method can sometimes be more important than number of lag selections, instead of the quadratic spectral kernel, I use the Bartlett kernel as Newey and West (1994) suggest in their research.

4. DATA

4.1. Introduction

This section explains the structure of the data and the steps taken to process the data. In the next subsection, I present the general structure and features of the data. Section 4.3 covers the contracts and days that are excluded from this study and the explanation. In section 4.4, I introduce the method that implemented to merge all sequential transactions. Section 4.5 discusses the corn futures price behavior and summary statistics for the BBO dataset. In the last subsection, I discuss the identification of the USDA announcement days.

4.2. Data Structure

The data are from the CBOT BBO database. The data contains electronic GLOBEX trading records for each corn futures contract from 01/14/2008 to 10/31/2011. Each best bid price pairs with a best offer price, along with the number of contracts at the specific BBO quotes. When a better bid or offer prevails, a new pair of BBO quotes are generated and recorded. This top-of-the-book data includes top bid, bid size, top ask, ask size, transaction volume, and transaction price, with a time-stamp for each quote revision to the nearest second.

For each futures contact, the CBOT records all quotes and trades with continuous serial numbers. Each pair of BBO quotes is assigned a new serial number and an updated time stamp. If a trade occurs right after a certain pair of BBO quotes, the trade serial number is continuous with the BBO quotes serial number.[9] Understanding this important fact of the data structure

[9]The quotes and trades serial number in the data is created with a minimum increment of 10. For example, if a best bid and offer is recorded with a serial number of a, and a transaction is triggered by this best bid and offer, this transaction is recorded with serial number of $a + 10$. If the best bid and offer changed after transaction $a + 10$, the

enables me to match each transaction with the quotes that triggered that particular transaction. If the demanded transaction size is larger than the bid/offer size from the top of the book, then part of the order size will be executed according to the second best bid and offer prices. Under such circumstances, I still use the original best bid and offer quotes to match the entire transaction.

For each pair of BBO quotes, the BAS is the best offer minus the best bid. The spread midpoint is the arithmetic mean of the best bid offer prices. The volatility measurement is the standard deviation of intraday quote midpoints in cents/bushel. Daily volatility is measured by the standard deviation of quote midpoints using only the quotes at the time of trade each day, this method is same as Chung and Kim (2009). I use the quote midpoints instead of trade prices so that the volatility measure is not subject to bid-ask bounce effect. Wang, Garcia, and Irwin (2013) use the same method to measure return volatility but they calculate the standard deviation of quote midpoints based all recorded quote revisions. In the thesis, I mainly study bid-ask quotes at the time of transactions, so I adopt the volatility measure from Chung and Kim (2009). The monthly volatility is the mean of all trading days' volatility measurements for each month. The same process is applied to calculate yearly volatilities. The trade indicator is generated by applying the procedure that was developed by Lee and Ready (1991). If the transaction price is higher than the quote midpoint, it is a buyer-initiated transaction. Seller-initiated transactions are those transaction prices that are lower than the quotes midpoints.

Based on the H-S model from equation (8.2), I need quote midpoint change, BAS, and trade indicators to conduct the empirical analysis. For each transaction in the dataset, I combine the transaction price and the bid-ask quotes recorded immediately before this transaction to be one

---

new best bid and offer quotes will be recorded with $a + 20$. The data recording process from the futures market is essentially different from the stock market, where the transaction data and quotes data are recorded separately.

observation. So that each observation includes transaction price, trade size, bid and price quotes, and sizes (see Table 1 for details).

4.3. Contracts/Trading Days Selection

According to Smith (2005), due to the uncertainty of old/new crop delivery, the September corn futures contract in each year is usually lightly traded. Wang, Garcia, and Irwin (2013) find similar results that the September contracts have the least trading interest to traders among all contracts. However, each December contract becomes actively traded as early as five months prior to its expiration. Therefore, the September corn contracts may not reflect public information as fully and efficiently as the December contracts. In this thesis, I exclude all the September contracts from 2008 to 2011, and roll over to the first deferred contract in the first trading day of the contract settlement month. The December contract becomes nearby on July 1 each year. Lastly, the data are missing data from January 1 to January 14 in 2009, and April 15. In addition, I exclude July 7, 2008 and July 5, 2011 because there are exceedingly low trading volume in these two days after the July 4 holidays. There are two other days, October 8, 2010 and March 31 2011 that have identical trade indicators throughout the day.[10]

Irwin and Sanders (2012) discover that approximately 92% of the total trading volume in the corn futures market are electronically traded by 2009. In my dataset, the average daily electronic trading volume (both day and evening sessions) is around 83,035 contracts. The total daily trading volume data for the nearby corn futures contract is also obtained from the Commodity

[10]Both October 8, 2010 and March 31 2011 are USDA reporting days. October 8, 2010 is the WASDE and Crop Production reports days, while March 31 2011 is the prospective planting acreage report day. The bid, ask, transaction prices, and trade indicators barely change during those days. As I will show later on, under such conditions, it is not possible to estimate the H-S model due to zero trade indicator covariance. However, it also show that the adverse selection components in those days are exceedingly high.

Research Bureau (CRB). I calculate the ratio of daily recorded electronic trading volume over the total daily trading volume. The mean ratio for all 945 days is 61.19%. In other words, on average, only 61.19% of the trading volume is documented in the BBO dataset. The trading volume from the CRB includes number of contracts traded for all CME Group venues, the Globex, open outcry, ClearPort/PNT, and all other executions.[11] In contrast, my dataset only shows trading volume from the Globex trading platform.

Out of 945 sampled days, there are 25 days in which electronic trading volume from the BBO dataset are below 20% of the total volume from the CRB. The BBO tick data may not be able to fully reflect the market quotes and price changes when too few contracts are electronically traded. Under those situations, I often find the adverse selection component (α) estimations to be much larger than 1, which is theoretically unreasonable. When large and information trading are pre-negotiated and not electronically traded, it will shock the electronic trading market. On days that electronic trading volume is little compared with the CRB volume, electronic transactions occur less frequently than normal but with relatively larger price jumps. When the price moves up or down largely (i.e. much more than a tick) from the last transaction, the change of quote midpoint is also large but the BAS may be still a tick. Based on equation (8.2), larger change of quote midpoint and small BAS will result in very large and unreasonable α estimations.

With 945 trading days in the full sample, 83 α coefficients fall outside the interval of [0,1], of which 5 of them are greater than 1 and 78 of them are smaller than 0. For example, the smallest α coefficient appears on 04/29/2011, and only 17% of the total volume was record in the data. The

[11] ClearPort is CME Group's clearing service for Over-The-Counter markets. PNT is for Privately Negotiated Trades.

largest α coefficient appears on 04/29/2008, and the GLOBEX electronic trading volume is only 11% of the CRB total trading volume. In contrast, vast majority of β coefficients fall inside the interval of [0,1]. Out of 945 trading days, only three β estimations are greater than 1 and eight β estimations are less than 0. Such out-of-range estimation issues from the H-S model has also been pointed out by previous research (i.e. Clarke and Shastri 2000, Van Ness, et al. 2001). Therefore, in this thesis, I only select days that have 50% or above recorded electronic trading volume over total volume. Using this filtering standard, 763 out of 945 days are selected.[12]

4.4. Bunching Sequential Trades

H-S and other researchers have shown that the H-S model estimates are downward biased when many transactions occur without any change of price and quotes. Traders separate large orders into small ones to prevent leaking their private information. As a consequence, I identify sequential transactions and bunch them into a single transaction. There are different ways to identify sequential trades. Shah and Brorsen (2010) assume that orders are split if they are recorded in the same second. For my dataset, there are often many transactions occuring in a second, and quotes and/or transaction prices change within a second. In Huang and Stoll (1997), if more than one consecutive transactions have the same bid or ask, same execution price, and trade direction, these trades are defined as sequential trades. However, H-S also argue that their bunching procedure overcorrects the order split issue due to aggregating split orders and independent orders all together. Their empirical results of adverse selection components and probability of trade reversal are more appropriate to be regarded as upper bounds of those coefficients. Here, I revise the H-S bunching procedure and only merge all sequential

[12] Please see Table 2 for more details about different volume filtering standards and outcomes.

transactions if they occur within the same minute. All sequential transactions are collapsed into one, and their volume is aggregated and the last transaction time stamp is retained for the merged transaction (see Table 3 for an example). After all sequential transactions are bunched, the bunched data is what I use to the H-S three-way decomposition model.

4.5. Recent Corn Price Behavior and Summary Statistics

Figure 1 demonstrates the trend of the nearby corn futures contract daily settlement price. The time interval from this figure covers the daily corn prices from 01/14/2008 to 10/31/2011. The price of corn increased substantially from January to early July 2008. Soon after the rapid price increase, it slumped from above $7/bushel to around $4/bushel by the end of 2008. Although corn price fluctuated in 2009 and 2010, the magnitude of the 2009/2010 price volatility is comparatively less than in 2008. By the end of 2010, corn price started a new and abrupt upward trend. From late 2010 to summer 2011, the corn price more than doubled from around $3.5/bushel to more than $7.5/bushel. In 2011, the corn price moved up and down several times and the overall price level was generally much higher than 2009 and 2010.

There are generally more than 40,000 recorded bid-ask quote revisions in each trading day. However, only around 12,000 to 18,000 transactions occur during each day time trading hours from 9:30 AM to 1:15 PM (See Table 4 Column 6 and 7). Table 4 also provides descriptive statistics for the daily average of major measurements for the 763 trading days in my sample. With high price volatility, the mean quoted BAS in 2008 and 2011 are higher than the other two years. The 2008 average BAS, 0.271 cents/bushel, of corn nearby futures is the highest in the sampled period. It is around 8.1% higher than the current one tick size in the CBOT corn futures market. From another perspective, the 2008 BAS is 6.06%, 6.45%, and 3.36% higher than the

BAS in 2009, 2010, and 2011 respectively. The volatile years of 2008 and 2011 have a similar average BAS, while the stable prices years of 2009 and 2010 are similar to one another but have smaller average BAS than the volatile years. From 2008 to 2011, the BAS and volatility measurements are highly correlated; years with high BAS are also years with high volatility. This is in line with findings by Wang, Garcia, and Irwin (2013). On the context of their study, high volatility may be explained by large information inflows into the market. Furthermore, the average daily electronic trading volume increased 30.8% from 2008 to 2011, but the standard deviation of the daily trading volume in 2010 is much higher than all other years.

