THREE ESSAYS ON COMMODITY MARKETS

BY

ZHEPENG HU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Agricultural and Applied Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Doctoral Committee:

Professor Philip Garcia, Co-Chair
Professor Maria Teresa Serra Devesa, Co-Chair
Assistant Professor Todd Hubbs
Associate Professor Mindy Mallory, Purdue University

ABSTRACT

This dissertation consists of three essays that investigate issues in agricultural commodity futures and cash markets. The first essay uses price discovery measures and intraday data to quantify the proportional contribution of nearby and deferred contracts in price discovery in the corn and live cattle futures markets. On average, nearby contracts reflect information more quickly than deferred contracts in the corn market but have a relatively less dominant role in the live cattle market. In both markets, the nearby contract loses dominance when its relative volume share dips below 50%, which typically occurs when the nearby is close to maturity. Regression results indicate that the share of price discovery is mainly related to trading volume and time to expiration in both markets. In the corn market, the price discovery share between nearby and deferred contracts is also related to inverse carrying charges, crop year differences, USDA announcements, market crashes, and commodity index position rolls. Differences between corn and live cattle markets are consistent with differences in the contracts' liquidity and commodity storability.

The second essay investigates the effect of algorithmic trading activity, as measured by quoting, on the corn, soybean, and live cattle commodity futures market quality. Using the CME's limit-order-book data and a heteroskedasticity-based identification approach, we find more intensive algorithmic quoting (AQ) is beneficial in multiple dimensions of market quality. On average, AQ improves pricing efficiency and mitigates short-term volatility, but its effects on liquidity costs are somewhat mixed. Increased AQ significantly narrows effective spreads in the corn and soybean markets, but not in the less traded live cattle futures market. The narrowing in effective spreads emerges from a reduction in adverse selection costs as more informed traders

lose their market advantage. There also is evidence that liquidity provider revenues increase with heightened AQ activity in the corn futures market, albeit the effect is not statistically significant in the soybean and live cattle futures markets.

The third essay investigates how export prices and sales responses to exchange rate movements are affected by the level of the stocks-to-use ratio. The analysis is performed in the corn, soybean, and wheat export markets using Threshold Vector Autoregressive (TVAR) models and monthly data for the January 1990-December 2019 period. Both importer and exporter exchange rates are considered in our analysis. Results show that the effects of both importer and exporter exchange rates on corn export prices and sales are either insignificant or have small economic value due to the relatively small export share of production. In the more export-oriented soybean and wheat markets, an increase in the value of the dollar relative to other exporters' currencies causes an expected and significant decrease in the export price, but export sales are not significantly affected which reflects the low substitutability between the U.S. exports and competitors' exports in terms of marketing seasons and crop classes. The effects of importer exchange rates present significant threshold effects in soybean and wheat markets as export prices and sales are more responsive in the low regime of stocks-to-use ratio. Similar threshold effects are also found in the exporter exchange rate impacts on corn export prices and sales. However, the impacts across regimes are not largely different in economic value.

ACKNOWLEDGMENTS

I would like to gratefully thank my advisors, Dr. Philip Garcia and Dr. Serra Terresa, for their great support and guidance throughout my graduate life. I am also pleased to thank Dr. Mindy Mallory for providing thoughtful suggestions to my dissertation. I would also like to thank Dr. Todd Hubbs for providing his knowledge about agricultural commodity markets.

I also have received great support from other faculty members including Dr. Robe Michel, Dr. Scott Irwin and my peers in the world-class price analysis research group in the department of ACE. Last but by no means least, special thanks are granted to my parents. Without them, I would not have been able to make it this far.

This dissertation was finished in the middle of a global pandemic. I would like to thank all healthcare workers who are fighting on the frontlines.

TABLE OF CONTENTS

CHAPTER 1:

INTRODUCTION

In the past decade, agricultural commodity futures and cash markets have witnessed structural changes and market events that created important impacts. In agricultural futures markets, the transformation from open outcry trading to electronic trading has changed the nature of futures trading. The speed at which information enters the market has increased, but this has caused concerns that market quality and efficiency have been harmed. In agricultural cash markets, U.S. exports are facing increasing competition in global export markets. In addition, recent trading tensions between the U.S. and major trading patterners cause tremendous uncertainty in U.S. export markets. In this context, this dissertation examines three contemporary issues in agricultural commodity futures and cash markets, aiming to provide implications for understanding the nature of electronic trading and challenges in U.S. agricultural exports.

The first two essays focus on market microstructure in agricultural commodity futures markets. The first essay uses high frequency trade data and price discovery shares to study price discovery along the futures forward in storable and non-storable markets. The second essay investigates the effects of algorithmic trading on agricultural futures market quality. The last essay focuses on exchange rate effects on agricultural export prices and sales.

The first essay, "**Measuring Price Discovery between Nearby and Deferred Contracts in Storable and Non-Storable Commodity Futures Markets**," uses high frequency intraday transaction data and price discovery measures, including Putniņš' (2013) information leadership share, to measure the share of price discovery between nearby and deferred contracts in corn and live cattle markets. It has been widely shown in previous studies that the nearby contract provides most price discovery in agricultural futures markets. However, price discovery along the forward curve is a dynamic process. Using intraday futures transaction data, this essay studies for the first time when, and the speed at which price discovery switches from the nearby contract to the next nearby contract. The results are helpful to researchers who need to build time series of rolling nearby futures prices. In addition, regression analysis is used to identify the factors that relate to price discovery, considering for the first time Working's and Tomek's predictions about the location of price discovery, as well as the effects of commodity index trading and pit trading closure.

While electronic trading changes the nature of price discovery in agricultural commodity futures markets, the change in the speed of trading also influences agricultural futures markets in many other ways. While a growing number of studies on the microstructure of agricultural commodity futures markets have emerged in recent years (e.g., Wang, Garcia, and Irwin 2013; Couleau, Serra, and Garcia 2018, 2019), they do not provide direct identification of the effects of high frequency trading. The second essay, "**Algorithmic Quoting, Trading, and Market Quality in Agricultural Commodity Futures Markets**," provides the first empirical evidence for the effects of algorithmic quoting, on the corn, soybean, and live cattle commodity futures market quality. Using limit order book data and a heteroskedasticity based identification approach, this essay shows algorithmic quoting is beneficial to market quality in multiple dimensions. However, there is evidence that heightened algorithmic quoting is associated with higher liquidity provider revenues. These findings point to a tradeoff between the dimensions of market quality, and the need for continued monitoring of algorithmic trading activity in agricultural commodity futures markets.

While the effect of electronic algorithmic trading on commodity markets can't be denied, market fundamentals continue to be relevant for understanding agricultural commodity markets. Exchange rate effects of agricultural commodity prices and exports have long been studied since Schuh's (1974) classic article. However, only a few studies have shown how the underlying market supply-demand conditions affect the exchange rate-exports relationship. The third essay, "**Exchange Rate Effects on Agricultural Export Prices and Sales in High-Low Stock Regimes**," studies export prices and sales responses to exchange rate movements in different stocks-to-use conditions in the corn, soybean, and wheat export markets. The results provide important implications for both policymakers and market participants that stocks-to-use conditions need to be considered for accurate evaluations and forecasts on exchange rate effects in agricultural markets.

CHAPTER 2:

MEASURING PRICE DISCOVERY BETWEEN NEARBY AND DEFERRED CONTRACTS IN STORABLE AND NON-STORABLE COMMODITY FUTURES MARKETS

2.1 Introduction

Price discovery is a main function of futures markets. Traditional research on price discovery in agricultural futures markets has developed in three main areas: determining which dominates, cash or futures price (Garbade and Silber, 1983; Schroeder and Goodwin, 1991; Ahumada and Cornejo, 2016); which of several geographically differentiated markets dominates (Koontz, Garcia and Hudson, 1990; Janzen and Adjemian, 2017; Arnade and Hoffman, 2018); and whether there is a difference in the quality of price discovery in storable versus non-storable commodities (Leuthold, Junkus and Cordier, 1989; Yang, Bessler and Leatham, 2001). Overwhelming evidence suggests that futures markets lead cash markets in price discovery.

The recent introduction of electronic trading in futures markets has heightened their liquidity and increased the speed of response to new information. Relative to cash markets, which typically report prices daily, this has strengthened the leadership of futures markets in price discovery. But futures markets are not completely homogeneous; instead a market contains contracts with different maturities to meet trader needs that differ in time. Little is known about where along the futures forward contract curve new information gets impounded into prices. Understanding how each contract contributes to price discovery is essential for market participants making sound hedging and trading decisions.

Working (1948, 1949) developed a theory explaining price relationships along the futures forward curve for storable commodities. Prices for a storable commodity are linked through time by the net costs of carrying inventories. Exceptions may occur in periods of inverse carrying charges as low inventories break down the normal storage linkage. Working's view of deferred futures prices implies that they may play a less dominant role in price discovery compared to the nearby futures prices as they only adjust to nearby prices based on changes in storage costs. However, Tomek and Gray (1970) and Tomek (1997) argue that commodity futures not only provide guidance for carrying inventories, but also forecasts of expected futures prices reflecting future supply and demand conditions. Hence, their view implies that price discovery along the forward curve is not always dominated by the nearby contract. In particular, when traders act on market news that affects their expectations about the supply and demand conditions in a deferred month, price discovery is more likely to occur first in a deferred contract.

While Working's and Tomek's theories offer different predictions on the price discovery role of the nearby contract, empirical studies find the nearby contract, on average, provides most price discovery in futures markets for agricultural commodities (Sanders, Garcia and Manfredo, 2008; Schnake, Karali and Dorfman, 2012), as well as other financial assets (Chen and Tsai, 2017). However, price discovery along the forward curve is dynamic. The forward curve shifts as days to expiration decrease for each contract and new contracts are added. The nearby contract loses importance as the delivery period approaches, evidenced by falling volume and open interest. However, no research has directly examined when and the speed at which price discovery switches from the nearby contract to the next. In addition, while long-run cointegration between commodity prices has been widely examined in previous studies on price discovery, the cointegration relationship between agricultural commodity futures prices for different time horizons within a day has never been studied.

In this paper we measure price discovery between nearby and deferred futures for each day from 2008 to 2015. We use Chicago Mercantile Exchange (CME) transactions data for corn and live cattle that are time-stamped to the second. We employ the information leadership share (Putniņš, 2013), which is designed for high frequency data samples and is robust to differences in noise in price series. This price discovery share (*PS*) measure enables us to determine the relative proportion of information impounded in nearby and deferred futures prices. The use of high frequency data allows us to measure price discovery daily and offer a day-to-day dynamic characterization of how futures price discovery switches from one contract to the next as the nearby nears expiration.

We first document patterns in daily *PS*s between nearby and deferred futures contracts. Findings reveal *PS*s are strongly related to the contracts' relative volume shares (*VS*s). The nearby contract dominates deferred contracts in price discovery when it has more trading volume, which typically occurs until several days before the nearby enters the delivery period. The nearby contract systematically loses dominance when its relative *VS* dips below 50%. Also, the nearby contract plays a more important role in price discovery in the corn than in the live cattle market. Using regression analysis, we investigate the factors that are related to *PS* between nearby and deferred futures. We find *PS* is strongly related to trading volume and days-to-expiration. In corn, *PS* is also related to inverse carrying charges, the nearby and deferred contracts representing different crop years, USDA reports, price declines, and commodity index rolls. Differences between corn and live cattle markets are consistent with differences in liquidity, storability and other market characteristics.

This paper contributes to the literature in two ways. First, this is the first paper that uses high frequency data to study daily price discovery dynamics in physical commodities with different degrees of storability. Previous studies typically use daily data and focus on price discovery in the long run (Covey and Bessler, 1995; Yang, Bessler and Leatham, 2001). We provide day-to-day dynamics on how fast price discovery switches from the nearby contract to the next. This is especially useful to empirical research that relies on futures prices and builds a time series of rolling nearby futures prices. Two rolling techniques are common: rolling on a fixed number of days prior to expiration and rolling when volume in the first deferred overtakes volume in the nearby. Our paper shows the latter is preferable from a price discovery perspective. Second, we identify the factors that relate to *PS*, considering for the first time Working's and Tomek's predictions about the location of price discovery, as well as the effects of commodity index trading and pit trading closure.

2.2 Price Discovery Measures

Garbade and Silber (1983) first developed a measure (the *GS* measure) to quantify price discovery. It is based on lead-lag relationships captured by the following model of price behavior:

$$\begin{bmatrix} p_{1,t} \\ p_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} 1-\beta_1 & \beta_1 \\ \beta_2 & 1-\beta_2 \end{bmatrix} \begin{bmatrix} p_{1,t-1} \\ p_{2,t-1} \end{bmatrix} + \begin{bmatrix} \omega_{1,t} \\ \omega_{2,t} \end{bmatrix} \tag{2.1}$$

where $p_{1,t}$ and $p_{2,t}$ are the prices for nearby and deferred futures contracts at time t, repectively. The coefficients β_1 and β_2 measure the effect of one-period lagged deferred futures price on the current nearby futures price and vice versa, respectively. The shares:

$$GS_1 = \frac{\beta_2}{\beta_1+\beta_2}\,,\ GS_2 = \frac{\beta_1}{\beta_1+\beta_2} \tag{2.2}$$

are used for measuring the proportional contribution of each contract to the price discovery process. However, the *GS* measure ignores the possibility that the two prices share a common stochastic trend that represents the common efficient price being discovered.