Some of the effects from bunching the data are provided in columns 5 to 8 in Table 4. Before sequential transactions are bunched, the day session average transaction size is approximately 5.84 contracts/transaction from 2008 to 2011. The average transaction sizes are not significantly different among all four years. After bunching all sequential transactions, the average transaction size dramatically increases for all years. The transaction size in 2010 is the highest both before and after data bunching. The last two columns in Table 4 are numbers of transactions during daytime trading sessions. Before the data bunching process, the 2011 daily average transaction count is the highest, which corresponds to the largest trading volume in that year. When all sequential transactions are bunched, the 2008 daily average number of transactions decreases by 53.81%. The transaction counts in 2010 decreases from 13,243 to 4,597 trades/day after bunching, this 65.28% decrement of daily number of transactions is the highest in all four years. This fact indicates that in 2010, many consecutive transactions are executed with the same price and quotes. The average transaction size is also the largest in 2010. It means that traders tended to submit relatively larger orders during that year because the corn price was stable.

4.6. Identification of USDA Announcement Days

The USDA releases different reports regarding the crop demand, supply, yield, and other key measurements of agricultural commodities. The Crop Production (CP) reports are generated by the National Agricultural Statistics Service (NASS), which includes acreage, yield per acre, total production, and general descriptions of crops' production. The World Agricultural Supply and Demand (WASDE) reports provide farmers and market traders with more comprehensive views on world supply, demand, stock levels, imports, exports, and etc. Both of these two reports are released around the 10^{th} of each month. During my sampled periods, all reports were released at 8:30 EST before the market opening. Here, I study the same USDA reports as Lehecka, Wang, and Garcia (2014) do in their research. The reports are WASDE, CP, Grain Stocks report (GS), Prospective Plantings report (PP), and the Acreage report (AR). The regular USDA announcements disclose important public information into the markets. Different capability of traders to process public information may cause higher adverse selection component on USDA announcement days. In addition, it is also important to discover how fast does it take for markets to react to the announcements. In later sections, I will study these issues from different approaches.

5. BAS THREE-WAY DECOMPOSITION

5.1. Introduction

Section 5.2 provides analysis on the daily adverse selection and inventory cost components. For each trading day, I apply the H-S model once and generate the adverse selection (α) and inventory cost (β) component estimations. Section 5.3 discusses the monthly average of the daily adverse selection components. The USDA announcement days BAS components are shown and discussed in section 5.4.

5.2. Daily BAS Components

For the 763 trading days that remain after filtering, the averages of the adverse selection, inventory, and order processing components are 12.1%, 53.1%, and 34.8%, respectively. The estimate of the adverse selection component is similar to Huang and Stoll (1997). Figure 2 plots the daily adverse selection and inventory cost components. Out of 763 trading days, 531 of the adverse selection component and 762 of the inventory cost component (β) estimations are significant at the 5% significance level. In particular, most of the adverse selection component (α) estimations are significant in 2008 and 2011. In 2009 and 2010, many α estimations are lower than the overall average and no statistically different from 0 (See Table 5 for statistical significance summary of all coefficient estimations). In general, the β coefficient is larger than the adverse selection component, and it is also statistically significant.

The empirical results show that the adverse selection proportion is small, and the major cost component of the BAS is inventory cost, which deviates from the H-S conclusion. On average, from 2008 to 2011, around 34.8% of the BAS is attributable to the order processing cost — larger than the adverse selection cost but smaller than the inventory cost. Order processing cost represents the labor, computer programming, and other types of fixed costs. It is expected that order processing cost is lower here compared with Huang and Stoll (1997) since their study was from 1992. The rise of electronic trading utilizing computers and networks has naturally reduced order processing cost since the original H-S study. In the short run order processing cost is determined by labor and capital cost, so it does not vary much. The adverse selection component of the BAS was unusually high in the summers of 2008 and 2011, periods in which corn and many other commodity prices spiked. In June 2008, the price of corn went up by more than $1/bushel in that single month, which is the highest monthly price increment in the period consider. For the first half of 2009 and the entirety of 2010, asymmetric information was mostly below four year's average. Not only the proportion is very small, the quoted BAS is generally below 0.26 cents/bushel in those two years as Table 1 shows, so the magnitude of transaction cost is also low. When not much information flows into the futures market, the transaction prices tend to bounce between the bid and ask quotes, and the probability of trade reversals is closer to 0.5 (Roll 1984).

Table 6 demonstrates the monthly adverse selection cost, inventory cost, the BAS, and volatility measures. In addition, the trade indicator covariance is shown for both the bunched and unbunched data. In the 3rd column of Table 6, the covariance of the trade indicators are all negative but more negative in 2009 and 2010 than the other two years. This shows that the trade direction is more likely to reverse; in other words there was a paucity information that would cause the market to lean toward one side.[13] In the same table, it also demonstrates that the volatility in 2009 and 2010 is not as high as in 2008 and 2011, which can be another indication

[13] Since the Cov(Q_t, Q_{t-1}) =(1-2π), negative trade indicator covariance means that trade direction is more likely to reverse than when it is positive. The lower bound of the trade indicator covariance value is -1, when Q_t is always the opposite of Q_{t-1}.

of low level of information inflow. The monthly average asymmetric information component in June 2008 is not only the highest in 2008 but also in all months from January 2008 to October 2011. Soon after June 2008, the corn price declined sharply, and we also see that the adverse selection component gradually decreases through the entire second half of 2008. Based on these empirical results, the monthly adverse selection component and the BAS in the corn futures market were high in both 2008 and 2011. Intense and unpredictable information inflow to the corn futures market leads to higher market BAS, the adverse selection cost also becomes a larger proportion of the BAS. Therefore, the magnitude of the adverse selection cost becomes high.

It is also interesting to notice that not only the adverse selection costs (α) are relatively low in 2009 and 2010 but also the inventory costs (β) (Table 6). When price is generally less volatile, the inventory risk for liquidity providers can become smaller compared with high price volatile periods. In general, the adverse selection and inventory components tend to have a negative correlation according to the daily analysis. Indeed, the correlation between the adverse selection and inventory components is -0.515. The summation of adverse selection and inventory components has to be less than 1, so the negative correlation is expected. More importantly, liquidity and volatility both are important factors to influence the liquidity provider's inventory cost. When the adverse selection cost component is relatively low and not much new information flows into the market, the inventory cost may become high due to low liquidity. When the market is less liquid, it is relatively more difficult for liquidity providers to balance their inventory positions. From the volatility point of view, higher volatility can also lead to higher inventory cost. When prices are volatile, the inventory value can change quickly and unpredictably. When the corn price was more volatile in 2008 and 2011, more than 55% of the BAS is attributable to the inventory cost. Although trading volumes were higher in 2008 and

2011 than the other two years, the exceptionally erratic corn price was possibly the main driving force for high inventory cost component. The inventory cost components from 2010 are mostly below the four-year average value of 53.1%. Table 4 shows the relatively low volatility and high trading volumes in 2010, which can be the explanation for the relatively low inventory cost in 2010. When the market is liquid and less volatile, liquidity providers concern less with their inventory cost, because the value of their inventory does not fluctuate much and they can also balance their inventory positions more easily.

With relatively low adverse selection and inventory cost in 2009 and 2010, the market order processing cost becomes higher in those two years. The respective order processing cost in 2009 and 2011 are 37.1% and 44.5%, while the respective order processing cost for 2008 and 2011 are 25% and 29.4%. As it was shown in Table 4 (columns and 4 and 5), the average size of transactions in 2009 and 2010 are larger than in 2008 and 2011.Previous research has found positive relationships between order processing costs and order sizes. As a result, larger order processing costs in 2009 and 2010 could be caused by bigger order sizes. Huang and Stoll (1997) show that the inventory cost increases as order size increases. However, we are seeing slightly lower inventory cost in 2009 and 2010 than 2008 and 2011.

5.3. Monthly Adverse Selection Component Pattern

The monthly average adverse selection component is plotted in Figure 3 for all years. The adverse selection cost was higher and more volatile in 2008 and 2011, and the plotted curves for 2009 and 2010 are flatter than the other two years. In the U.S., the earliest corn planting starts in late March, while the latest can take place in May. Adverse selection component drops from March to April in all four years. Suggesting that some traders have private information regarding the progress of planting. Figure 3 shows that all years have decreasing trends of the adverse selection component from October to November, key harvest months. Some traders may have better private information regarding the true production level early in the harvest, but as time goes on a more accurate picture emerges to the public; hence, the asymmetric information level tends to decline. Although it is intriguing to see that the asymmetric information component rises again from November to December 2008. Increased adverse selection components during the last month of 2008 may be due to uncertainty of next year's planting plans or consumption levels. Also, the true production level of each year may be the most accurate in the January United States Department of Agriculture (USDA) reports from following years. Therefore, the upward trend of α in December 2008 may be a result of crop production revisions regarding the true yield and consumption. Overall, there is not an obvious market adverse selection seasonal trend that applies to all four years. In the next subsection, I analyze daily adverse selection, and inventory cost components on USDA announcement days, and then make comparisons with the weekly and monthly averages.

5.4. Daily BAS Components on USDA Announcement Days

With the empirical estimations of adverse selection for both USDA announcement and no-announcement days, my first step is to compare the no-announcement daily asymmetric information components with announcement days. Table 7 shows all the adverse selection (α) and inventory cost (β) for each of the announcement day, and the p-values for the coefficients. The adverse selection and inventory cost components are estimated for each day. In addition, the weekly and monthly α and β coefficients are calculated by taking their averages of every day's coefficient excluding the announcement days. There are a total of 39 reporting days that are listed in Table 7. For each announcement day, at least one of the WASDE, CP, GS, PP, or AR is released. Each estimated adverse selection and inventory cost component is between 0 and 1, while three of them are not significant at the 5% level. Based on the entire day analysis, the overall mean of daily α on all USDA announcement days is 15.028%. The average adverse selection component on the announcement days in the volatile years of 2008, and 2011 are 24.187%, 14.144% respectively; but the values of this measurement are lower in less volatile years of 2009 and 2010, at 10.404%, 8.532% respectively.

The highest announcement day adverse selection component happened on March 31, 2008, which was 0.606, when the GS and PP reports were released and the market demand was positively shocked. Along with huge price spikes of many other commodities, corn price also increased sharply in March 2008. In the March 2008 PP report, the USDA announced that the 2008 intended planting acreage for corn was 86.0 million acres. That was an 8% decrease from 2007, which had historically high corn planting acreage. Also, the market expectations for the corn stock was around 7.076 billion bushels but the USDA reported the stock level to be only 6.859 billion bushels. In addition, the winter pig crop that year was 7 percent higher than the previous year (Good 2008). With strong corn demand outlook due to ethanol production, and animal feeding, the significant decline of corn planting acreage and stock level caused high adverse selection cost and intraday volatility. Indeed, the respective price volatility measurement (standard deviation of quote midpoints) and BAS of March 31, 2008 are 6.993 and 0.293. Both of these numbers are much higher than the yearly and monthly averages which can be seen in Table 4.

The second highest α coefficient also happened in 2008, which was 0.507 on August 12. The high adverse selection cost on the August 2008 announcement day may be attributable to the Midwest flood of that year. The June 30, 2008 report was the first USDA official report that produced relatively precise estimates of the flood's impact.[14] More importantly, on June 19, 2008, the USDA notified the public via a news statement that more extensive and accurate results of the flood effect would be released in the August CP report.[15] In the 2008 August CP report, the reported after-flood corn production was around 12.288 billion bushels, which was actually 350 million bushels over the average market expectations (Good 2008). Therefore, the asymmetric information proportion is high on the August 2008 CP report day and stands out among all USDA announcement days. In general, most of the announcement day adverse selection components are below 20%, with only a few of them above 50%.

The respective average inventory cost components on the USDA announcement days in 2008, 2009, 2010, and 2011are 50.266%, 54.509%, 47.80%, and 52.662%. When the adverse selection costs were exceedingly high in March and August 2008 announcement days, the inventory cost in those two days appeared to be both below 30%, which are unusually low. However, the order processing costs were not much different from the yearly averages. Although when the corn price was volatile and the adverse selection costs were high during 2008, the order processing cost was still stable as it is also shown in other paragraphs. With relatively stable order processing cost component, it is reasonable that the inventory cost component and adverse selection components seem to have a negative relationship. When the adverse selection component is the predominant component of the BAS, the proportion of the inventory cost

[14] The June 30, 2008 estimation is dropped because the recorded volumes from the BBO data is less than 50% of the CRB recorded trading volumes.

[15]For more details, see *http://www.usda.gov/wps/portal/usda/usdahome?navid=LATEST_RELEASES&parentnav=NEWSROOM&navtype=RT&edeployment_target=archived&edeployment_action=latestreleases*

declines. On USDA announcement days, liquidity providers expect higher probability to trade against informed traders and set a wide BAS accordingly.

Considering the fact that the corn price and BAS components estimations differ in each year, I decided to apply the paired t test instead of the unpaired t test, which allows me to compare the adverse selection and inventory cost components among the USDA announcement days, corresponding weeks, and corresponding months. The t test shows that the mean difference between the announcement day and weekly adverse selection components (α) are not statistically different at the 5% significance level; the monthly α coefficients means are also do not statistically differ from the announcement day α's. The same paired t-test is also applied to the inventory cost component (β). The result shows than neither weekly nor monthly average β's is statistically different from the inventory cost component on announcement days. Therefore, I do not find significant difference on either BAS component between announcement and no-announcement days here. In section, 6.3, I will further analyze the USDA announcement effects based on the intraday BAS components analyses.

6. INTRADAY ANALYSIS AND USDA ANNOUNCEMENT EFFECTS

6.1. Introduction

In this section, I conduct the intraday analysis of the BAS components and then make comparisons of each component between the USDA announcement and no-announcement trading days. Section 6.2 explains the intraday patterns of adverse selection component, inventory component, market volatility, BAS, number of transactions, and the trading volume. Section 6.3 discusses the USDA announcement effects on the corn futures market. It shows the intraday comparisons between the announcement and no-announcement trading days for all six measures listed above. The empirical results show that the intraday adverse selection cost in both announcement and no-announcement days appear to be U-shaped, while it is higher during the first trading hour on the USDA announcement days.

6.2. Intraday BAS Components Behaviors

Because trading volume and BAS appear to be U-shaped throughout the trading day (Wang and Garcia 2012), it is likely that the components of the BAS are non-constant throughout the trading day as well. In my sample period, the daytime CBOT corn futures market session was from 9:30 to 13:15 (3 hours and 45 minutes). I separate each trading day into 7 different time intervals — with six 30-minute intervals plus a final 45-miunte interval at the end of the trading day (12:30:01-13:15:00). I then estimate equation 8.1 and 8.2 for each time interval from each day. I drop all days in which any interval contains less than 3 transactions (after bunching), if there is any interval in which there is no change in the trade direction, or if there is any interval in which there is an unchanged BAS.[16] There are 707 days out of 763 days left in my sample. In the remaining sample, 39 days are USDA announcement days.

Figure 4 also displays comparisons of USDA announcement days with no announcement days. I estimate equations 8.1 and 8.2 for each time interval on each day that does not contain a USDA announcement (668 days in total). Then, the BAS, the adverse selection component, inventory cost component, and other market measures are averaged across days and displayed in Figure 4 oriented with the appropriate time interval on the x-axis. The dash lines in Figure 4 reflects the same procedure applied to only USDA announcement days (39 days in total). In the remainder of this section I will discuss the no announcement day results; discussion of USDA announcement days will follow in section 6.3.

The top left plot in Figure 4 shows that the adverse selection component sharply decreases from the first to the second time interval. Therefore, the probability of liquidity providers trading against informed traders is high in the first 30 minutes after market openings, and more than 15% of the BAS is attributable to the adverse selection cost during that time. However, from 10:00AM-12:30PM, the market adverse selection cost is on average less than 10% and appears to be very steady throughout this 2.5-hour time interval. In the last 45 minutes of the daytime session, the adverse selection cost is around 15% of the BAS, similar to the market open and much higher than in the mid-day hours. The U-shaped adverse selection pattern indicates that asymmetric information is high during both market opening and closing periods, and the proportion of such cost is stable and lower during the mid-day hours (10:30AM-12:30PM). The results match previous studies from the foreign stock markets and foreign futures market (e.g.,

[16] The H-S model does not produce estimations with less than 3 observations, unchanged trade direction, or unchanged BAS.

Ryu 2011, Ahn et al. 2002). However, my findings also differ from other studies such as MRR (1997), who find that the intraday adverse selection cost appears to have a L-shape, so it consistently declines throughout each trading day. They claim the intraday L-shaped adverse selection cost in the NYSE stock markets is due to liquidity providers' learning process. By the end of each trading day, liquidity providers learn more about the fundamental values of each stock, so they become less dependent on only surmising information from the order flow patterns. Anh et al. (2002) explain the U-Shaped pattern of the adverse selection cost as informed traders taking advantage of high liquidity and low transacting cost during market opening and closing periods. From another point of view, they also argue that informed traders tend to close their intraday positions before market closure to avoid overnight price fluctuation risks.

I also investigate the intraday patterns of the inventory cost, price volatility, BAS, number of transactions, and volume. In Figure 4, the intraday inventory cost component appears to be an inverted U-shape. This particular pattern shows that the liquidity providers' inventory cost is the lowest during market opening and closings. Since a big proportion of informed traders choose to liquidate their positions before market closure, the liquidity providers' inventory cost sharply decreases during the last 45 minutes of each trading day. By the end of each trading day, liquidity providers' net positions are relatively "neutral" so that at the beginning of the next trading day, their net positions are also relatively neutral. Because the night trading volume is only a small proportion of the total electronic trading volume, it is unlikely that liquidity providers build either large long or short positions in evening trading sessions.

From Figure 4, the inventory cost component is low during the first and last time intervals. Simultaneously, market volatility and trading volume is both high. Here, liquidity seems to have a greater impact on liquidity providers' inventory cost than volatility. During market openings and closings, liquidity providers can easily liquidate or balance their inventories due to relatively large number of transactions and volume. Under more liquid market conditions, there is also less uncertainty about liquidity providers' inventory return. During mid-day trading hours, from 10:00 to 12:30, the market is generally less volatile and less liquid. More importantly, the inventory cost component is much higher in mid-day hours than openings and closings, which can be resulted in low liquidity.

The middle right and bottom right plots of Figure 4 exhibit that the number of transactions and trading volume is both the highest in the first time interval (9:30-10:00) from January 2008 to October 2011. Therefore, the corn futures market is more liquid during the first 30 minutes of each trading day than the mid-day trading hours. In a normal trading day, the average number of transactions in the first 30-minute time interval is slightly more than 4,000. However, it dramatically drops more than 50% to around 2,000 transactions in the second 30-minute time interval. The intraday trading volume behave similarly to the transaction counts, and the average trading volume in the first 30 minutes of no-announcement trading days is approximately 25,000 contracts and it declines to around 10,000 contracts after 30 minutes. After the first trading hour each day, both the transaction counts and trading volume seem to be stable until the last 45 minutes of day trading sessions. On average, there are less than 2,000 transactions and 10,000 contracts for every 30-minute time interval during the two mid-day trading hours (10:30AM-12:30PM). Before market closing, both the transaction counts and trading volume increase, which can be due to the fact that many informed traders do not keep overnight positions so they liquidate their positions before market closure as Anh et al. (2002) demonstrate.

The top right plot in Figure 4 displays the intraday change of the BAS. During day time trading sessions, the BAS appears to have an L-shaped pattern, with the highest BAS occurring in the first 30 minutes after market opening and declining to a steady level afterwards. This result corresponds to the finding from Wang, Garcia, and Irwin (2012). Here the BAS in the first 30 minutes is around 0.265 cents/bushel, which is only 5.8% higher than the current minimum tick size in the U.S. corn futures market. When BAS reaches a steady level from 10:00-13:15, it is approximately 0.260 cents/bushel. Although the difference between 0.265 and 0.260 is minimal, the clear intraday L-shaped BAS pattern is apparent.