More recent price discovery measures are derived from structural models of the data generating process based on cointegration and error correction models. Hasbrouck (1995) information share (*IS*) and Harris, McInish and Wood (2002) component share (*CS*) are the most widely used. The fundamental value of a commodity at contract maturity (w) is unknown but discovered through a dynamic process. Let w_t be the fundamental value of price conditional on the information available at time *t*. w_t is assumed to follow a random walk:

$$w_t = w_{t-1} + v_t, \quad v_t \sim N(0, \sigma_v), \tag{2.3}$$

where v_t is *i.i.d*. Market participants incorporate information and expectations about fundamentals with a delay of δ_i periods into the observed futures price $p_{i,t}$, as they need time to interpret the information and take appropriate positions. As a result, $p_{i,t}$ is:

$$p_{i,t} = w_{t-\delta_i} + s_{i,t}, \;\; s_{i,t} \sim N\left(0, \sigma_{s_i}\right) \tag{2.4}$$

where i = 1 and 2 are nearby and deferred contracts and $s_{i,t}$ represents *i.i.d* noise. Thus, price deviations from the fundamental value are only transient which results in cointegration between prices for nearby and deferred contracts. Both *IS* and *CS* are derived by estimating a (bivariate) VECM:

$$\Delta \boldsymbol{p_t} = \boldsymbol{\alpha}(\boldsymbol{\beta}' \boldsymbol{p_t} - \mu) + \sum_{j=1}^{J} \boldsymbol{\Gamma_j} \Delta \boldsymbol{p_{t-j}} + \boldsymbol{e_t} \tag{2.5}$$

where $\boldsymbol{p_t} = (p_{1,t}, p_{2,t})'$ is a vector of nearby and deferred futures prices at time t. $\boldsymbol{\beta} \in \mathbb{R}^2$ is a cointegrating vector of parameters that allows for a constant term μ which reflects the difference between nearby and deferred prices. Since storage costs are typically quoted in cents per day in commercial settings, it is reasonable to assume that they are constant within the day and thus reflected in μ. The parameter vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)'$ contains error correction coefficients that measure the speed at which disruptions of the long-run price equilibrium are corrected. $\boldsymbol{\Gamma_j} \in \mathbb{R}^{2\times2}$ is a vector of autoregressive coefficients representing short-run dynamics and J is the number of lags in the model. The error term $\boldsymbol{e_t}$ is a zero-mean vector of white noise residuals with covariance matrix:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}. \tag{2.6}$$

Harris, McInish, and Wood (2002) show *CS* can be calculated from the normalized orthogonal to the vector of error correction coefficients, $\alpha_\perp = (\gamma_1, \gamma_2)'$. By noting that CS_1+CS_2=1,

$$CS_1 = \gamma_1 = \frac{\alpha_2}{\alpha_2 - \alpha_1}, CS_2 = \gamma_2 = \frac{\alpha_1}{\alpha_1 - \alpha_2} \tag{2.7}$$

are the *CS* measures for nearby and deferred contracts, respectively.

The *IS* measures for nearby (IS_1) and deferred (IS_2) contracts can be derived from the error correction coefficients and the variance-covariance matrix of the error terms as follows (Hasbrouck, 1995):

$$IS_1 = \frac{(\gamma_1 m_{11} + \gamma_2 m_{12})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2}, IS_2 = \frac{(\gamma_2 m_{22})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2}, \tag{2.8}$$

where γ_1 and γ_2 are the *CS* measures in equation (2.7), and m_{11}, m_{12}, and m_{22} are from the Cholesky factorization of the VECM residual covariance matrix, $\Sigma = MM'$, where

$$M = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2(1 - \rho^{1/2})^{1/2} \end{pmatrix} = \begin{pmatrix} m_{11} & 0 \\ m_{12} & m_{22} \end{pmatrix}. \tag{2.9}$$

The Cholesky factorization eliminates the contemporaneous relationship between price innovations (Hasbrouck, 1995). However, this procedure makes the *IS* results order dependent. Following Baillie et al. (2002) and others, we calculate *IS* by averaging the measures of the two price orderings.

Price discovery metrics are designed to reflect the leadership in the speed in impounding new information (Hasbrouck, 1995). However, Yan and Zivot (2010), and Putniņš (2013) show that *IS* and *CS* measure a combination of speed in impounding new information and noise due to trading frictions. Although contract specifications such as tick size and price limits are the same for nearby and deferred contracts, differences in noise levels in nearby and deferred prices can be large due to differences in trading frequency or high frequency trading activities (Wang, Garcia and Irwin, 2013; Couleau, Serra and Garcia, 2019). Price discovery incorporates information into the market through active trading and higher volume, which leads to increased price updating and more microstructure noise. In contrast, less trading activity is associated with less information entering the market, but also less noise. When the difference in noise levels between nearby and deferred is larger than the difference in the speed at which information is impounded, *IS* and *CS* may lead to an over-stating of the price discovery contribution of the contract with lower trading volume. Yan and Zivot (2010) propose a combination of *IS* and *CS* that nets out transitory frictions which cause the noise. Their measure, termed "information leadership" (*IL*) in Putniņš (2013), is expressed as:

$$IL_1 = \left|\frac{IS_1}{IS_2}\frac{CS_2}{CS_1}\right|, IL_2 = \left|\frac{IS_2}{IS_1}\frac{CS_1}{CS_2}\right| \tag{2.10}$$

where IL_1 and IL_2 are the *IL* measures for nearby and deferred contracts, respectively. The *IL* is not a "share." For comparability and interpretation, Putniņš (2013) defines information leadership shares for nearby (ILS_1) and deferred (ILS_2) as:

$$ILS_1 = \frac{IL_1}{IL_1+IL_2}, ILS_2 = \frac{IL_2}{IL_1+IL_2}. \tag{2.11}$$

Since the *ILS* is more robust to differences in noise, when prices are cointegrated, we use the *ILS* measure.

2.3 Data

The analysis uses corn and live cattle futures contracts traded at the Chicago Mercantile Exchange (CME). These markets represent the most actively traded storable and non-storable agricultural commodities. The sample period studied for corn is from January 14, 2008 through December 14, 2015, and the period used for live cattle ranges from January 1, 2008 to December 31, 2015. This period is characterized by a growing relevance of electronic trading in agricultural commodity futures markets. The electronic platform's shares of corn and live cattle futures trades were about 80% and 10% at the beginning of 2008 (Irwin and Sanders, 2012), and both rose to over 95% in July 2015, after which CME closed pit trading (Gousgounis and Onur, 2017). The period examined also includes pit trade closure, price boom-bust cycles, as well as periods when the markets were inverted, i.e., when the price of deferred futures contracts was lower than the price of the nearby contract.

We use high frequency transactions prices time stamped to the second and ordered chronologically by sequence numbers. Data are obtained from CME Group's Top-of-Book Electronic Platform database. To study price discovery, we need to define an intraday sampling frequency. Janzen and Adjemian (2017) use 1-minute sampling intervals and take the first transaction price in each 1-minute interval. When there is no transaction within a given minute, they replace the missing value using the most recent transaction price. However, this can generate two problems. First, since trading is becoming more frequent, multiple trades can occur even within such a short-time interval, making it difficult to accurately identify which price moves first. Second, prices can vary little during periods of the day and replacing missing observations can lead to stale prices, increasing difficulties in model specification, making residuals serially dependent, and reducing the ability to accurately identify price discovery.

Researchers facing similar problems in empirical microstructure price discovery studies (Brogaard, Hendershott, and Riordan, 2014; Hansen and Lunde, 2006; Hasbrouck, 2018) have switched to event time analysis which in our case limits the analysis to prices which correspond to an actual transaction in at least one of the contracts. Our event time analysis is consistent with the evidence that information flows take place through trading (Kyle and Obizhaeva, 2016; Evans and Lyons, 2008). Beginning with seconds, we keep only seconds when at least one transaction occurs in either the nearby or deferred contract. In the situation where only one contract has transactions, these prices are matched with the last transaction price in the other contract.[1] When the two contracts have a different number of transactions within a second, we first match them by their sequence number, then match any remaining (unmatched) transactions in one contract with the most recent transaction in the other contract. While CME electronic trading system is open nearly 24 hours a day, we only consider the day-time trading session for both the corn and live cattle contracts when the most active trading occurs. On each day, we have an average between 19 to 30 thousand observations for each contract pair in the corn market, and between 4 and 7 thousand observations for contract pairs in the live cattle market (see supplementary result 2 for detailed summary statistics on the number of daily observations).

The corn futures contract has five delivery months (March, May, July, September, and December) and live cattle futures have six delivery months (February, April, June, August, October, and December). Since volumes in the distant deferred contracts are quite low, we use the first five (four) nearby contracts for corn (live cattle), and refer to them as the nearby,

[1] See supplemental result 1 for the percentage of cases where a transaction's price in one contract is matched with the last transaction in the other contract. We also tried using sampling intervals of 1-second and results are similar.

deferred 1, deferred 2, and so on. Corn futures contracts expire on the business day prior to the 15th calendar day of the delivery month and live cattle futures contracts expire on the last business day of each maturity month. We define a contract to be the nearby from the business day after the previous nearby contract expires through the current nearby contract expiration. We do not roll the nearby contract to the next, since we clearly want to identify how price discovery share in the nearby declines as expiration approaches.

2.4 Empirical Results

2.4.1 Cointegration Tests

Since the *ILS*, as well as *CS* and *IS*, are based on a VECM, we test for cointegration first. Daily nearby and deferred futures prices are often found to be cointegrated in the literature. However, intraday prices for nearby and deferred futures may not be cointegrated due to the presence of inverse carrying charges in a storable commodity, short-run market inefficiency (Schroeder and Goodwin, 1991), and pricing of the nearby contract being altered by delivery conditions (Garcia, Irwin and Smith, 2015).

For each sample day, we employ Johansen tests to assess cointegration between the nearby and each deferred contract. Lags for the test are selected based on the Bayesian Information Criteria (BIC) for each day. Consistent with equation (2.5), a constant term is included in the cointegrating vector to allow for storage costs. We have 1992 and 2015 sample trading days for corn and live cattle, respectively. However, we excluded slightly more than 10 days for live cattle as prices varied little due to primarily limit moves, but also because deferred contracts were not sufficiently active to allow for testing.

We follow Fricke and Menkhoff (2011) and use Johansen rank test to categorize data in each day into three categories: 1) Stationarity: intraday nearby and deferred futures prices are both stationary $I(0)$ series, in which case we fail to reject the null hypothesis of a rank of 2 at the 5% significance level. 2) Cointegration: intraday nearby and deferred futures prices are cointegrated $I(1)$ series, in which case we fail to reject the null hypothesis of a rank of 1 at the 5% significance level. 3) Non-cointegration: intraday nearby and deferred futures prices are both non-stationary and not cointegrated, in which case we fail to reject the null hypothesis of a rank of 0 at the 5% significance level.

Table 2.1 summarizes the percentage of days that belong to each category. The probability of both prices being stationary ranges from 8% to 11% in the corn market and about 6% across contract pairs in the live cattle market. The percentage of days in which the nearby and deferred futures are cointegrated is about 80% across all contract pairs for corn and 70% for live cattle. The percentage of non-cointegration days ranges from 4% to 12% for corn and from 16% to 27% for live cattle. The percentage of non-cointegration generally increases at more deferred contracts. This is consistent with Tomek and Gray (1970) and Tomek (1997) and shows that contracts at more distant maturities may reflect different price information (i.e., expected supply and demand), particularly in live cattle where prices are not linked through storage costs.

Figure 2.1 and figure 2.2 present the distribution of Johansen test results through time for corn and live cattle, respectively. Each observation is colored coded to reflect the test results and located relative to the vertical axis to represent the nearby contract volume share (*VS*). *VS* equals the volume of the nearby contract divided by the total volume of the nearby and deferred contracts on the same day. Shaded areas represent periods of inverse carrying charges. In both figures, we see *VS* presents a cyclic pattern, with the nearby contract's *VS* decreasing as expiration approaches and then increasing with the shift to the next nearby contract. In figure 2.1, a clear pattern emerges for corn, with most non-cointegration days (red squares) appearing in periods of inverse carrying charges. This result shows that an inverted market reduces the link between different maturities in storable commodities and is consistent with Working's (1948 and 1949) theory. Consistent with live cattle's non-storable character, figure 2.2 shows non-cointegration in live cattle does not concentrate in periods when the market was inverted. In both markets, we find that as expiration approaches, the number of non-cointegration days increases.[2] Further, in the corn market nearby and deferred futures prices are less likely to be both stationary in the first few weeks after entering the nearby period.

2.4.2 Price Discovery Shares for the Nearby Contract Relative to Deferred Contracts

We calculate *ILS* for each day when intraday nearby and deferred transaction prices are cointegrated. For days when both prices are stationary, we use the *GS* measure. Non-cointegration days are not included in the price discovery analysis as these prices do not share and discover a common efficient price (Fricke and Menkhoff, 2011). We calculate daily price discovery shares for the nearby contract relative to each deferred contract separately. For *ILS,* the BIC recommends estimating a VECM that has between 1 and 10 lags for both commodities. Following Garbade and Silber (1983), we set negative estimates of β_1 and β_2 to 0 when calculating the *GS* measure, since they have no conceptual meaning. Hereafter, we refer to *PS* as the combination of *ILS* and *GS* as they reflect the same basic notion of price discovery.

Table 2.2 reports the averages of daily price discovery and volume shares for the nearby contract and deferred contracts for corn and live cattle. Although *ILS* is the preferred price discovery measure when data are cointegrated, we also report *CS* and *IS* for comparison. In both markets, price discovery shares as well as *VS* for the nearby contract generally increase with the

[2] To save space, details are in supplementary result 3.

temporal distance between the nearby and deferred contract.[3] This term structure is expected because volume and accompanying liquidity at distant horizons are usually lower which implies less market information. Although *CS*, *IS*, and *ILS* suggest the same term structure, in both markets, *CS* and *IS* for the nearby contract are consistently lower than the *ILS* when the nearby contract has a higher *VS*. As the *CS* and *IS* give higher share to less noisy price series relative to the *ILS*, their relatively lower values suggest that the more actively traded nearby contracts are noisier than deferred contracts.[4] Since *CS* and *IS* are biased compared to *ILS*, hereafter, we focus on *PS* as the price discovery measure. For corn futures, the nearby contract only slightly dominates the first deferred contract with an average *PS* of 53%. However, the nearby *PS* rises quickly when the nearby is compared to more deferred contracts. By the deferred 4 contract the nearby *PS* has reached 83%.