Wang, Garcia, and Irwin (2013) find a strong and significant negative relationship between trading volume and BAS, and a significant positive relationship between volatility and BAS for both the nearby and first deferred corn futures contracts. Wang and Yau (2000) also discover an inverse relationship between trading volume and the BAS of the S&P 500 index futures market in the CME from 1990 to 1994. However, from my intraday analysis results, the BAS is high during the first 30 minutes after market opening. Given prior research, it is surprising to see concurrent high transaction counts, trading volume, and BAS in the first 30 minutes of trading, however volatility is also at its highest level in this time interval so it appears volatility is the dominant determinant of BAS. Wang, Garcia, and Irwin (2013) study the overall relationships among the BAS, volume, and volatility for each contract, while I focus on the intraday analysis here. So, Wang, Garcia, and Irwin (2013) may have averaged away the results we see in the open and closing periods of the daytime session. Green (2004) also finds concurrent existence of high liquidity cost and high asymmetric information cost in the U.S. Treasury market from 1991 to 1995 around macroeconomic announcements. Green (2004) argues that despite high adverse selection cost around announcement periods, the Treasury market is liquid because the release of economic information generates uncertainty about the appropriate level of the new fundamental value of underlying asset.

Table 8 also demonstrates the intraday order processing cost component (1- α- β). This cost component seems to be higher in market opening and closing than mid-day hours; however, the difference is really small. Throughout each trading day, the order processing cost is always between 30 and 40 percent of the BAS, which appears to be more stable than the other two cost components. Slightly higher order during the first and last time intervals may indicate that order processing costs increase when trading volume increases.

6.3. Intraday USDA Announcement Effects

In contrast to the daily adverse selection and inventory cost components comparisons between USDA announcement and no-announcement days in section 5.4, the intraday analyses of the BAS components and other market liquidity measures here provide more insights of the USDA announcement effects on the corn futures market. Markets are more volatile and information-asymmetric during opening and closing hours, while mid-day market conditions are more stable (e.g. Ryu 2011, Ahn et al. 2002, etc). When estimating each day as a set of observations with the H-S model, the model does not capture intraday market conditions. Therefore, the entire day BAS component estimations are similar for all trading days. However, the intraday analysis exhibits apparent information asymmetry and inventory cost component differences between USDA announcement and no-announcement days.

The general intraday patterns for all variables are very similar on both announcement and no-announcement days, the actual values for each variable on announcements days differ from other trading days. However, the values are especially different immediately after market openings. During the first hour of day-time trading sessions (9:30:00-10:30:00) in the CBOT corn futures market on USDA announcement days, the proportion of the BAS that is attributed to the adverse selection is much higher than on no-announcement days. Especially in the first 30 minutes after market openings, the adverse selection component is around 25% on announcement days while this number is only around 15% for all trading days in my sample, which is in agreement with many previous studies (e.g. Krinsky and Lee 1996, Chung, Elder, and Kim 2013). In the second time interval (10:00:01-10:30:00), the adverse selection component is slightly below 10% on announcement days and less than 5% for no-announcement days. Both of these two numbers sharply decline from the first time interval (9:30:01-10:00:00); the average adverse selection cost is still higher on announcement days but the difference between those two is significantly smaller than in the first time interval. Another interesting fact is that, on no-announcement days, the adverse selection component reaches its lowest intraday level at the second time interval. However, the lowest adverse selection cost proportion occurs in the third time interval (10:30:01-11:00:00) during announcement days, which is later than other trading days. This evidence indicates that after the USDA releases their reports on announcement days, it takes longer for the futures market to reflect all new information than on no-announcement trading days. The adverse selection cost for both types of trading days are nearly identical at the fourth time interval (11:30:01-12:00:00). In another context, the USDA announcements affect the corn futures market trading activities for around 1.5 hours after market openings. Chung, Elder, and Kim (2013) find that the monetary policy announcement effects are short lived, lasting about 1.5 hours, in the U.S. stock markets. I find similar results in the corn futures market.

Before market closings, the asymmetric information cost climbs again in the last 45 minutes. The adverse selection cost difference between announcement and no-announcement days is around 2.5% in the second to last time interval, but the difference is nearly zero in the last time interval. It slightly contradicts the evidence from the volatility plot. In the last time interval, the volatility are always higher than the second last time interval. In addition, the volatility on USDA announcement days is higher than other trading days. Both transaction numbers and trading volume is significantly higher on announcement days, and the high BAS in the last 75 minutes of announcement day trading sessions suggests that the demand for liquidity is higher than the supply. Therefore, the cost of demanding liquidity is higher but the asymmetric information component is not necessarily larger.

The other important BAS component, the inventory cost, also seems to have an inverted U-shaped intraday pattern from 2008 to 2011. Krinsky and Lee (1996) also conclude that the inventory cost is approximately 13% lower during event period than their benchmark period for the NYSE stock markets. On announcement days, the inventory cost is much lower than all other days in the first time interval. From 10:30AM to 12:30PM, the announcement days inventory cost are nearly identical with no-announcement days. After 12:30PM on announcement trading days, the inventory cost component starts declining earlier than on no-announcement days. At the same time, the BAS, volatility, transaction numbers, and volume break their mid-day stable patterns and begin to increase earlier than on no-announcement days. This suggests that traders start liquidating their positions earlier on USDA announcement days than other days because positions built during the market opening periods. As traders liquidate their inventories earlier, liquidity providers' inventory levels also decline. Therefore, liquidity providers' inventory cost decreases earlier during day-trading hours on announcement days.

Based on my empirical results, the BAS and adverse selection cost are high after market openings, and they become even higher on USDA announcement days than no-announcement days. This finding corresponds to Wang, Garcia, and Irwin's (2013) inference that the adverse selection drives high BAS on USDA announcement days. Lehecka, Wang, and Garcia (2014) argue that after the release of USDA reports, the rapid transmission of information into the corn futures market takes place in ten minutes. Strong market reactions to the USDA announcements are represented by high return variability, volume, and volatilities after market opening. They also show that the excess trading volume continues for around an hour. Most importantly, the excess return variability lasts only about ten minutes in market openings on USDA announcement days. Findings here partially match the results from Lehecka, Wang, and Garcia (2014). On the basis of the intraday analysis, both studies show strong market reactions to the USDA announcements. However, the asymmetric information proportions are in average higher on announcement days than other days during the first trading hour. Although it is true that once a report is released, it becomes common public information, the high asymmetric information cost in post-announcement periods do not necessarily reflect the presence of private information holders. Instead, the increment of adverse selection cost on USDA announcement days is due to the presence of information processors who have superior ability to assess new asset fundamental values on the basis of the public announcements (Kim and Verrecchia 1994).

7. DISCUSSION AND CONCLUSIONS

7.1. Summary and Review

As the first research that studies the BAS components in the CBOT corn futures market, this thesis provides insights on the proportions of different BAS components, the adverse selection, inventory, and order processing cost components. The BBO dataset from the CBOT corn futures market enables me to decompose the market BAS with the Huang and Stoll (1997) model. The data is from January 2008 to October 2011, and there are 945 trading days in my sample. After omitting days that have low electronic trading volume, the sample size is 763 days.

Various research has shown order splitting is a common phenomenon in today's futures markets. Before applying the H-S model to decompose the corn futures market BAS, I revise the standard that Huang and Stoll (1997) used to bunch consecutive sequential transactions. If consecutive transactions occur with the same transaction price, trade direction, the same bid or ask, and within the same second, I identify them as split orders and bunch then into one transaction. This process is data bunching. I use the GMM procedure to estimate the H-S model for each selected trading day and time interval (e.g. Huang and Stoll 1997, Madhavan, Richardson, and Roomans 1997, Ahn et al. 2002, etc). All the components are expressed as proportions of the time-varying effective BAS. I show the daily adverse selection and inventory costs of the CBOT corn futures market. This thesis also studies on the monthly pattern of the market adverse selection cost, although I do not find an apparent monthly pattern. Additionally, I estimate the USDA announcement days' adverse selection and inventory cost components and compare them with their corresponding weekly and monthly averages. I do not detect any significant difference between USDA announcement daily adverse selection cost and the weekly or monthly costs. This result also applies to the daily inventory cost component. However, the intraday analysis shows that the market conditions are very different on USDA announcement days than other trading days, market are especially more volatile and liquid in the first hour of announcement days trading sessions.

7.2. Conclusion

In the absence of informed traders, the BAS contains only the inventory cost and order processing cost. If liquidity providers know a proportion of traders in the market are informed, they must widen the BAS to offset the losses when trading against informed traders. Under such situations, adverse selection cost is also an important component of the BAS. In this thesis, I find that from January 2008 to October 2011, the average proportion of the BAS that is attributable to the adverse selection is around 12.1%. Some previous studies on the stock markets find similar results that the adverse selection cost is only a small proportion of the BAS (e.g. Huang and Stoll 1997; McInish and Wood 1992). However, my result contradicts the study by MRR (1997) who conclude that adverse selection cost is the predominant source of the BAS. The predominant BAS component in my study is the inventory cost. The average proportion of the inventory cost is around 53.1% of the BAS in the CBOT corn futures market. This part of my findings differ from with Huang and Stoll (1997), who find the order processing cost is dominant. In this thesis, the market order processing cost proportion is stably around 35% of the BAS, which is the second largest component of the market liquidity cost.

During times that corn prices are high and volatile (2008 and 2011), the adverse selection cost proportion were larger than they were in 2009 and 2010. The respective adverse selection component from 2008 to 2011 are 17.63%, 9.03%, 7.33%, and 15.45%. When corn prices are more volatile, not only can higher proportion of the BAS be explained by the adverse selection cost, but also the magnitude of the BAS in those periods are high. With higher BAS and higher adverse selection cost proportion, the adverse selection cost becomes larger. The monthly adverse selection components in 2008 and 2011 seem to be more volatile than the other two years. During years that corn prices are stable (2009 and 2010), the monthly adverse selection seem to be stable as well.

The intraday analysis shows that the adverse selection cost, price volatility, BAS, number of transactions, and trading volume appear to have strong U-shaped patterns throughout each trading day, while the intraday inventory cost component has an inverted U-shaped pattern. The U-shaped pattern of the intraday adverse selection cost component indicates information flows into the market more intensely during market opening and closing hours than mid-day trading hours. This result in line with some previous research like Ryu (2011) and Ahn et al. (2002). The strong inverted U-shaped inventory cost suggests that liquidity providers net positions are relatively neutral during the beginning and ending of each trading day. In addition, the market is most liquid in opening and closing hours, despite high volatility in those hours, the inventory cost is low when liquidity providers can offset their inventory positions easily under liquid market conditions. In contrast, transactions occur less frequently during mid-day trading hours. The market is relatively less liquid, and the uncertainty about when the next transaction will occur is also high; therefore, the inventory cost component becomes high.

On days when the USDA does not have official announcements, market conditions generally are volatile during the first 30 minutes of trading and become quieter afterwards. Volatility and liquidity cost tend to increase during the last 45 minutes before market closings. Although I do not find much difference of the daily adverse selection and inventory cost components between USDA announcement days and no-announcement days, the intraday analysis reveals that the market conditions between these two types of days are exceedingly different during the first trading hour. On USDA announcement days, the adverse selection cost is around 10% higher and the inventory cost is around 10% lower than no-announcement trading days during the first 30 minutes. My results correspond to Krinsky and Lee (1996) who find that inventory cost is lower but adverse selection cost is higher around announcement event periods. From 10:30AM to 12:30PM, market conditions do not seem to deviate between no-announcement trading days and USDA announcement days as much as the opening and closing hours. During those 2 mid-day trading hours, the magnitudes of all the six liquidity measurements are nearly equal to each other for each time interval regardless of the USDA announcements. My empirical result shows that the market may take longer to settle down on USDA announcement days than what Lehecka, Wang, and Garcia (2014) conclude. They demonstrate that market excess volatility and returns only last around ten minutes on USDA announcement days. However, both their study and my study show that the CBOT corn futures market conditions are obviously different between announcement and no-announcement days. More precisely, the market is more volatile during opening and closing hours on USDA announcement days than no-announcement days.

7.3. Suggestions for Future Work

Despite the value of this study, there are several possible topics that can be studied in the future to enrich the literature in this research area. In the process of estimating the H-S model, I find that the H-S model generates α and/or β values that are below 0 or greater than 1. When the trade indicators covariance is exceedingly low, this issue tends to happen more often. However, vast majority of coefficient estimations falls into the interval of [0,1]. In the future, it will be useful to develop a more precise model to cope with the negative coefficients issue.

Since January 2013, the USDA started releasing major reports at 12:00PM EST. In the time period of my study, the reports were released at 8:30AM EST, which was 2 hours before market openings. It will be intriguing to investigate the immediate effect of the USDA reports on agricultural commodity markets since the releasing time was postponed. Moreover, this article does not study any correlation between the nearby and the deferred contracts. Mallory, Garcia, and Serra (2015) have found evidence that quote adjustments between the nearby and first deferred contracts are completed within a second. Decomposing the BAS and comparing its components between the nearby and deferred contacts can be an important milestone to answer how information arrivals shock nearby and deferred contracts similarly and/or differently.

REFERENCES

Amihud, Yakov, and Haim Mendelson. 1980. "Dealership market: Market-Making with Inventory." *Journal of Financial Economics* 8 (1). Elsevier: 31–53.

Ahn, H., Cai, J., Hamao, Y., & Ho, R. 2002. "The components of the bid–ask spread in a limit-order market: Evidence from the Tokyo Stock Exchange." *Journal of Empirical Finance*, 9(4), 399–430.

Bryant, H., & Haigh, M. (2004). Bid–ask spreads in commodity futures markets. *Applied Financial Economics*, 923-936.

Chelley-Steeley, Patricia, and Keebong Park. 2010. "The Adverse Selection Component of Exchange Traded Funds." *International Review of Financial Analysis* 19 (1). Elsevier: 65–76.

Collin-Dufresne, P. I. E. R. R. E., & Fos, V. 2015. Do prices reveal the presence of informed trading? *The Journal of Finance*.

Copeland, Thomas E, and Dan Galai. 1983. "Information Effects on the Bid-Ask Spread." the *Journal of Finance* 38 (5). Wiley Online Library: 1457–69.

Chung, K. H., & Kim, Y. (2009). Volatility, Market Structure, and the Bid-Ask Spread*. *Asia-Pacific Journal of Financial Studies, 38*(1), 67-107.

Demsetz, Harold. 1968. "The Cost of Transacting." *The Quarterly Journal of Economics*. JSTOR, 33–53.

Easley, David, Nicholas M Kiefer, Maureen O'hara, and Joseph B Paperman. 1996. "Liquidity, Information, and Infrequently Traded Stocks." *The Journal of Finance* 51 (4). Wiley Online Library: 1405–36.

Frank, Julieta, and Philip Garcia. 2010. "Bid-Ask Spreads, Volume, and Volatility: Evidence from Livestock Markets." *American Journal of Agricultural Economics*. Oxford University Press, aaq116.

Fridman, Moshe, and Lawrence Harris. 1998. “A Maximum Likelihood Approach for Non-Gaussian Stochastic Volatility Models.” *Journal of Business & Economic Statistics* 16 (3). Taylor & Francis Group: 284–91.

Garman, Mark B. 1976. “Market Microstructure.” *Journal of Financial Economics* 3 (3). Elsevier: 257–75.

Glosten, Lawrence R. 1987. “Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices.” *The Journal of Finance* 42 (5). Wiley Online Library: 1293–1307.

Glosten, Lawrence R, and Lawrence E Harris. 1988. “Estimating the Components of the Bid/Ask Spread.” *Journal of Financial Economics* 21 (1). Elsevier: 123–42.

Glosten, Lawrence R, and Paul R Milgrom. 1985. “Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders.” *Journal of Financial Economics* 14 (1). Elsevier: 71–100.

Grossman, Sanford J. 1986. “An Analysis of the Role of Insider Trading’ on Futures Markets.” *Journal of Business*. JSTOR, S129–46.

Hansen, Lars Peter. 1982. “Large Sample Properties of Generalized Method of Moments Estimators.” *Econometrica: Journal of the Econometric Society*. JSTOR, 1029–54.

Henker, T., & Wang, J. X. (2006). On the importance of timing specifications in market microstructure research. *Journal of Financial Markets*, *9*(2), 162-179.

Ho, Thomas, and Hans R Stoll. 1981. “Optimal Dealer Pricing Under Transactions and Return Uncertainty.” *Journal of Financial Economics* 9 (1). Elsevier: 47–73.

Huang, Roger D, and Hans R Stoll. 1994. “Market Microstructure and Stock Return Predictions.” *Review of Financial Studies* 7 (1). Soc Financial Studies: 179–213.

———. 1997. “The Components of the Bid-Ask Spread: A General Approach.” *Review of Financial Studies* 10 (4). Soc Financial Studies: 995–1034.

Irwin, Scott H., and Dwight R. Sanders. 2012. "Financialization and structural change in commodity futures markets." *Journal of agricultural and applied economics* 44（03): 371-396.

Kaniel, R., & Liu, H. 2006. So What Orders Do Informed Traders Use?*. *The Journal of Business J BUS*, 79(4), 1867-1913.

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

Krinsky, Itzhak, and Jason Lee. 1996. Earnings announcements and the components of the bid-ask spread. *The Journal of Finance*, 51(4), 1523-1535.

Kyle, Albert S. 1985. “Continuous Auctions and Insider Trading.” *Econometrica: Journal of the Econometric Society*. JSTOR, 1315–35.

Lee, Charles, and Mark J Ready. 1991. “Inferring Trade Direction from Intraday Data.” *The Journal of Finance* 46 (2). Wiley Online Library: 733–46.

Lehecka, Georg V, Xiaoyang Wang, and Philip Garcia. 2014. “Gone in Ten Minutes: Intraday Evidence of Announcement Effects in the Electronic Corn Futures Market.” *Applied Economic Perspectives and Policy* 36 (3). Oxford University Press: 504–26.

Lin, Ji-Chai. 1993. “Order Persistence, Adverse Selection, and Gross Profits Earned by NYSE Specialists.” In *Journal of Finance*, 48:1108–9. 3. AMER FINANCE ASSN 44 WEST FOURTH ST, STE 9-190, NEW YORK, NY 10012.

Lin, Ji-Chai, Gary C Sanger, and G Geoffrey Booth. 1995. “Trade Size and Components of the Bid-Ask Spread.” *Review of Financial Studies* 8 (4). Soc Financial Studies: 1153–83.

Madhavan, Ananth, Matthew Richardson, and Mark Roomans. 1997. “Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks.” *Review of Financial Studies* 10 (4). Soc Financial Studies: 1035–64.

Mallory, M., P. Garcia, and T. Serra. 2015. “Nearby and Deferred Quotes: What They Tell Us about Linkages and Adjustments to Information.” *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]

McInish, Thomas H, and Robert A Wood. 1992. “An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks.” *Journal of Finance*. JSTOR, 753–64.

O’Hara, Maureen. 1995. Market Microstructure Theory. Vol. 108. Blackwell Cambridge, MA.

Petersen, Mitchell A, and David Fialkowski. 1994. “Posted Versus Effective Spreads: Good Prices or Bad Quotes?” *Journal of Financial Economics* 35 (3). Elsevier: 269–92.

Roll, Richard. 1984. “A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market.” *The Journal of Finance* 39 (4). Wiley Online Library: 1127–39.

Ryu, Doojin. 2011, “Intraday price formation and bid–ask spread components: A new approach using a cross-market model.” *Journal of Futures Markets*, 31(12), 1142-1169.

———. 2012. “The Effectiveness of the Order-Splitting Strategy: an Analysis of Unique Data.” *Applied Economics Letters* 19 (6). Taylor & Francis: 541–49.

Shah, Samarth, and B Wade Brorsen. 2010. “Electronic Vs. Open Outcry: Side-by-Side Trading of KCBT Wheat Futures.” *Open Outcry: Side-by-Side Trading of KCBT Wheat Futures* (November 23, 2010).

Smith, Aaron. 2005. “Partially Overlapping Time Series: a New Model for Volatility Dynamics in Commodity Futures.” *Journal of Applied Econometrics* 20 (3). Wiley Online Library: 405–22.

Serednyakov, Alexey Sergeevich. 2006. *A model of the components of the bid-ask spread*. ProQuest.

Stoll, Hans R. 1976. “Dealer Inventory Behavior: an Empirical Investigation of NASDAQ Stocks.” *Journal of Financial and Quantitative Analysis* 11 (03). Cambridge Univ Press: 359–80.

———. 1989. “Inferring the Components of the Bid-Ask Spread: theory and Empirical Tests.” *The Journal of Finance* 44 (1). Wiley Online Library: 115–34.

Tannous, George, Juan Wang, and Craig Wilson. 2013. “The Intraday Pattern of Information Asymmetry, Spread, and Depth: Evidence from the NYSE.” *International Review of Finance* 13 (2). Wiley Online Library: 215–40.

Van Ness, Bonnie F, Robert A Van Ness, and Richard S Warr. 2001. “How Well Do Adverse Selection Components Measure Adverse Selection?” *Financial Management*. JSTOR, 77–98.