Compared to corn, the nearby live cattle contract is less dominant in price discovery. On average, the nearby contract does not provide more price discovery than the next nearby contract with an average *PS* of 37%. Compared to the second and third deferred contracts, the nearby contract contributes about 57% and 67% of the price discovery. Informatively, *VS*s for deferred 1 contract are appreciably below 50%, and only reach 50% for the deferred 2 contract suggesting much less trading in the nearby contract which again contrasts with the corn contract. For live cattle futures, since there is no storage arbitrage to link the contracts with differing maturities, contracts for different delivery dates provide information of equilibrium conditions at different

[3] *VS* and *PS* are not statistically different at the 5% level only in the first contract pair for corn and the first and last contract pair for live cattle

[4] Noise levels are presented in supplementary result 4.

future dates (Leuthold, Junkus and Cordier, 1989). Thus, the difference in the dominance of the nearby contract for corn and live cattle can be attributed to their difference in storability.

Figures 2.3 and 2.4 show the daily dynamics of *PS* for the nearby contract relative to each deferred contract for corn and live cattle.[5] In both figures, we observe a cyclic pattern where the nearby contract's *PS* decreases as expiration approaches which is similar to the behavior of *VS* presented in figures 2.1 and 2.2. This pattern is stable in the corn market, except when the market is inverted during which price discovery shares become more volatile. In the live cattle market, although *PS* follows a similar general cyclic pattern, it is more volatile relative to the corn market. Examination of the plots also suggests a dominance of the nearby contract relative to distant contracts particularly in the corn market. The nearby *PS* is progressively more concentrated near the top of the *PS* range at more deferred contracts, which is consistent with the term structure of *PS* (table 2.2).

2.4.3 Price Discovery Shares in the Nearby Period for Each Contract Month

To examine more closely the behavior of price discovery, we average *VS* and *PS* for nearby contracts across years and plot them for each contract (figures 2.5 and 2.6). The horizontal axis in all plots measures the trading days to contract expiration. Because the December corn contract becomes important for hedging and pricing early in the marketing year as it reflects information on the new crop, we also compare the nearby July contract to the December contract (figure 2.5, panel 6).

*PS*s for corn are presented in figure 2.5 and exhibit similar patterns for most contract months, except when September is the nearby contract. Initially, most nearby contracts have a *PS* around 80% and continue to dominate (i.e. a share higher than 50%) price discovery until 2-3 weeks

---

[5] We present information on the error correction parameters in supplementary result 5.

prior to contract expiration. In general, the *PS* moves in tandem with the *VS*, declining sharply as trading volume decreases in the nearby contract. The nearby contract loses its dominance in price discovery nearly at the same time as it loses dominance in volume. This typically happens 2-3 weeks prior to contract expiration which roughly coincides with the beginning of the delivery window.

The notable exception to the price discovery pattern described is the September-December contract pair (figure 2.5, panel 4). While most contracts begin with a *PS* at nearly 80%, the September contract only initially and briefly breaks 50%, and then remains well below 40% to expiration. This suggests that the September contract does not have a dominant role in price discovery even when it is the nearby contract. Since the December contract reflects information on the new crop which becomes highly relevant in the summer months, we also examine the *PS* and *VS* for the July contract relative to December contract (figure 2.5, panel 6). This exhibits a pattern more similar to the other corn panels.[6] Hence, as the July contract approaches expiration, price discovery shifts rapidly to the December contract. The relative lack of importance of the September contract in the pricing process is likely due to its position between two crop years. Other researchers have suggested avoiding using the September contract for constructing nearby series based on its price volatility patterns (Smith, 2005) and ability to predict subsequent cash prices (Leath and Garcia, 1983). While the September contract does not dominate the price discovery process, it still accounts for 30-40% of the price discovery before entering the delivery month and therefore should not be totally ignored as a price signal. However, the leadership of

[6] We also compare the December and May contracts. Results show a similar pattern, except the May contract is more dominant. This is found in supplementary result 6.

the December contract for such an extended period indicates that December futures prices become the focus of hedgers and market participants even early in the summer months.

Figure 2.6 plots the *PS* for the live cattle market. Both *PS* and *VS* exhibit a similar general downward trend across all contract months. The nearby contract loses its pricing dominance almost at the same point when *VS* declines below 50%. Both the dominance of price discovery and trading volume of the nearby contract switch to the next nearby contract about 2 weeks prior to the delivery, which is earlier than in the corn market. The earlier switch is likely because of the added days in delivery months in the live cattle market. Compared to the corn market, *PS* is more volatile in delivery in the live cattle market. The *PS* initially decreases but declines more slowly or remains relatively stable close to expiration. In some contract months, *PS* can even increase in the final trading days,[7] which reflects unstable intertemporal price relationships near expiration in the live cattle market that has been widely identified in the literature (Leuthold, 1972; Naik and Leuthold, 1988). Similar behavior, although to a much less extent, can be found in the corn market which could reflect price distortions in the nearby futures in delivery found in grain markets in recent years (Garcia, Irwin and Smith, 2015).

2.4.4 Regression Analysis

In this section, we assess the relationships between the nearby contracts' *PS* relative to deferred contracts and several factors using a regression framework. The analysis focuses on equations (2.12) and (2.13) for the corn and live cattle markets, respectively:

$$PS_{corn,d} = f(VS_d, Expiration_d, Expiration^2{}_d, Inverted_d, Tomek_d,$$
$$WASDE\&CP_d, GRS_d, Drought_d, Crash_d, Pit_d, Indexroll_d, Stationarity_d). \qquad (2.12)$$

[7] Examination of the individual *PS*s revealed that the values in figure 2.6 are not driven by outliers or non-cointegration days.

$$PS_{livecattle,d} = f(VS_d, Expiration_d, Expiration^2{}_d, Inverted_d,$$

$$CF_{d-1}, WASDE\&CP_d, GRS_d, Drought_d, D2015_d, Indexroll_d, Stationarity_d). \quad (2.13)$$

where *PS* is the day *d* price discovery share for the nearby relative contract. *VS* and *Expiration* are the nearby contract volume share and a variable that counts the number of days to expiration, respectively. Similar factors were identified as relevant in explaining *PS* in the VIX (Chen and Tsai, 2017) and bond futures markets (Mizrach and Neely, 2008; Fricke and Menkhoff, 2011). To allow for the non-linear pattern observed in the price discovery (Figure 2.3 and 2.4), we include a quadratic term, $Expiration^2$. Because we use a combination of *GS* and *ILS* as our price discovery measure, the magnitude of the *PS* on a given day may be affected by the price discovery measure applied. Thus, we introduce a dummy variable $stationary$ that equals one if the prices are both stationary in which case the *GS* is used.

To measure the relationship between *PS* and an inverted market, we create a dummy variable $Inverted$ that equals one on days when deferred futures settlement price is below the nearby futures settlement price. We expect $Inverted$ to be non-significant in the non-storable live cattle futures' regressions. However, consistent with Working's (1948, 1949) view of price discovery, an inverted market may be negatively correlated with *PS* in the storable corn market. To test Tomek's hypothesis, we create a dummy variable *Tomek*, which equals 1 when the deferred contract represents a new crop year and the nearby contract represents the old crop. As the nearby and deferred 4 contracts always represent prices for different crop years, the dummy variable *Tomek* is not included in the equation for those contracts. As predicted by Tomek and Gray (1970) and Tomek (1997), we expect the nearby contract to have a smaller share of price discovery when nearby and deferred contracts represent prices for different crop years.

In addition, we include several market factors which may be related to the price discovery process. The first is USDA market reports which impart important information about fundamentals that is quickly reflected in futures prices across maturities (McKenzie, 2008; Adjemian, 2012; Dorfman and Karali, 2015). For the corn market, we consider three important USDA grain market reports: World Agricultural Supply and Demand Estimate (WASDE) report, Crop Production (CP) report and Grain Stocks (GRS) report. Since the WASDE and CP reports are usually released in the second week of each month on the same day, we create a single dummy variable $WASDE\&CP$ for the two reports. The GRS report releases are quarterly and release days, captured by the dummy variable GRS, are usually in mid-January and the end of March, June and September. For live cattle, we use the Cattle on Feed (captured by the dummy variable CF) report that has the largest impact on the live cattle market among all USDA reports (Isengildina, Irwin, and Good, 2006). We also include corn market reports as corn is used as feed for live cattle.[8]

Another set of factors is included to account for period events. Both markets experienced dramatic price declines which we capture through the dummy $Crash$. For corn, $Crash$ equals one from July 03, 2008 when prices peaked to December 08, 2008 when prices hit bottom. The brief period of rapid corn price declines from August 09, 2012 to September 14, 2012, which followed a run up due to the severe drought in the summer of 2012, is captured by the dummy

[8] Grain market reports are released either before market opening or during the regular trading session and thus should affect prices on the release day. *CF* reports are released on the third Friday of each month after regular trading hours and thus should affect price the next day. Therefore, the dummy variables for the grain market reports equal 1 on the release day and the dummy variable for *CF* reports equals 1 on the following trading day.

variable $Drought$. During this one-month period, corn nearby futures price declined 7.89%. Since the drought also affected the live cattle market, the variable $Drought$ for live cattle equals 1 from December 19, 2012, when nearby futures' price peaked at 134.40 cents/lb after the drought, to May 20, 2013 when the price bottomed at 118.00 cents/lb. A more sustained collapse in cattle prices occurred in 2015 causing concerns about the price discovery function of the live cattle futures market. This period overlaps the time of CME's pit trading closure, which was announced in February 2015 and started officially in July 2015. Therefore, we create the dummy $D2015$ which equals 1 in the year of 2015, to capture the possible joint price decline and pit closure. The corn market remained relatively stable during 2015, providing a good opportunity for identifying changes related to the pit trading closure. Consistent with Gousgounis and Onur (2017), we create a dummy variable Pit that equals 1 after February 4, 2015 when the closing of pit trading was announced for corn.[9] Commodity index funds have increased investments in commodity futures markets (Irwin and Sanders, 2012). These funds typically follow a predetermined schedule to roll their positions from the nearby to the next nearby contract. Considering the vast position changes involved in the rolling process, changes may occur in the discovery process. We include a dummy variable $Indexroll$ that equals 1 between the fifth and tenth business days of the month prior to expiration, which includes the roll periods of the two largest commodity indices: S&P Goldman Sachs and Dow Jones UBS commodity indices.

2.4.5 Regression Results

Regression models are estimated for each pair of contracts in corn and live cattle markets using the OLS and presented in Table 2.3 and 2.4, respectively. Heteroscedasticity and autocorrelation

[9] Similar results are obtained using the period after July 2015 when this decision was officially executed.

robust standard errors using the Newey-West (1994) estimator are presented in parentheses and adjusted *R*-squared values are in the lower panel in each table. Adjusted *R*-squared values in corn models are consistently higher than in live cattle models, reinforcing the graphical analysis which demonstrates that price discovery in the corn market is more correlated with observable variables than in the live cattle market. In each market, *R*-squared values decrease as the length of time between the nearby and deferred contracts increases, indicating the model fits better for contract pairs with closer maturities. The coefficients for stationarity are not significant suggesting that our results are robust to the selection of price discovery measures.

Consistent with the economic intuition that information is incorporated in the market through volume, the *VS* coefficient is significant and positive across contract pairs in both markets. A 1% increase in the nearby contract's *VS* is associated to a 0.56% to 0.69% increase in the nearby contract's *PS* in the corn market, and between 0.30% and 0.39% in the live cattle market.

In the corn market, the coefficients of the days to expiration variable and its quadratic term are both significant. The coefficients of *Expiration* are positive and indicate that one day closer to contract expiration is associated with a 0.6%-1.1% decrease in the nearby contact's *PS*. The significant negative $Expiration^2$ parameters indicate the decline in *PS* occurs in a non-linear fashion, first declining gradually and then dropping off more sharply as expiration approaches. The relationship between days-to-expiration and *PS* is only significant in the first contract pair in the live cattle market. The coefficient of *Expiration* in the nearby and deferred 1 contract pair is significant and negative, while its quadratic term is significant and positive. This evidence supports the earlier observation in figure 2.6 that the *PS* for the live cattle nearby contract declines as expiration approaches but can increase in the last few days.

The coefficients of *Inverted* and *Tomek* provide some support for the implications of Working's and Tomek's theories on price discovery. Consistent with Working's theory, price discovery along the forward curve is uncorrelated with the inverted market indicator in the non-storable live cattle market. In the corn market and for the first deferred pair, the nearby contract's *PS* increases about 6.8% when the market is inverted. However, in the other corn contract pairs, the *PS* is not significantly related to *Inverted*. As expected, the coefficient for *Tomek* has a negative sign in all cases, albeit it is not statistically different from zero in the nearby and first deferred contract pair. These findings indicate a decline in the price discovery dominance of the nearby contract when the deferred contract represents a different crop year.

The correlations of *PS* with USDA reports, price declines and commodity index position rolls, are significant in the corn market but not in the live cattle market. In addition, no statistically significant relationship between the closure of pit trading and *PS* along the forward curve are found in either market. On average, the *PSs* for corn nearby contracts are generally lower on USDA report days. Coefficients for *WASDE&CP* are negative across all contract pairs and statistically different from zero in the second and third pairs. There is a statistically significant and negative correlation between *GRS* and *PS* in the nearby and deferred 4 contract pair. The result likely reflects that grain reports contain outlook information and market participants use this information to adjust their forecasts for longer horizons, therefore improving price discovery in deferred contracts. These findings are consistent with Tomek and Gray's (1970) and Tomek's (1997) view that futures along the forward curve provide information about expected future supply-demand information.

Several small positive relationships between price declines and *PS* are found in the corn market. *PS* is correlated with the dummy *Drought*. During that period, the *PS* of the nearby contract relative to the second deferred contract significantly increased by 4.4%. However, no significant coefficients are found in other contract pairs. The response to the July-early December 2008 crash in prices appears more relevant. Nearby and deferred 1 contracts responded similarly to the sharp decline in prices. At more distant horizons, the nearby contracts, which were most closely tied temporally to the initial sharp decline and possessed greater liquidity, responded more quickly. The insignificant coefficient in the nearby and deferred 4 contract pair is positive and follows what appears to be a declining importance of the nearby contract through time.

Index commodity roll periods are positively correlated with the *PS* of the nearby contract for the first two contract pairs. However, the magnitude and significance of the parameter decreases from 3.9% in the first contract pair to 3.6% in the second contract pair and becomes insignificant in the third and fourth contract pairs. Several studies have documented a "sunshine trading effect," which consists of a predetermined commodity index roll period attracting counterparties and increasing liquidity supply (Shang, Mallory and Garcia, 2018). Since traders are highly concentrated in the nearby contract during index rolling periods (Aulerich, Irwin and Garcia, 2014), one possible explanation for the positive correlation between index rolling and *PS* could be that liquidity improvement caused by the "sunshine trading effect" is more pronounced in nearby contracts than deferred contracts.

2.5 Conclusions

Understanding price discovery along the futures forward curve is important for market participants in making sound trading, hedging, and production decisions. In the corn and live cattle futures markets, we quantify price discovery using intraday data, and graphical and statistical analysis for the 2008-2015 period—a period characterized by highly volatile prices and the closure of pit trading. We measure the price discovery share between nearby and deferred contracts and identify when the dominance of price discovery switches between contracts. We also estimate the importance of the factors related to price discovery.

Our results provide nuanced support for the theory of price of storage. We find price discovery is more dominated by the nearby contract in the storable corn market than the non-storable live cattle market. In addition, except when the market is inverted, intraday nearby and deferred futures prices are more likely to be cointegrated in the storable corn market than in the non-storable live cattle market. However, deferred contracts play a non-trivial role in price discovery not only in the non-storable live cattle market, but also in the storable corn market, particularly when the deferred contract prices represent a new crop year. This demonstrates the importance of futures' forward pricing role in price discovery as argued by Tomek and Gray (1970) and Tomek (1997).

The price discovery share of the nearby contract decreases in both markets as expiration approaches and trading becomes less active. The nearby contract leadership in price discovery is tightly related to trading volume. This finding has a practical implication for researchers and practitioners who need to construct a continuous series of nearby contracts. Since price discovery is closely linked to volume share, we recommend rolling to the next nearby contract when it achieves more than 50% of the volume share, instead of using date-based methods. Another informative finding is that the September corn contract rarely dominates the next nearby (December) contract. This suggests that price discovery is dominated by the December contract as early as the beginning of July.

The regression results show in both markets the share of price discovery along the forward curve is strongly linearly correlated with trading volume and nonlinearly correlated with time to expiration. Other market-related factors only have statistically meaningful correlations with the share of price discovery in the corn market. During periods of price declines and commodity index rolls, price discovery in the corn market is more likely to concentrate in the more traded nearby contract than in deferred contracts. Consistent with Working's theory, we find that an inverted market increases price discovery in the corn nearby contract, though only in the first deferred pair. In addition, Tomek and Gray's (1970) and Tomek's (1997) view that futures not only provide guidance for carrying inventories but also forecasts of expected futures prices, is supported. We find that deferred contracts' price discovery role becomes more important when they represent a new crop year and on days when the USDA releases important forward-looking market information.

Overall, while the result that price discovery occurs principally in nearby futures contracts is consistent with earlier more descriptive analyses based on daily prices (e.g., Working, 1948, 1949; Tomek, 1997; Leuthold et al., 1989), our findings highlight the dynamic and systematic aspects of the price discovery process in agricultural markets. Differences exist between storable and non-storable markets, but their importance in price discovery can also vary in time and by market. Despite these differences, the clear relationship across markets between volume share and relative price discovery is striking. Informatively, while nearby contracts tend to dominate the price discovery process, we identify the non-trivial role that deferred contracts play in today's fast-moving markets. This information should be of value to the pricing and hedging decisions that market participants make and to researchers interested in uncovering relevant relationships in electronically traded agricultural markets. Future research might expand our analysis to other markets, examine the relationships at more disaggregate temporal units and intervals within the day, and focus more specifically on how the intraday price discovery process changes on USDA announcement days.

2.6 Tables and Figures

Table 2.1 Distribution of Days Based on Johansen Rank Test Results, 2008-2015

	Stationarity	Cointegration	Non-cointegration	Total
Corn				
Nearby and Deferred 1	11.75%	84.14%	4.11%	1992
Nearby and Deferred 2	9.59%	83.89%	6.52%	1992
Nearby and Deferred 3	8.89%	83.28%	7.83%	1992
Nearby and Deferred 4	8.38%	79.62%	12.00%	1992
Live Cattle				
Nearby and Deferred 1	6.85%	75.73%	16.92%	2005
Nearby and Deferred 2	5.91%	72.26%	21.29%	2003
Nearby and Deferred 3	5.61%	65.76%	27.99%	2002

Note: Results are based on Johansen rank hypothesis tests between intraday nearby and deferred futures prices at the 5% significance level using trace statistics. Percentages are given as percentage of total number of days.

Table 2.2 Average Price Discovery and Volume Shares for the Nearby Contract, 2008-2015

Contract Pair	*CS*	*IS*	*ILS*	*GS*	*PS*	*VS*
Corn						
Nearby vs Deferred 1	0.515	0.504	0.535	0.571	0.537	0.532
Nearby vs Deferred 2	0.473	0.590	0.697	0.724	0.699	0.700
Nearby vs Deferred 3	0.524	0.662	0.784	0.761	0.783	0.797
Nearby vs Deferred 4	0.438	0.701	0.830	0.801	0.831	0.855
Live cattle						
Nearby vs Deferred 1	0.425	0.394	0.369	0.571	0.367	0.332
Nearby vs Deferred 2	0.569	0.550	0.574	0.724	0.571	0.504
Nearby vs Deferred 3	0.517	0.637	0.680	0.761	0.672	0.621

Note: Days with cointegrated $I(1)$ intraday nearby and deferred prices are used for *CS*, *IS*, and *ILS*. Days with stationary intraday nearby and deferred prices are used for the *GS*. *PS* represents the combination of *ILS* and *GS* estimates. *VS* is the nearby contract's volume share that equals the volume of the nearby contract divided by the total volume of the nearby and deferred contracts on the same day.

Table 2.3 Regression Results for the Corn Futures Contracts, 2008-2015

	Nearby and Deferred 1	Nearby and Deferred 2	Nearby and Deferred 3	Nearby and Deferred 4
VS	0.683***	0.560***	0.625***	0.687***
	(0.037)	(0.041)	(0.043)	(0.057)
Expiration	0.007***	0.011***	0.009***	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)
Expiration^2	-0.000**	-0.000***	-0.000***	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Inverted	0.068***	0.009	-0.007	-0.005
	(0.014)	(0.017)	(0.014)	(0.013)
Tomek	-0.026	-0.043***	-0.027*	
	(0.016)	(0.013)	(0.013)	
WASDE & CP	-0.022	-0.044**	-0.071***	-0.043
	(0.013)	(0.017)	(0.021)	(0.025)
GRS	0.006	0.012	-0.021	-0.101*
	(0.035)	(0.042)	(0.038)	(0.049)
Drought	-0.005	0.044*	0.038	-0.002
	(0.023)	(0.021)	(0.050)	(0.046)
Crash	-0.010	0.062*	0.067***	0.044
	(0.017)	(0.027)	(0.019)	(0.024)
Pit	0.011	-0.004	-0.014	-0.009
	(0.010)	(0.014)	(0.017)	(0.017)
Indexroll	0.039***	0.036**	0.006	0.012
	(0.010)	(0.012)	(0.014)	(0.012)
Stationarity	0.013	0.036***	0.010	-0.006
	(0.008)	(0.010)	(0.011)	(0.011)
Intercept	0.038**	0.081***	0.112***	0.103*
	(0.014)	(0.024)	(0.032)	(0.042)
Adjusted R^2	0.78	0.70	0.64	0.52
Observations	1908	1861	1834	1752

Note: *VS* is the nearby contract's volume share. *Expiration* is the number of days to the nearby contract's expiration. *Inverted* is a dummy variable for days in which deferred futures settlement price was below nearby futures settlement price. *WASDE&CP* is a dummy variable for USDA WASDE and Crop Production report days. *GRS* is a dummy variable for USDA Grain Stocks report days. *Drought* is dummy variable for the period of declining corn prices following the 2012 drought. *Crash* is a dummy variable for the corn market crash period in 2008. *Pit* is a dummy variable for days after CME's pit closure announcement. *Tomek* is a dummy variable which equals 1 when the deferred contract represents a new crop year and the nearby contract represents prices for the old crop. *Stationarity* is a dummy variable for days in which intraday nearby and deferred prices were both stationary. Heteroscedasticity and autocorrelation robust standard errors are reported in parenthesis. Asterisks ***, **, and * indicate significance at the 0.1%, 1%, and 5% levels.

Table 2.4 Regression Results for the Live Cattle Futures Contracts, 2008-2015

	Nearby and Deferred 1	Nearby and Deferred 2	Nearby and Deferred 3
VS	0.305***	0.387***	0.348***
	(0.081)	(0.079)	(0.080)
Expiration	-0.008**	0.005	0.006
	(0.003)	(0.003)	(0.004)
Expiration^2	0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)
Inverted	-0.011	0.020	0.011
	(0.013)	(0.016)	(0.019)
CF	0.036	-0.009	-0.041
	(0.028)	(0.030)	(0.033)
WASDE & CP	-0.026	-0.049	-0.013
	(0.020)	(0.026)	(0.028)
GRS	0.041	0.041	0.015
	(0.049)	(0.052)	(0.043)
Drought	-0.019	-0.032	0.029
	(0.015)	(0.024)	(0.027)
D2015	-0.007	-0.045	0.014
	(0.023)	(0.050)	(0.050)
Indexroll	0.008	-0.002	-0.004
	(0.013)	(0.016)	(0.018)
Stationarity	-0.025*	-0.021	-0.028*
	(0.011)	(0.013)	(0.014)
Intercept	0.294***	0.242***	0.295***
	(0.026)	(0.026)	(0.034)
Adjusted R^2	0.29	0.27	0.21
Observations	1661	1570	1436

Note: *VS* is the nearby contract's volume share. *Expiration* is the number of days to the nearby contract's expiration. *Inverted* is a dummy variable for days in which deferred futures settlement price was below nearby futures settlement price. *CF* is a dummy variable for the trading day following the release of Cattle on Feed report. *WASDE&CP* is a dummy variable for USDA WASDE and Crop Production report days. *GRS* is a dummy variable for USDA Grain Stocks report days. *Drought* is dummy variable for the period of declining live cattle prices following the 2012 drought. *D2015* is a dummy variable for the year of 2015. *Indexroll* is a dummy variable for commodity index rolling periods. *Stationarity* is a dummy variable for days in which intraday nearby and deferred prices were both stationary. Heteroscedasticity and autocorrelation robust standard errors are reported in parenthesis. Numbers are rounded to the third decimal place. Asterisks ***, **, and * indicate significance at the 0.1%, 1%, and 5% levels.

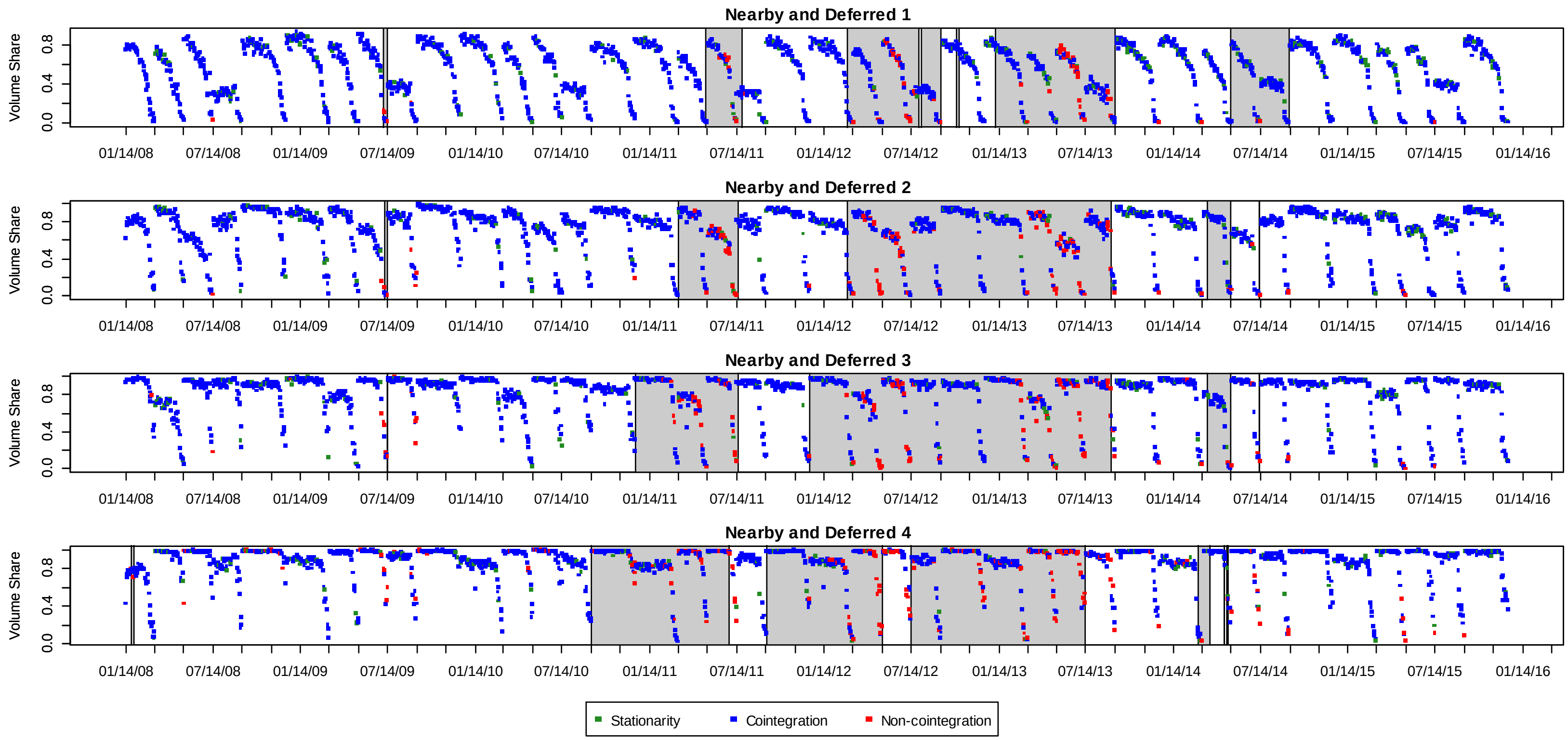


Figure 2.1 Johansen Rank Test Results and Volume Shares for the Corn Futures Contracts, 2008-2015

Note: Shaded areas represent backwardation periods. Corn futures contracts expire on the business day prior to the 15th calendar day of the maturity month.