Wang, George HK, Jot Yau, and others. 2000. “Trading Volume, Bid-Ask Spread, and Price Volatility in Futures Markets.” *Journal of Futures Markets* 20 (10): 943–70.

Wang, Xiaoyang, Philip Garcia, and Scott H. Irwin. "Intraday Bid Ask Spread Variation in the Electronically Traded Corn Futures Market." *In 2012 Annual Meeting, August 12-14, 2012, Seattle, Washington*, no. 124899. Agricultural and Applied Economics Association, 2012.

———. 2013. “The Behavior of Bid-Ask Spreads in the Electronically-Traded Corn Futures Market.” *American Journal of Agricultural Economics*. Oxford University Press, aat096.

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

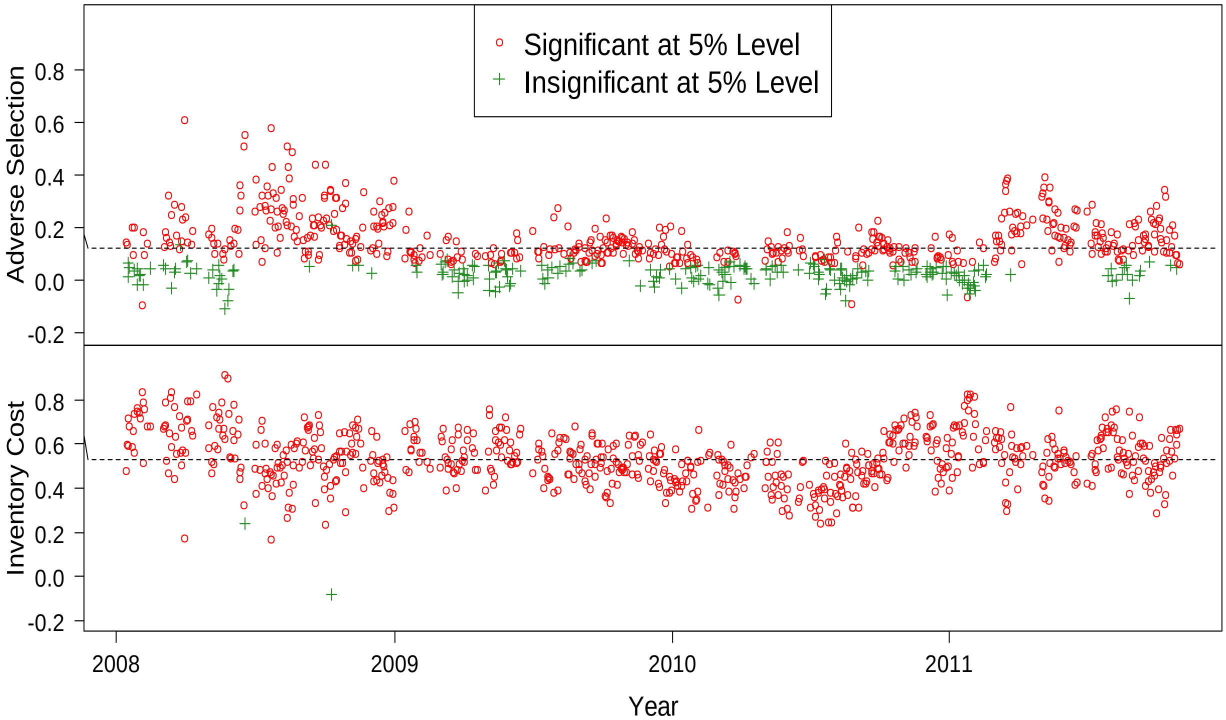
———. 1970. "Economic functions of futures markets." *Futures Trading in Livestock–Origins and Concepts, Chicago: Chicago Mercantile Exchange*, 267–97

FIGURES

Figure 1: Daily Nearby Corn Futures Settlement Price

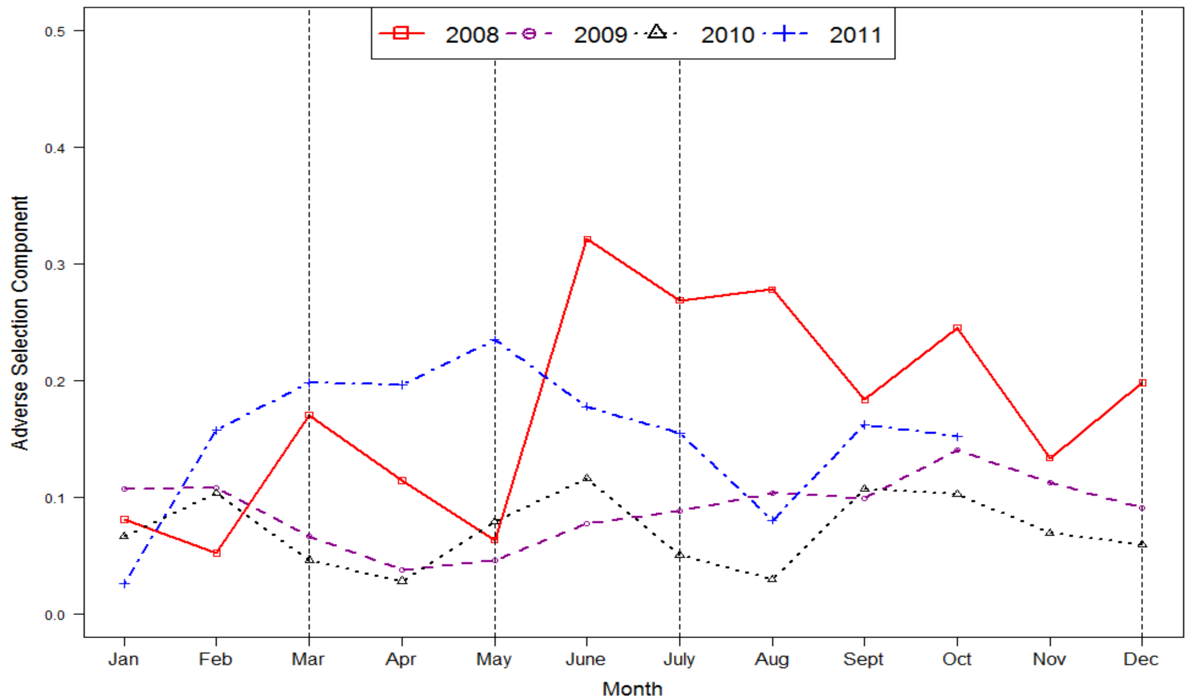
NOTES: Daily nearby corn futures settlement prices are obtained from the CRB database. The nearby contract is rolled over to the first deferred on the first trading day of nearby expiration moth. The time horizon is from January 14, 2008 to October 31, 2011.

Figure 2: Adverse Selection and Inventory Cost Components



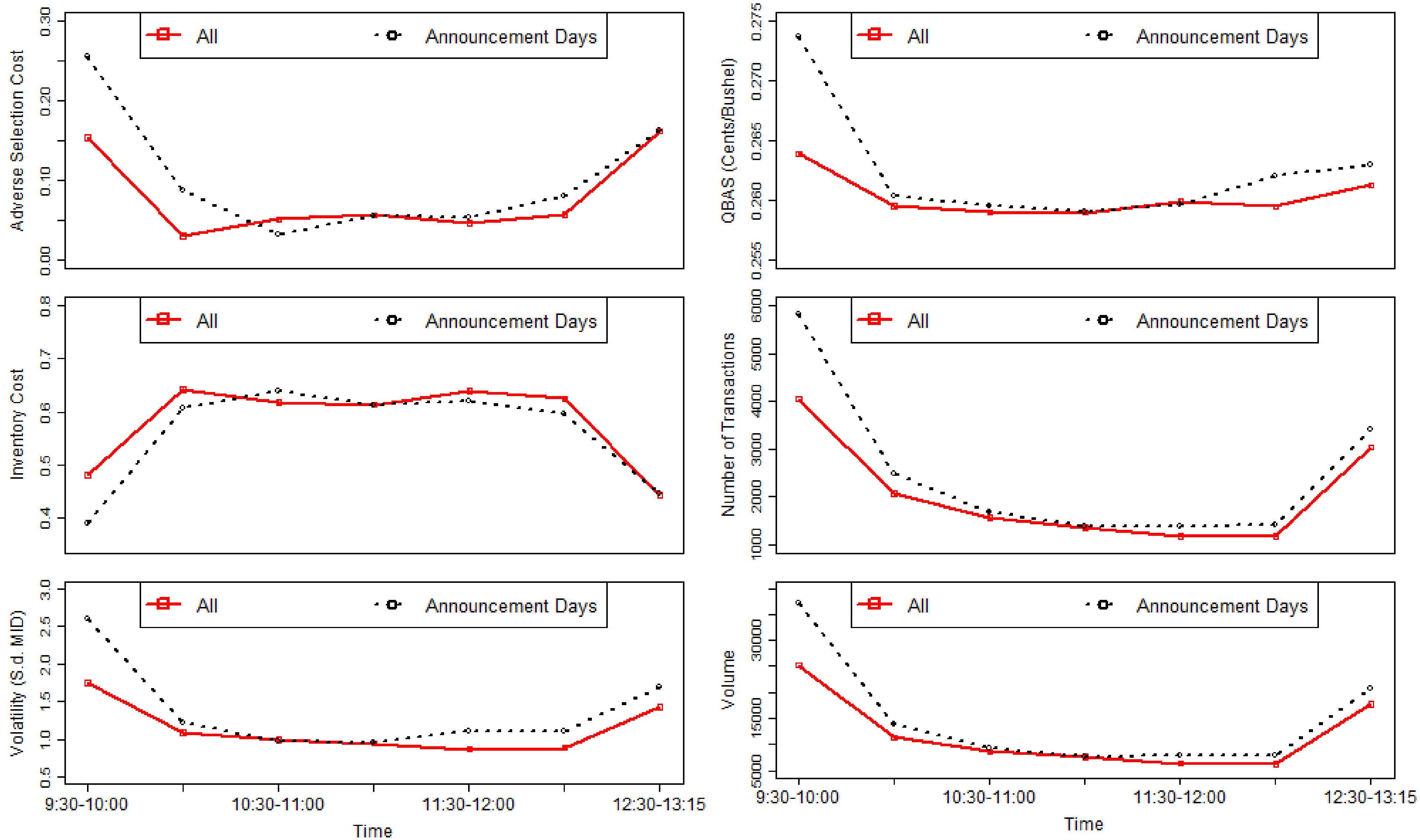
NOTES: Figure 2 are daily estimates for the adverse selection cost component (α) and inventory cost component (β). The dash line on each panel represents the mean value of the α and β, where the mean for α is 0.121 and 0.531 is the mean of all β's. Each circle/cross represents a day, and I use crosses and circles to distinguish from significant days (circles) from non-significant ones (crosses) at the 5% significance level.