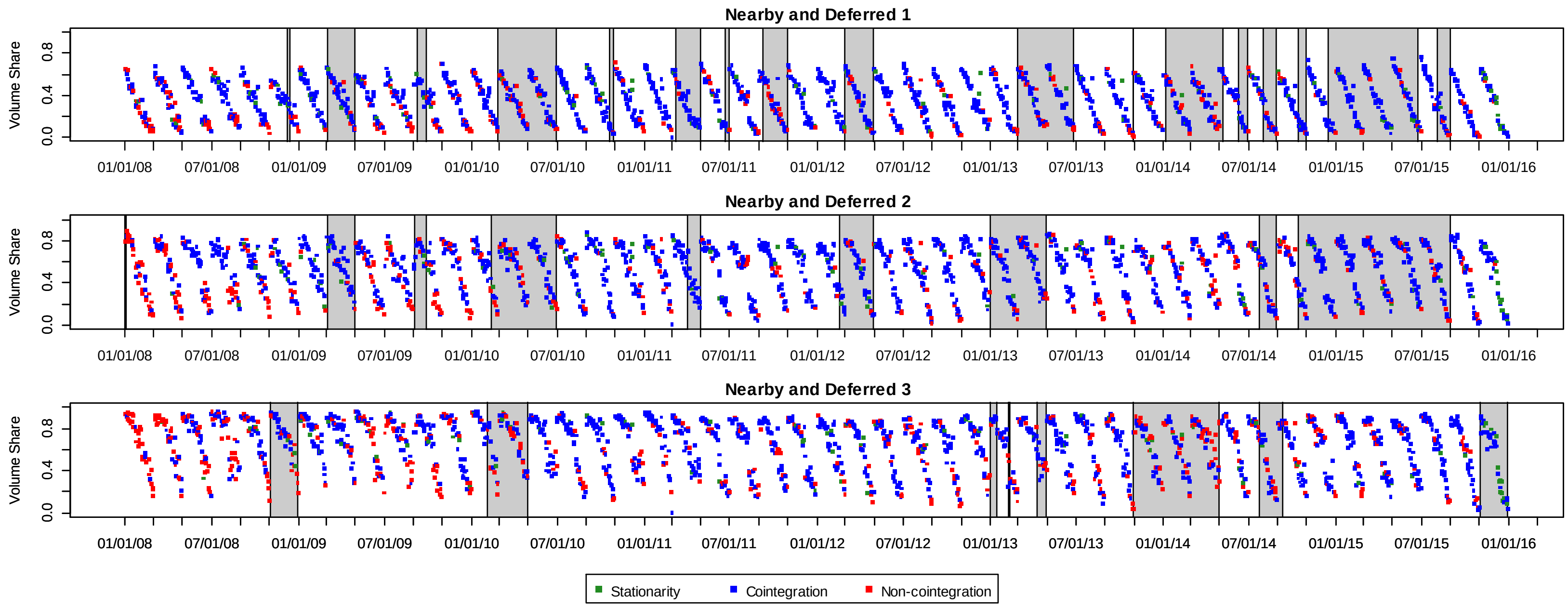


Figure 2.2 Johansen Rank Test Results and Volume Shares for the Live Cattle Futures Contracts, 2008-2015

Note: Shaded areas represent backwardation periods. Live cattle futures contracts expire on the last business day of the maturity month.

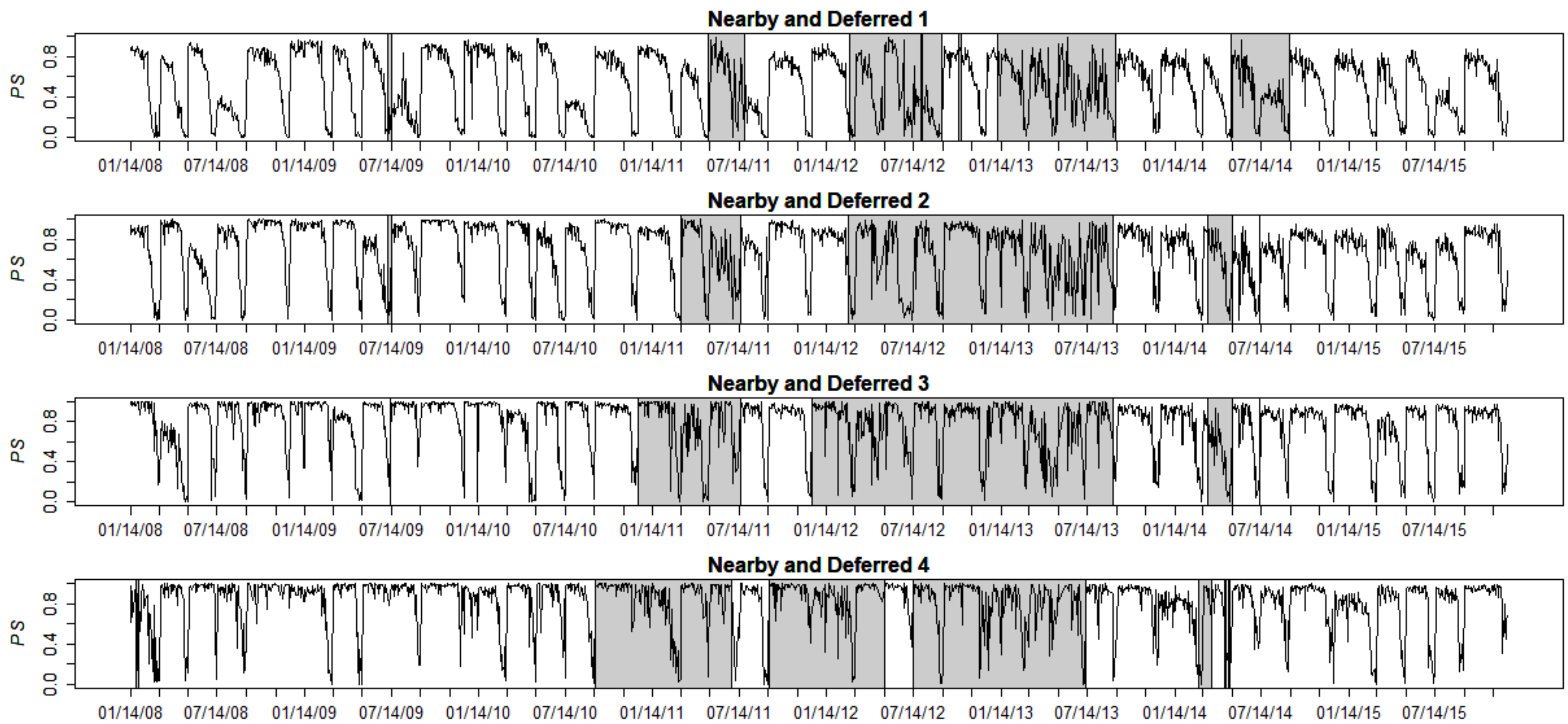


Figure 2.3 Price Discovery Shares for the Nearby Contract Compared to Deferred 1, 2, 3, And 4 Contracts in the Corn Futures Market, 2008-2015

Note: Shaded areas represent backwardation periods. Corn futures contracts expire on the business day prior to the 15th calendar day of the maturity month.

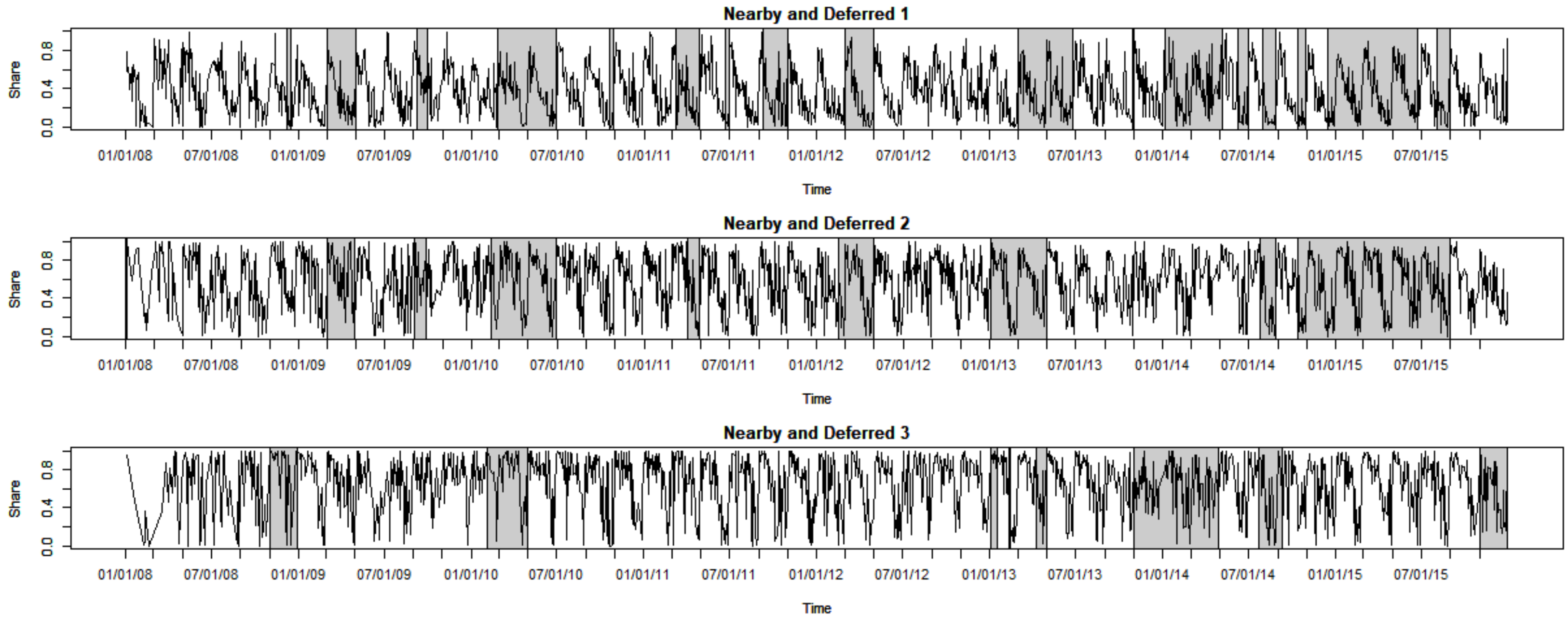


Figure 2.4 Price Discovery Shares for the Nearby Contract Compared to Deferred 1, 2, and 3 Contracts in the Live Cattle Futures Market, 2008-2015

Note: Shaded areas represent backwardation periods. Live cattle futures contracts expire on the last business day of the maturity month.

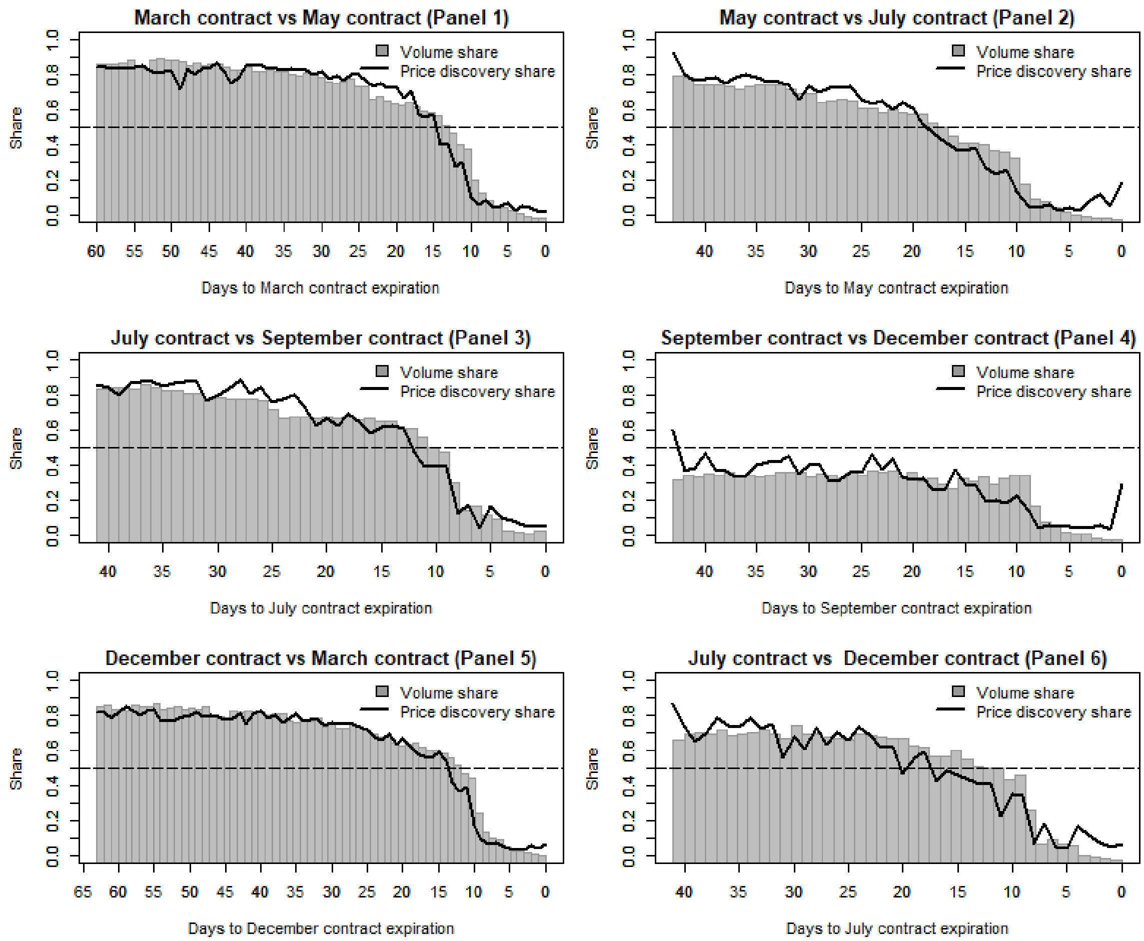


Figure 2.5 Price Discovery and Volume Shares in the Nearby Period for each Contract Month in the Corn Futures Market, 2008-2015

Note: Panels show the average over years of volume share of nearby relative to the first deferred and price discovery share between nearby and first deferred contract. The information is organized along the x-axis by days to maturity of the nearby.

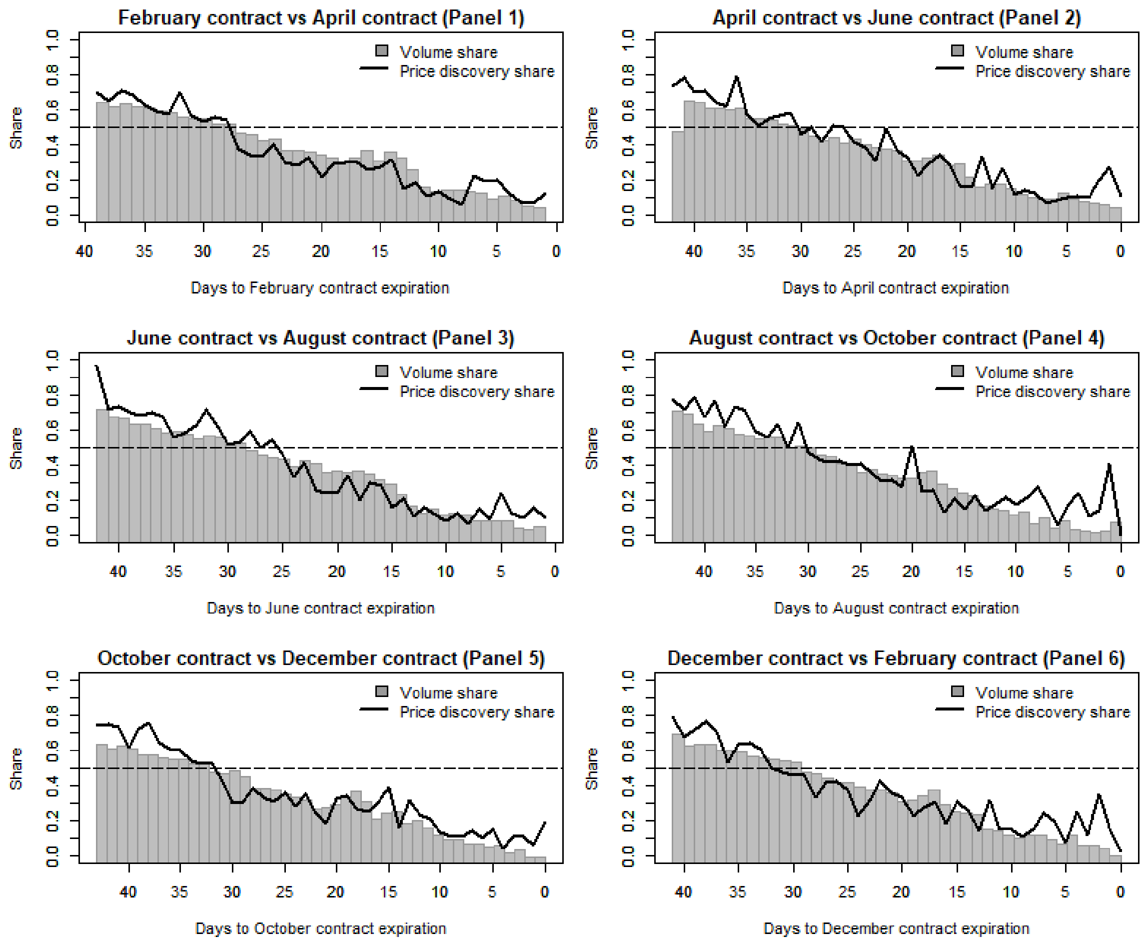


Figure 2.6 Price Discovery and Volume Shares in the Nearby Period for each Contract Month in the Live Cattle Futures Market, 2008-2015

Note: Panels show the average over years of volume share of nearby relative to the first deferred and price discovery share between nearby and first deferred contract. The information is organized along the x-axis by days to maturity of the nearby.

2.7 Supplementary Results

Supplementary Result 1

Supplemental Table 2.1 Percentage of Replaced Observations for each Contract Pair in the Corn Market, 2008-2015

	Contract Pair 1		Contract Pair 2		Contract Pair 3		Contract Pair 4	
	Nearby	Deferred 1	Nearby	Deferred 2	Nearby	Deferred 3	Nearby	Deferred 4
Mean	37.33%	46.84%	22.14%	68.17%	16.03%	78.60%	12.34%	83.71%
Min	0.56%	0.00%	0.03%	0.04%	0.01%	0.14%	0.00%	0.91%
Max	99.96%	93.40%	99.88%	99.40%	99.79%	99.75%	98.82%	99.90%
S.D	37.43%	35.20%	32.29%	32.33%	28.39%	29.49%	24.50%	25.34%

Note: percentage of cases where a transaction's price in one contract is matched with the last transaction in the other contract. While not shown here, these percentages follow a dynamic pattern consistent with volume (they grow for the nearby and decline for the first deferred contract as the nearby contract approaches expiration).

Supplemental Table 2.2 Percentage of Replaced Observations for each Contract Pair in the Live Cattle Market, 2008-2015

	Contract Pair 1		Contract Pair 2		Contract Pair 3	
	Nearby	Deferred 1	Nearby	Deferred 2	Nearby	Deferred 3
Mean	57.53%	23.17%	39.06%	50.10%	29.46%	63.97%
Min	3.99%	0.00%	2.44%	0.49%	1.27%	0.59%
Max	99.89%	80.98%	99.39%	89.80%	99.41%	96.76%
S.D	28.83%	21.30%	27.21%	24.26%	24.70%	23.87%

Note: percentage of cases where a transaction's price in one contract is matched with the last transaction in the other contract. While not shown here, these percentages follow a dynamic pattern consistent with volume (they grow for the nearby and decline for the first deferred contract as the nearby contract approaches expiration).

Supplementary Result 2

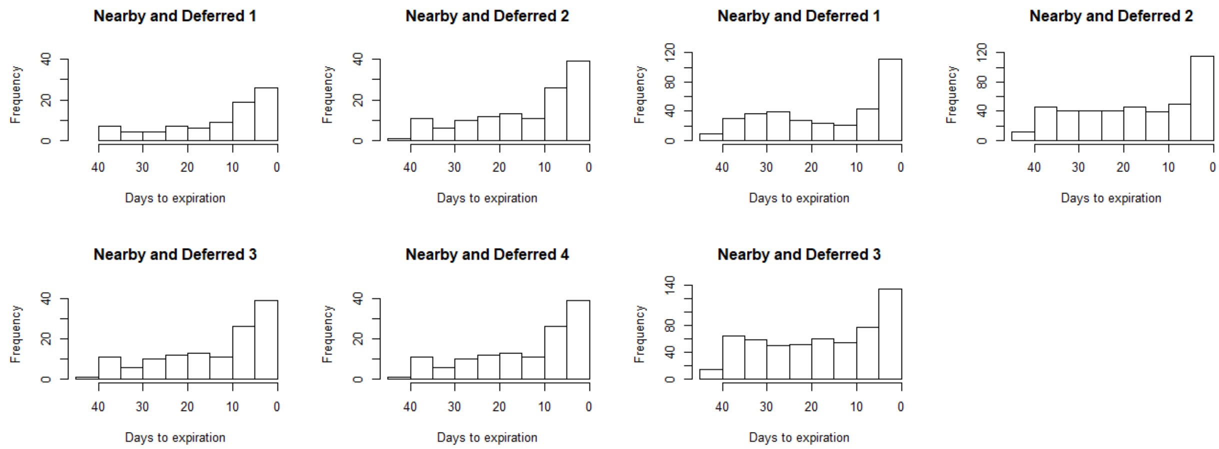
Supplemental Table 2.3 Number of Daily Observations for each Contract Pair in the Corn Futures Market, 2008-2015

	Nearby and Deferred 1	Nearby and Deferred 2	Nearby and Deferred 3	Nearby and Deferred 4
Mean	29,662	22,191	20,168	19,473
Minimum	696	665	253	131
Maximum	152,565	152,059	150,931	154,009
Median	23,868	15,793	14,382	13,756
Standard Deviation	19,607	19,222	18,383	18,423

Supplemental Table 2.4 Number of Daily Observations for each Contract Pair in the Live Cattle Futures Market, 2008-2015

	Nearby and Deferred 1	Nearby and Deferred 2	Nearby and Deferred 3
Mean	7,224	4,471	3,901
Minimum	249	96	28
Maximum	30,815	29,024	27,689
Median	6,778	3,509	2,732
Standard Deviation	5,106	3,778	3,651

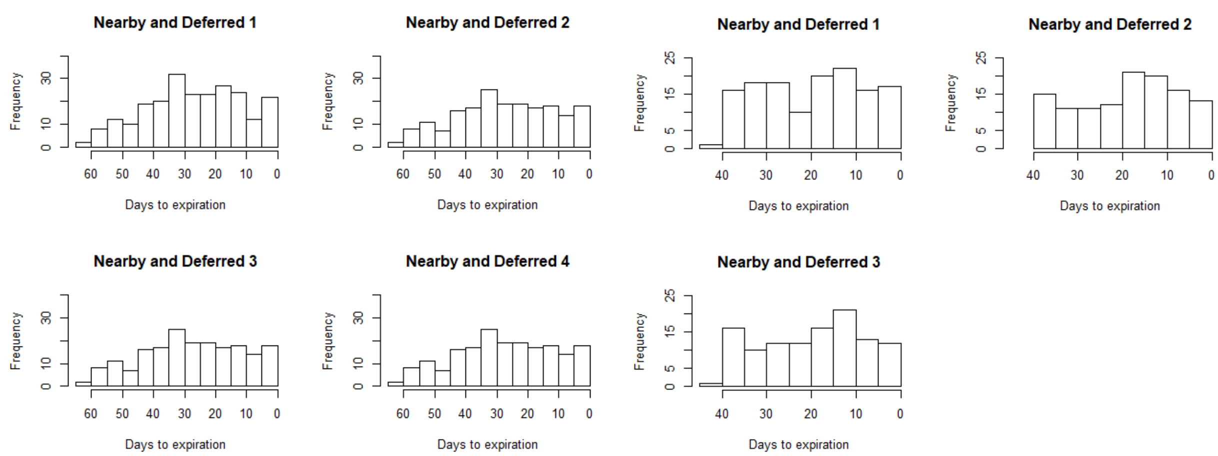
Supplementary Result 3



Corn Live cattle

Supplemental Figure 2.1 Histograms of Days when Intraday Nearby and Deferred Futures Prices were not Cointegrated, 2008-2015

Note: Histograms for corn and live cattle are in the left and right panels, respectively.

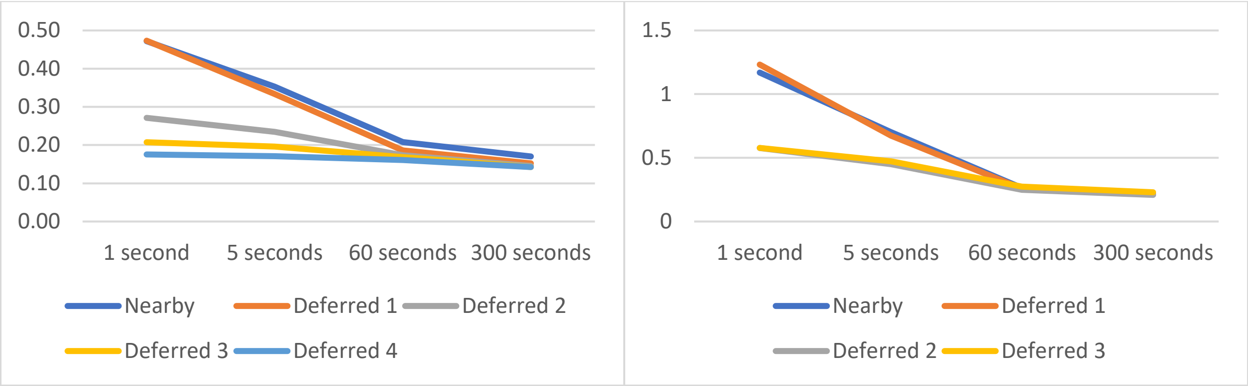


Corn Live cattle

Supplemental Figure 2.2 Histograms of Days when Intraday Nearby and Deferred Futures Prices were both Stationary, 2008-2015

Note: Histograms for corn and live cattle are in the left and right panels, respectively.