Figure 3: Monthly Average Adverse Selection Component



NOTES: Figure 3 demonstrates the monthly average adverse selection component of the BAS. The α coefficient for each month is the mean of all daily adverse selection components. The vertical dash lines indicate the time that I roll over from each nearby contract to the 1st deferred contract. Since the data ends at Oct 31, 2011, there are no November and December estimations for 2011.

Figure 4: Intraday Analysis on USDA Announcement and No-announcement Days



NOTES: Figure 4 is the result of intraday analysis, and the empirical results are shown in Table 8. The x-axis represents 7 time intervals, which are 9:30:00-10:00:00, 10:00:01-10:30:00, 10:30:01-11:00:00, 11:00:01-11:30:00, 11:30:01-12:00:00, 12:00:0-12:30:00, and 12:30:01-13:15:00. In each plot, the dash line represents estimation of a variable on USDA announcement days, the solid line is a no-announcement day estimation. The 3 plots on the left (from top to bottom) are intraday adverse selection component (α), liquidity providers' inventory cost component (β), and the market volatility. The volatility is the standard deviation of all quote midpoints. The 3 plots on the right (from top to bottom) are the BAS, number of transactions (without bunching data), and the trading volume.

TABLES

Table 1: First Ten Observations from the Data Entry

1-A: Top Entry for Quote Revisions only

DATE & TIME	EX	SYMBOL	OFRSIZ	OFR	BIDSIZ	BID
1/14/2008 09:30:00	110460	803	7	510.75	238	511.25
1/14/2008 09:30:00	110480	803	7	510.75	204	511.25
1/14/2008 09:30:00	110500	803	240	511.5	204	511.25
1/14/2008 09:30:00	110510	803	239	511.5	204	511.25
1/14/2008 09:30:00	110540	803	238	511.5	204	511.25
1/14/2008 09:30:00	110550	803	237	511.5	203	511.25
1/14/2008 09:30:00	110570	803	237	511.5	202	511.25
1/14/2008 09:30:00	110580	803	237	511.5	212	511.25
1/14/2008 09:30:00	110590	803	237	511.5	112	511.25
1/14/2008 09:30:00	110600	803	237	511.5	62	511.25
1/14/2008 09:30:00	110620	803	235	511.5	62	511.25
1/14/2008 09:30:00	110630	803	199	511.5	62	511.25
1/14/2008 09:30:00	110640	803	204	511.5	62	511.25
1/14/2008 09:30:00	110660	803	204	511.5	58	511.25
1/14/2008 09:30:00	110680	803	202	511.5	58	511.25
1/14/2008 09:30:00	110700	803	202	511.5	57	511.25
1/14/2008 09:30:00	110720	803	201	511.5	57	511.25

1-B: Top Entry of Transactions only

DATE & TIME	EX	SYMBOL	SIZE	PRICE
1/14/2008 09:30:00	110470	803	34	511.25
1/14/2008 09:30:00	110490	803	1	511.5
1/14/2008 09:30:00	110520	803	1	511.25
1/14/2008 09:30:00	110560	803	1	511.25
1/14/2008 09:30:00	110610	803	38	511.5
1/14/2008 09:30:00	110650	803	4	511.25
1/14/2008 09:30:00	110670	803	2	511.5
1/14/2008 09:30:00	110690	803	1	511.25
1/14/2008 09:30:00	110710	803	1	511.5
1/14/2008 09:30:00	110730	803	20	511.25

1-C: Quotes and Trades Combined

DATE & TIME	EX	PRICE	SIZE	OFR	OFRSIZ	BID	BIDSIZ
1/14/2008 09:30:00	110470	511.25	34	510.75	7	511.25	238
1/14/2008 09:30:00	110490	511.5	1	510.75	7	511.25	204
1/14/2008 09:30:00	110520	511.25	1	511.5	239	511.25	204
1/14/2008 09:30:00	110560	511.25	1	511.5	237	511.25	203
1/14/2008 09:30:00	110610	511.5	38	511.5	237	511.25	62
1/14/2008 09:30:00	110650	511.25	4	511.5	204	511.25	62
1/14/2008 09:30:00	110670	511.5	2	511.5	204	511.25	58
1/14/2008 09:30:00	110690	511.25	1	511.5	202	511.25	58
1/14/2008 09:30:00	110710	511.5	1	511.5	202	511.25	57
1/14/2008 09:30:00	110730	511.25	20	511.5	201	511.25	57

Table 2: BBO Dataset Volume with CRB Volume

(BBO Dataset Volume)/ (CRB Volume)	Number of Trading Days Left	$\alpha > 1$	$\alpha < 0$	Significant at 10%
20% or above	920	3	73	599
30% or above	889	1	67	589
40% or above	848	1	63	568
50% or above	763	1	54	522
60% or above	599	1	40	428
70% or above	311	0	20	230
80% or above	37	0	0	23

NOTES: There are 945 trading days from January 14, 2008 to October 31, 2011. The first column in this table are different data filtering standards. For example, when I omit all trading days in which the daily trading volume from the BBO dataset is less than 20% of the trading volume from the CRB database, 25 days are omitted. So 920 days out of 945 days are left (in column 2, row 1). Out of all the 920 trading days, three of the daily adverse selection estimations are above 1 and 73 of them are below 0. Among these 920 daily adverse selection estimations, 599 of them are significant at 10% level. When changing the filtering standard to be 30%, 40%, and etc., the results are presented in similar ways. In this thesis, I choose the 50% standard.

Table 3:Bunch Sequential Transactions Example

3-A: Original Transaction and Quotes Data

DATE & TIME	EX	PRICE	SIZE	OFR	OFRSIZ	BID	BIDSIZ
1/14/2008 10:30:11	167570	514	10	514.25	25	514	31
1/14/2008 10:30:11	167600	514	6	514.25	20	514	21
1/14/2008 10:30:11	167620	514	1	514.25	20	514	15
1/14/2008 10:30:12	167640	514	2	514.25	19	514	14
1/14/2008 10:30:12	167660	514	10	514.25	19	514	12
1/14/2008 10:30:12	167680	514	2	514.25	19	514	2
1/14/2008 10:30:19	167750	513.75	10	514	7	513.75	17
1/14/2008 10:30:19	167770	513.75	7	514	7	513.75	7
1/14/2008 10:30:21	167820	513.5	5	513.75	12	513.5	30
1/14/2008 10:30:22	167850	513.5	25	513.75	14	513.5	25
1/14/2008 10:30:22	167880	513.25	2	513.5	2	513.25	17
1/14/2008 10:30:25	167920	513.25	10	513.5	6	513.25	15
1/14/2008 10:30:25	167940	513.5	2	513.5	6	513.25	5
1/14/2008 10:30:27	168000	513.25	2	513.5	7	513.25	15
1/14/2008 10:30:28	168030	513.25	10	513.5	7	513.25	15
1/14/2008 10:30:29	168050	513.25	5	513.5	7	513.25	5

3-B: After Bunching Sequential Transactions

DATE & TIME	EX	PRICE	SIZE	OFR	BID
1/14/2008 10:30:12	167680	514	31	514.25	514
1/14/2008 10:30:19	167770	513.75	17	514	513.75
1/14/2008 10:30:22	167850	513.5	30	513.75	513.5
1/14/2008 10:30:25	167920	513.25	12	513.5	513.25
1/14/2008 10:30:25	167940	513.5	2	513.5	513.25
1/14/2008 10:30:29	168050	513.25	17	513.5	513.25

Table 4: Summary Statistics

	BAS	Volume	Volatility	Average Transaction Size (not bunched data)	Average Transaction Size (bunched data)	Daily Transaction Counts (not bunched data)	Daily Transaction Counts (bunched data)
Mean							
2008	0.271	75389	3.925	5.598	12.385	13735	6343
2009	0.256	69280	2.443	5.595	15.835	12209	4368
2010	0.254	85991	2.342	6.475	18.794	13215	4687
2011	0.261	102385	4.040	5.726	13.875	17643	7392
Maximum							
2008	0.325	149940	12.173	12.554	30.692	26046	14313
2009	0.274	196443	7.419	7.278	20.947	26775	9876
2010	0.276	217849	8.382	8.938	29.027	31739	12638
2011	0.316	167019	11.853	7.150	21.935	27497	13349
Minimum							
2008	0.218	9897	0.472	4.009	6.603	643	263
2009	0.248	15519	0.614	4.238	9.938	3258	1521
2010	0.250	35619	0.460	4.679	10.874	6396	1885
2011	0.254	43956	1.146	4.424	9.106	6745	2763
Standard Deviation							
2008	0.011	21825	2.101	1.174	3.117	4332	2287
2009	0.003	21650	1.267	0.566	2.198	3398	1354
2010	0.004	32080	1.382	0.742	3.374	4986	2039
2011	0.007	23503	2.065	0.484	2.034	3541	1781

NOTES: The BAS here is the quoted spread, which is estimated directly from all transactions of each day. Volume is the total counts of transacted futures contracts of each daytime session. The volatility measurement is the standard deviation of the quotes midpoints of each day. The 4th column is the average transaction size (by number of contracts) before the applying the data bunching process. After bunching all sequential trades, the average size per transaction (by number of contracts) is shown in the 5th column. The 6th and 7th columns are number of non-zero-size transactions for both bunched and not bunched data.

Table 5: GMM Estimation Coefficients Significance Summary

Year	Number of Trading Days	0< α <1	10%	5%	1%	0< β <1	10%	5%	1%
2008	197	187	162	158	148	195	195	194	193
2009	188	176	135	127	103	188	188	188	188
2010	203	184	117	108	80	204	204	204	204
2011	175	162	140	138	129	175	175	175	175
Total	763	709	554	531	460	762	762	761	760

NOTES: I present the daily BAS component estimations in section 6. This table complements Figure 2. The second column in this table shows the sample size in each year after the filtering process by comparing the electronic volume and CRB trading volume. The third column is the number of adverse selection component estimations that are between 0 and 1. Columns 4 to 6 are number of α's that are significant at 10%, 5%, and 1% significance levels respectively. The last four columns are descriptive statistics for β estimations, and the structure of those columns are same as columns 3 to 6 of α's.

Table 6: Adverse Selection and Inventory Cost Component (Monthly Average)

	2008							2009					
	α	β	Cov-B	Cov-N	BAS	Volatility		α	β	Cov-B	Cov-N	BAS	Volatility
January	0.081	0.654	-0.502	0.344	0.266	2.883	January	0.108	0.609	-0.528	0.386	0.261	3.090
February	0.052	0.712	-0.516	0.356	0.263	3.007	February	0.108	0.530	-0.597	0.375	0.258	2.097
March	0.171	0.615	-0.376	0.326	0.276	3.698	March	0.067	0.535	-0.630	0.400	0.257	2.037
April	0.115	0.711	-0.444	0.338	0.266	3.516	April	0.038	0.568	-0.649	0.447	0.256	1.903
May	0.064	0.661	-0.412	0.353	0.276	3.475	May	0.046	0.587	-0.647	0.410	0.256	2.167
June	0.322	0.463	-0.288	0.337	0.287	4.842	June	0.078	0.561	-0.641	0.440	0.255	2.545
July	0.269	0.483	-0.356	0.350	0.276	5.076	July	0.088	0.502	-0.599	0.408	0.254	2.603
August	0.278	0.48	-0.388	0.357	0.271	4.515	August	0.104	0.544	-0.620	0.428	0.255	2.613
September	0.184	0.588	-0.405	0.347	0.269	3.952	September	0.100	0.523	-0.667	0.413	0.254	2.509
October	0.246	0.459	-0.405	0.356	0.269	3.886	October	0.141	0.483	-0.651	0.385	0.256	2.705
November	0.134	0.599	-0.471	0.391	0.263	3.511	November	0.113	0.549	-0.626	0.414	0.255	3.150
December	0.199	0.474	-0.500	0.359	0.264	3.718	December	0.092	0.481	-0.645	0.383	0.256	2.024
Average	*0.176*	*0.574*	*-0.421*	*0.351*	*0.270*	*3.839*	***Average***	*0.090*	*0.539*	*-0.625*	*0.407*	*0.256*	*2.453*

Table 6 (cont.)

	2010							2011					
	α	β	Cov-B	Cov-N	BAS	Volatility		α	β	Cov-B	Cov-N	BAS	Volatility
January	0.067	0.451	-0.741	0.440	0.253	1.746	January	0.027	0.659	-0.653	0.386	0.259	3.336
February	0.104	0.519	-0.613	0.374	0.258	3.195	February	0.158	0.547	-0.524	0.352	0.262	4.131
March	0.047	0.455	-0.708	0.465	0.253	1.596	March	0.199	0.548	-0.482	0.347	0.266	4.582
April	0.042	0.457	-0.719	0.412	0.253	2.085	April	0.197	0.538	-0.489	0.361	0.263	4.733
May	0.080	0.454	-0.73	0.389	0.252	1.902	May	0.235	0.519	-0.440	0.336	0.269	4.778
June	0.117	0.397	-0.696	0.404	0.252	2.046	June	0.178	0.506	-0.481	0.334	0.265	5.450
July	0.051	0.362	-0.648	0.355	0.252	1.978	July	0.155	0.574	-0.513	0.352	0.261	3.769
August	0.030	0.429	-0.632	0.349	0.254	2.430	August	0.081	0.576	-0.574	0.372	0.259	3.098
September	0.108	0.476	-0.759	0.388	0.254	2.839	September	0.163	0.505	-0.582	0.344	0.258	4.434
October	0.103	0.586	-0.688	0.387	0.256	3.535	October	0.152	0.544	-0.547	0.349	0.258	3.770
November	0.070	0.640	-0.619	0.418	0.260	3.834	***Average***	*0.154*	*0.551*	*-0.528*	*0.353*	*0.262*	*4.208*
December	0.060	0.559	-0.63	0.347	0.260	2.288							
Average	*0.073*	*0.482*	*-0.681*	*0.394*	*0.254*	*2.456*							

NOTES: Since some trading days have low recorded electronic trading volume, only 763 days (out of 945) are selected to generate in Table 6. All included days' recorded electronic trading volume to total volume ratio are greater than 0.5. There are 197, 188, 204, and 175 days in 2008, 2009, 2010, and 2011 respectively. α represents the adverse selection component; β is the inventory cost component. Cov-B and Cov-N are the trade indicator covariance for the bunched and not bunched dataset. BAS is the average of all quoted spreads. Volatility variable from the last column is the standard deviation of quote midpoints.

Table 7: USDA Announcement Days BAS Components

Date	α	α (P-Value)	β	β (P-Value)	Reports
3/11/2008	0.143	0	0.648	0	WASDE, CP
3/31/2008	0.606	0	0.175	0	GS, PP
4/9/2008	0.134	0	0.661	0	WASDE, CP
5/9/2008	0.1	0.008	0.590	0	WASDE, CP
6/10/2008	0.099	0.005	0.711	0	WASDE, CP
7/11/2008	0.32	0	0.409	0	WASDE, CP
8/12/2008	0.507	0	0.263	0	WASDE, CP
9/12/2008	0.167	0	0.506	0	WASDE, CP
9/30/2008	0.32	0	0.395	0	GS
10/28/2008	0.29	0	0.475	0	CP (Revision)
11/10/2008	0.074	0.029	0.677	0	WASDE, CP
12/11/2008	0.138	0	0.522	0	WASDE, CP
3/11/2009	0.075	0.081	0.531	0	WASDE, CP
3/31/2009	0.08	0.003	0.527	0	GS, PP
5/12/2009	0.061	0.016	0.451	0	WASDE, CP
6/10/2009	0.177	0.002	0.559	0	WASDE, CP
7/10/2009	0.081	0.015	0.600	0	WASDE, CP
8/12/2009	0.098	0.013	0.626	0	WASDE, CP
9/11/2009	0.068	0.037	0.546	0	WASDE, CP
9/30/2009	0.095	0.002	0.512	0	GS
10/9/2009	0.104	0	0.574	0	WASDE, CP
11/10/2009	0.183	0	0.546	0	WASDE, CP
12/10/2009	0.116	0	0.524	0	WASDE, CP
3/10/2010	0.07	0.022	0.522	0	WASDE, CP
3/31/2010	0.058	0.055	0.442	0	GS, PP
7/9/2010	0.092	0.003	0.269	0	WASDE, CP
8/12/2010	0.043	0.17	0.391	0	WASDE, CP
9/10/2010	0.118	0	0.423	0	WASDE, CP
9/30/2010	0.167	0	0.491	0	GS
11/9/2010	0.103	0	0.594	0	WASDE, CP
12/10/2010	0.028	0.364	0.692	0	WASDE, CP
1/12/2011	0.068	0.003	0.547	0	WASDE, CP, GS
2/9/2011	0.141	0	0.510	0	WASDE, CP
3/10/2011	0.135	0	0.622	0	WASDE, CP
6/9/2011	0.108	0	0.520	0	WASDE, CP
7/12/2011	0.168	0	0.510	0	WASDE, CP
8/11/2011	0.078	0	0.528	0	WASDE, CP
9/12/2011	0.113	0	0.605	0	WASDE, CP

Table 7 (cont.)

Date	α	α (P-Value)	β	β (P-Value)	Reports
10/12/2011	0.317	0	0.371	0	WASDE, CP

NOTES: There are 59 USDA announcement days from January 2008 to October 2011. The reporting days in January 2008 and 2009 are omitted due to missing data. As it was mentioned before, the October 2011 USDA WASDE and CP reports announcement day is also dropped because the transaction price, the best bid price, and the trade indicator are identical for that entire day. March 31st, 2011 is excluded due to the same reason. There are some other announcement days that are dropped due to small recorded electronic trading volume over the CRB total volume, which are February 8, 2008, June 30, 2008, February 10, 2009, June 30, 2009, January 12, 2010, February 9, 2010, June 10 and June 30 of 2010, April 8, 2011, and June 30, 2011. In addition, announcement days like October 10, 2008, April 9, 2009, April 9 and May 11 of 2010, and May 11 and September 30 of 2011 are omitted due to limited price movement, and/or too few recorded observations. After filtering out, there are 39 announcement days in my sample of study. Under such circumstance, the H-S model is not able to evaluate the true values of α andβ. When calculating the weekly and monthly average of alphas, I averaged all estimated values regardless of their significance values.

Table 8: Intraday Analysis Output

Time Interval	α	β	1- α- β	Volatility	BAS	Number of Transactions	Trading Volume
09:30:00-10:00:00	0.153 (0.255)	0.481 (0.389)	0.366 (0.356)	1.748 (2.603)	0.263 (0.274)	4047 (5826)	25220 (37054)
10:00:01-10:30:00	0.030 (0.087)	0.641 (0.608)	0.329 (0.305)	1.087 (1.228)	0.259 (0.26)	2060 (2480)	11354 (14003)
10:30:01-11:00:00	0.051 (0.033)	0.618 (0.639)	0.331 (0.328)	0.99 (0.976)	0.258 (0.26)	1561 (1678)	8614 (9193)
11:00:01-11:30:00	0.056 (0.054)	0.614 (0.614)	0.33 (0.332)	0.947 (0.957)	0.258 (0.259)	1350 (1391)	7570 (7585)
11:30:01-12:00:00	0.046 (0.053)	0.639 (0.619)	0.315 (0.328)	0.869 (1.101)	0.259 (0.26)	1170 (1393)	6366 (8022)
12:00:01-12:30:00	0.056 (0.08)	0.624 (0.596)	0.32 (0.324)	0.889 (1.112)	0.259 (0.262)	1165 (1413)	6328 (7774)
12:30:01-13:15:00	0.161 (0.163)	0.443 (0.445)	0.396 (0.392)	1.425 (1.701)	0.261 (0.263)	3033 (3417)	17792 (20663)

NOTES: Figure 4 is plotted based on the output from this table. Except the order processing cost (1- α- β), six other market measures are plotted in Figure 4. Values in parentheses are estimations for announcement days. For example, the estimation of α in the first time interval for all trading days is 0.159, while the α estimation is 0.255 on USDA announcement days in the same time interval. The volatility, BAS, and number of transactions are calculated based on the non-bunched data.