Supplementary Result 4



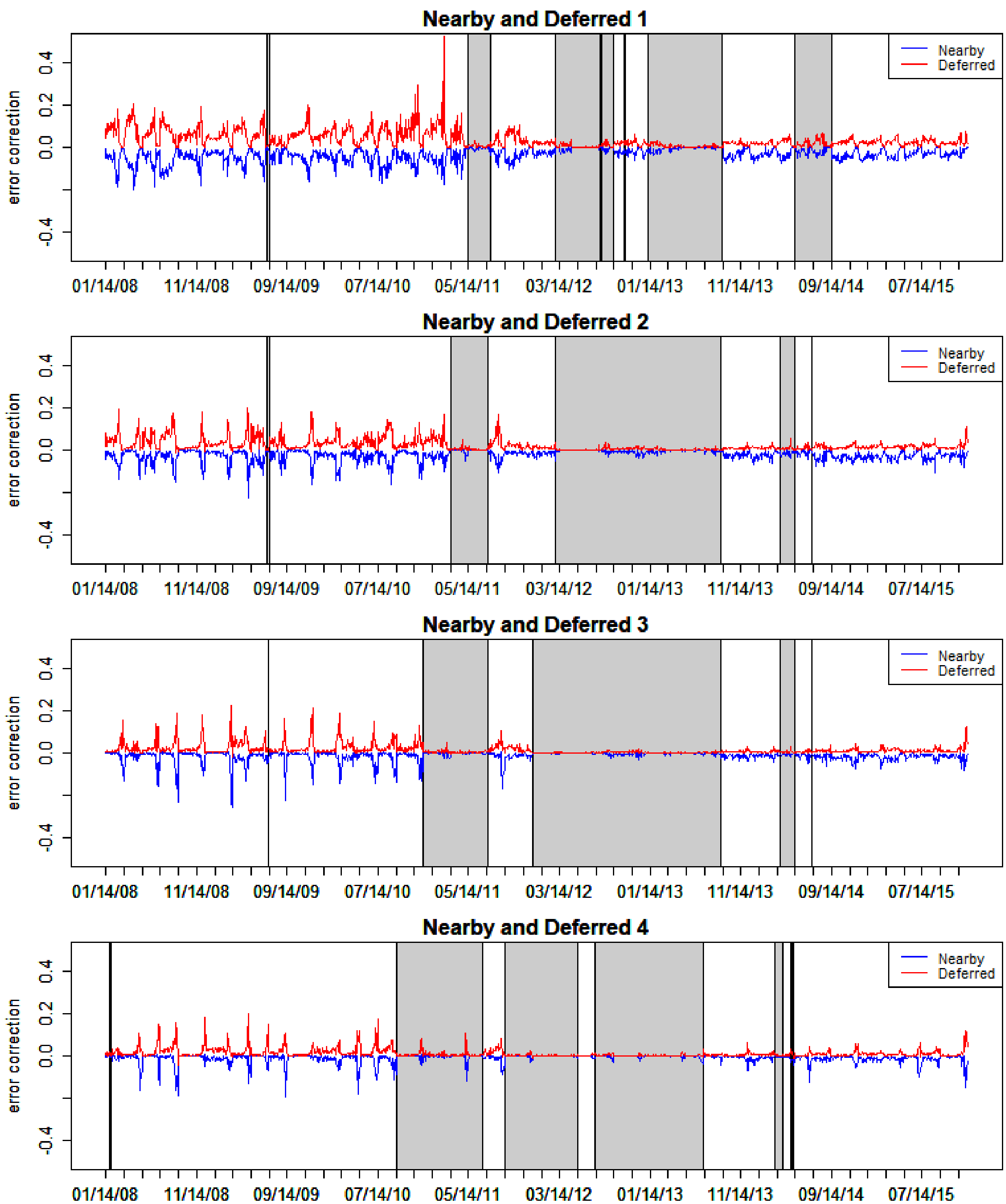
Supplemental Figure 2.3 Variance Signature Plots for Corn (left) and Live Cattle (right) Prices, 2008-2015

As explained in Hansen and Lunde (2005) and others, realized variance (RV) using higher sampling frequencies incorporate variance due to both changes in efficient price and the variance of noise, while the RV using low frequency is close to the variance of efficient price. Thus, RV using sampling frequency at higher frequencies can be used as a consistent estimator of noise (Bandi and Russell 2008), and the difference between RV using higher and lower sampling frequencies reflects the variance of noise. We calculate RV as the sum of non-overlapping squared intraday returns using *n* second sampling frequency for nearby and deferred contracts in corn and live cattle markets for each day during the whole sample period. We take *n* equals 1,5,60, and 300 seconds. The figure above shows daily average RV using different sampling frequencies for nearby and deferred contracts for both commodities. We find in both markets, nearby and deferred contracts have similar levels of efficient price variance, which can be proxied by the RV using 300 seconds sampling frequency (i.e., 0.15 for corn and 0.25 for live cattle). This evidence is consistent with the fact that prices for nearby and deferred contracts follow a common efficient trend in most days. However, the nearby contracts are associated with higher 1-sec RV and larger differences between RV using high (1 second) and low sampling

(500 seconds) frequencies. The differences between RV for nearby and first deferred contracts is small in both markets, consistent with the fact that differences between CS, IS and ILS are smaller between the nearby and first deferred contracts in both markets as shown in table 2 in the article.

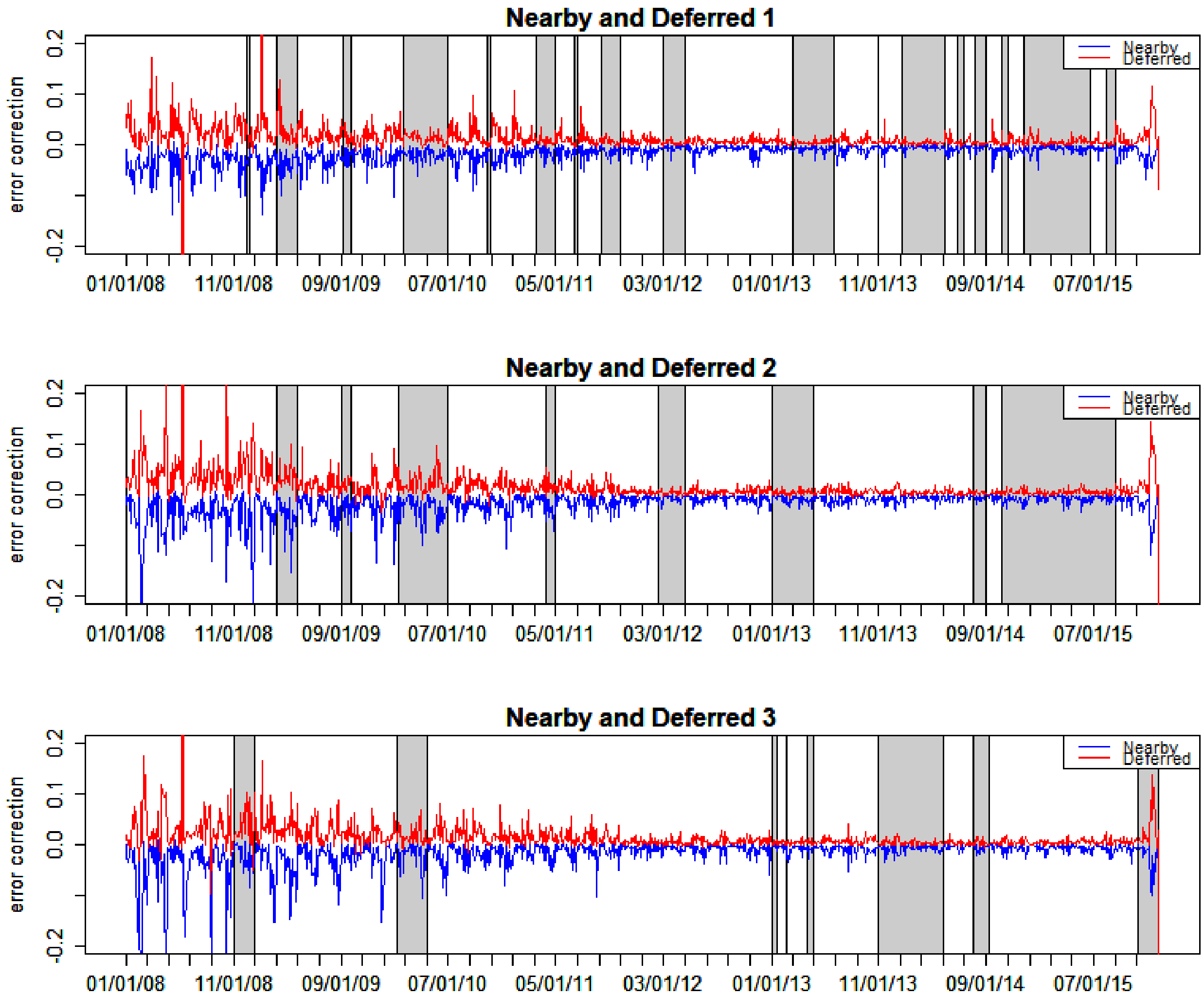
Supplementary Result 5

We present information on the error correction parameters below and find that a clear pattern emerges from supplemental Figures 2.4 and 2.5. After entering the nearby contract period, the magnitude of the error correction parameters peaks for both nearby and deferred contracts reflecting stronger adjustments to deviations from the intraday parity. As we approach the delivery period, adjustments are less strong, as reflected by a decline in the magnitude of the error correction parameters. Adjustments are larger during tumultuous periods (e.g., 2008-09) when market shocks are more likely to pull apart the two contract prices, requiring more active adjustment to maintain the equilibrium parity. Error correction coefficients are very small for inverted corn markets, reflecting the deterioration of the link between maturities during these periods. Weak exogeneity tests show that the null of weak exogeneity is usually rejected for both contract pairs.



Supplemental Figure 2.4 Error Correction Coefficients for the Corn Market, 2008-2015

Note: Shaded areas represent backwardation periods. Corn futures contracts expire on the business day prior to the 15th calendar day of the maturity month.



Supplemental Figure 2.5 Error Correction Coefficients for the Live Cattle Market, 2008-2015

Note: Live cattle futures contracts expire on the last business day of the maturity month.

Supplemental Table 2.5 Percentage of Days when Price is Exogenous in the VECM (Corn)

	Contract Pair 1	Contract Pair 2	Contract Pair 3	Contract Pair 4
Nearby	1%	5%	14%	22%
Deferred	4%	4%	3%	2%

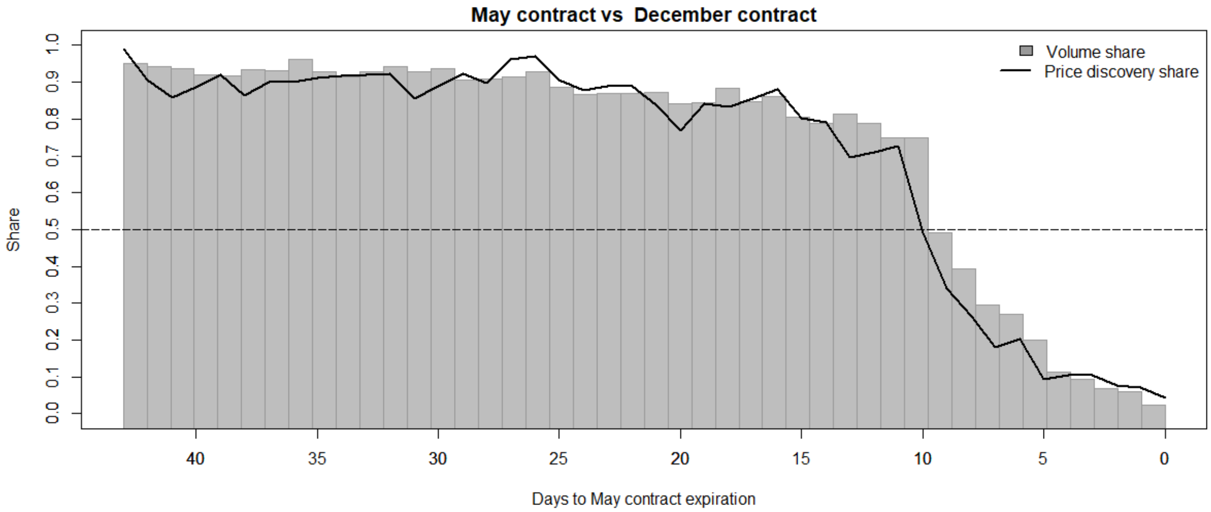
Note: these tests should be considered carefully due to the presence of microstructure noise.

Supplemental Table 2.6 Percentage of Days when Price is Exogenous in the VECM (Live Cattle)

	Contract pair 1	Contract pair 2	Contract pair 3
Nearby	16%	10%	9%
Deferred	6%	16%	24%

Note: these tests should be considered carefully due to the presence of microstructure noise.

Supplementary Result 6



Supplemental Figure 2.6 Price Discovery and Volume Shares for the May Contract, 2008-2015

CHAPTER 3:

ALGORITHMIC QUOTING, TRADING, AND MARKET QUALITY IN AGRICULTURAL COMMODITY FUTURES MARKETS

3.1 Introduction

"Traditional end-users— such as those from the agricultural community—are particularly concerned about the effects of automated trading on these markets. It is especially important for us to be able to respond to the concerns of those who are not so-called "flash boys," and are only moving at human speed."

Commodity Futures Trading Commission Former Chairman Timothy Massad[10]

Algorithmic trading (AT), which adopts computer programs to automatically monitor markets and implement trading strategies, has become common in commodity futures markets. According to Commodity Futures Trading Commission's (CFTC) studies by Haynes and Roberts (2015, 2017), in the period November 2014 to October 2016 the proportion of automated trading reached 49% in grain and oilseed markets, and 45.8% in livestock markets. Since the emergence of AT, there have been concerns about its effects on commodity futures market quality, particularly in the agricultural community.

Most research as well as policy concerns over AT have focused on whether it impairs pricing efficiency, increases volatility, and diminishes liquidity (SEC 2014). Studies find AT improves

---

[10] Statement by Chairman Timothy Massad regarding the approval of supplement proposal to automated trading regulation on November 4, 2016. https://www.cftc.gov/PressRoom/SpeechesTestimony /massadstatement110416

pricing efficiency (Conrad, Wahal and Xiang, 2015; Chaboud et al., 2014; Brogaard, Hendershott and Riordan, 2014; Carrion, 2013) and liquidity (Hendershott, Jones and Menkveld, 2011; Hasbrouck and Saar, 2013; Conrad, Wahal and Xiang, 2015), but have not reached consensus on how it influences market volatility. For instance, Hasbrouck and Saar (2013) show low latency trading reduces short-term volatility and Hagströmer and Nordén (2013) find passive market making algorithms mitigate intraday market volatility. More recently, Brogaard et al. (2018) focus on high frequency trading (HFT) and extreme price movements in the security market and find no evidence of HFT causing extreme price movements. However, a number of studies conclude AT is associated with increasing price volatility (Boehmer, Fong and Wu, 2012; Zhang 2010; Scholtus, Dijk and Frijns, 2014). None of this research has examined agricultural futures markets.

In recent years, a growing number of studies on the microstructure of agricultural commodity futures markets have emerged. Wang, Garcia and Irwin (2013) document that liquidity costs are lower in the transition to electronic trading in the corn futures. Couleau, Serra and Garcia (2018) investigate microstructure noises in the live cattle futures market and find no evidence that high frequency trading (HFT) is responsible for any economically meaningful market noise. Adjemian and Irwin (2018) investigate the USDA announcement effects in real time. They find that the average size per trade just after announcement time has fallen while the bid-ask spread has increased in the real-time era, indicating the possibility that high frequency traders increase trading costs in the short window following the report. While these studies document important liquidity, noise, and volatility measures in periods associated with different levels of AT (or HFT), they do not directly identify effects of AT on these market quality measures.

Motivated by the recent policy concerns and the limited information on AT behavior in agricultural commodity markets, this paper aims to identify the effects of AT on pricing efficiency, volatility and liquidity in the corn, soybean, and live cattle futures markets. Identifying the effects of AT is complicated because regulatory agencies such as the CFTC have not clearly defined AT, and public or regulatory data are not readily available. Even when available, data often do not contain the relevant measures. For instance, the CFTC's dataset contains transaction-level data, but evidence suggests that AT's effects are more related to the changes in the supply of quotations rather than trades (e.g., Conrad, Wahal and Xiang, 2015; Hendershott, Jones and Menkveld, 2011; Hasbrouck and Saar, 2013; Hasbrouck 2018).

This paper focuses on the effects of quotations rather than trades generated by trading algorithms. The main advantage of using algorithms is that there is virtually zero marginal cost to monitor the market and adjust quotes, therefore the effects of AT are more likely to be revealed in quotes (Hendershott, Jones and Menkveld, 2011; Conrad, Wahal and Xiang, 2015). In addition, regulatory agencies appear to focus on algorithmic quoting (AQ) rather than trading. Recently, several AQ practices in commodity futures markets have caused concerns on whether they impede the market efficiency and generate order-execution risks. For instance, the CME launched the Messaging Efficiency Program (MEP) aiming to restrict inefficient messaging such as massive order cancellations and nonmarketable order submissions for corn and soybean futures in 2013 and 2016 for the live cattle futures. The CFTC has also taken several enforcement actions against proprietary traders who use computer algorithms for spoofing, which is based on generating a large number of quotes. For example, on January 31, 2019 the CFTC charged a trader at a proprietary trading firm in Chicago for engaging in spoofing the

CME's soybean futures between August 2013 and June 2016.[11] Hasbrouck (2018), Li, Wang and Ye (2017) among many others, note that high speed changes in quote updates might cause order-execution risks, with slow traders losing to fast traders who are able to adversely select favorable orders to trade. Thus, traditional futures users such as farmers and ranchers who are not fast traders may lose to fast algorithmic traders and not be able to effectively hedge their business risks.

We follow Conrad, Wahal and Xiang (2015) and Hendershott, Jones and Menkveld (2011) who use the rate of electronic message traffic to measure AQ. Message traffic includes quote updates in the limit order book (LOB) caused by order submissions, cancelations, and trades. Following Hendershott, Jones and Menkveld (2011), we normalize quote updates by trade volume; the variation in this measure reflects the intensity of quote updates relative to trade volume. We examine the effect of AQ on different market quality measures in each market separately. AQ and market quality measures such as liquidity and volatility are simultaneously determined. To analyze the causal effects of AQ on different market quality measures, we follow Chaboud et al. (2014) and use a heteroskedasticity-based identification approach.

The results demonstrate that more intensive AQ is associated with more efficient prices and lower short-term volatility. AQ also narrows effective spreads, i.e. order execution costs, particularly in the more liquid corn and soybean markets. Lower effective spreads are a result of reduced adverse selection costs. Also, there is evidence that AQ significantly increases realized spreads, which represent liquidity provider revenues, in the corn futures market. Such evidence indicates that currently algorithmic liquidity providers have a competitive advantage in the market, which is likely because of their trading speed. No evidence emerges that more intensive

[11] https://www.cftc.gov/PressRoom/PressReleases/7865-19

AQ is associated with increased realized spreads in the soybean and live cattle markets. These results provide important implications on current policy debates on algorithmic trading in agricultural futures markets.

3.2 Data and Measurement

We use CME's market depth data for the period of November 20, 2015 to May 14, 2017 for three agricultural commodities: corn, soybeans, and live cattle. The sample period was chosen because CME's Market Data Platform (MDP) 3.0, which provides more accurate and detailed market depth data than previous versions of the dataset, is available beginning in November 2015.[12] The data are detailed LOB updates that are time-stamped to the nanosecond. Updates to the LOB occur as orders are (1) added, (2) deleted from the book due to cancellation or execution (3) or changed. The data also provide detailed trade information including the price, size, time, sequence, and direction of each trade.

In processing the LOB data, we only use quotes and trades in the outright book, while implied quotes and trades from the implied book (spread trading) are excluded from the analysis. Although spread trades also consume and supply liquidity to the outright book, algorithms used in the spread book are more likely to be intended to impact price spreads rather than price levels. Besides, implied trades are initiated in the spread book and do not have trade initiators in the LOB data that are required for calculating realized spreads and adverse selection costs used in

[12] Before, the CME Globex FIX format was used. The MDP data compared to the FIX data, provide nanosecond resolution and more accurate trade size information. These improvements help with reducing measurement errors when computing market quality measures, particularly for order execution cost measures.

our analysis.[13] On each day, we choose the contract with the highest message traffic because AQ is more prevalent in contracts that are more active.

3.2.1 AQ Measure

As alluded to earlier, our data do not allow us to observe whether an update in the LOB is generated by a computer algorithm. Following the literature (e.g., Conrad, Wahal and Xiang, 2015; Hendershott, Jones and Menkveld, 2011), we build the AQ measure using the rate of electronic message traffic which captures updates in the outright LOB per traded contract. This proxy is used by researchers and market participants based on the observed correlation between the increase in message traffic and the fast growth of AT. An AT strategy can submit and cancel orders multiple times before the transaction is executed or "slice and dice" a large order into multiple small orders, therefore generating more messages. However, an increase in message traffic can result from an increase in trading rather than the change in the trading practice. Hence, we normalize the raw number of quote updates by the dollar trading volume, and AQ_t is calculated as,

$$AQ_t = \frac{Message_t}{DollarVolume_t}, \tag{3.1}$$

where $Message_t$ is the number of updates in the LOB and $DollarVolume_t$ is the total dollar volume over the time interval t, respectively. Hence, AQ_t is a proxy for the intensity of AQ

[13] The Lee and Ready (1991) rule cannot be used for inferring trade directions for implied trades as they are initiated in the spread book. Additionally, realized spreads and adverse selection costs can be miscalculated as liquidity provider revenues and losses are not realized in the outright book.

relative to trading volume.[14] A high AQ_t indicates that quote updates are frequent compared to the volume of trade.

We measure AQ over 10-minute intervals for corn and soybeans, and 25-minute intervals for live cattle. The lengths of the measurement intervals are selected based on the trade-off between two purposes. First, the measurement interval t needs to be long enough to capture enough AQ activity. Second, t needs to be smaller than a day to capture possible intraday patterns in AQ activity, which will help in the statistical identification, and also increase the statistical power in the regression analysis. For corn and soybeans, we use day-time trading hours (8:30 a.m. – 1:20 p.m. CT), so each day has 29 10-minute intervals. In total, there are 11,020 measurement intervals for corn and soybeans. The live cattle futures market changed trading hours to 8:30 a.m. - 1:05 p.m. CT on February 29, 2016. Before, the live cattle futures market used different trading hours during the week.[15] For the period after February 29, 2016, we use the whole trading session and each day has 11 25-minute intervals. For the period prior to the change in trading hours, since prices are typically stale after 1:00 p.m. we use the first 9 and 12 25-minute intervals for Monday and Tuesday to Friday, respectively, i.e. only including time intervals before 1:05 p.m. In total, we have 4,209 observations for live cattle. The same measurement intervals are also used for market quality measures described below. In the robustness check section, we show

[14] Hendershott, Jones and Menkveld (2011) interpret AQ as a proxy for the amount of AT taking place in the market. They argue this measure essentially captures changes in liquidity supply caused by trading algorithms. The measure AQ_t is also similar to the message-to-fill ratio used in the CME's Messaging Efficiency Program for inferring the degree of algorithmic quoting's impact.

[15] 9:05 to 16:00 for Monday. 8:00 to 16:00 for Tuesday to Thursday. 8:00 to 13:55 for Friday.

that our conclusions remain unchanged when using different sample selections and subsample periods.

The number of quote updates, dollar volume, and *AQ* for the corn, soybean and live cattle markets are presented in figures 3.1-3.3, respectively. For most of the time, the AQ measure is relatively stable in the three markets, with no visible upward or downward trend, which is consistent with Haynes and Roberts' (2017) evidence that the proportion of automated trades in these markets does not change much during the sample period. However, large spikes of *AQ* are present in all the three markets, with more spikes in the corn and soybean markets and much fewer in the live cattle market. In the corn market, large spikes are clustered at the end of the sample period, the first half of the year 2017, when trading was light due to high South American corn yields and limited surprises in the USDA reports (Hubbs, 2017a,b). In the soybean market, large spikes of AQ are found when the low-volume August contract is the nearby contract, which is a typical seasonal pattern in the soybean market.

Massive quote updates are more likely generated by algorithms rather than humans when trade volume is small. For example, AQ reaches a high of 53 during the 13:00-13:10 period on Jan 26, 2016 in the corn futures market, with 18 trades but as many as 7,127 quote updates. In particular, we find 1,065 quote updates happened within a second (13:14:49) during this ten-minute interval. More interestingly, some of the largest spikes are associated with "strategic runs" as defined in Hasbrouck and Saar (2013), which are linked order submissions and cancelations that are likely to be part of an algorithm, particularly in the corn market. In supplementary result 1, we show an example of a "strategic run" that generates 135 messages in the LOB within just 41 milliseconds in the corn market.

It is not clear when and why “strategic runs” as well as other types of algorithmic quoting became more active and generated massive messages in these periods, since algorithmic trading strategies are confidential and predictions from theoretical models typically only focus on one type of algorithm and depend on restrictive (and sometimes unrealistic) assumptions (Hasbrouck and Saar, 2013; Li, Wang and Ye, 2017; Biais, Declerck and Moinas, 2016). Future research may want to explore the reasons for active algorithmic quoting during periods of low trade volume.

3.2.2 Pricing Efficiency Measure

Variance ratios are commonly used in empirical market microstructure studies to examine pricing efficiency (e.g., Bessembinder 2003; Conrad, Wahal and Xiang, 2015). Typically, the variance ratio compares the variance of price returns at two different time scales. This measure assumes that price changes in a long horizon are dominated by movements in the fundamental value, while price changes in a short horizon are more affected by noise. Consistent with the measurement interval for AQ, we calculate a variance ratio over each 10 minutes for corn and soybeans, and 25 minutes for live cattle. We divide each measurement interval into n equally spaced non-overlapping long intervals and q equally spaced non-overlapping short intervals in each long interval. Both n and q need to be integers greater than 1. Since microstructure noise is generally more important at a shorter scale (Charles and Darné, 2009), we chose the short interval to be 500ms therefore q=1,200 for grains and 3,000 for live cattle. To show variance ratios are robust to the selection of q, we also present results using 1-second short intervals. The longer interval needs to be long enough to be dominated by changes in fundamentals. As shown in Hu et al. (2017) and Couleau, Serra and Garcia (2018), price variance is dominated by fundamentals when using 5-minutes sampling intervals in corn and live cattle markets. As a result, we choose n to be two for corn and five for live cattle, and each long interval therefore is 5 minutes. Following specifications in Lo and MacKinlay (1989), we define X_k, where $k = 0,1,2 \ldots nq$, as the log quote midpoints process. The mean drift in prices, $\hat{\mu}$, is as follows,

$$\hat{\mu} = \frac{1}{nq}\sum_{k=1}^{nq}(X_k - X_{k-1}) = \frac{1}{nq}\left(X_{nq} - X_0\right). \tag{3.2}$$

The variance of non-overlapping short interval (s) price differences with bias correction then is

$$\hat{\sigma}_s^2 = \frac{1}{nq-1}\sum_{k=1}^{nq}(X_k - X_{k-1} - \hat{\mu})^2, \tag{3.3}$$

and the bias-corrected variance of $nq - q - 1$ overlapping long interval price differences (l) is

$$\hat{\sigma}_l^2 = \frac{1}{q(nq-q+1)\left(1-\frac{q}{nq}\right)}\sum_{k=q}^{nq}\left(X_k - X_{k-q} - q\hat{\mu}\right)^2. \tag{3.4}$$

The variance ratio for the time interval t is calculated as follows:

$$VR_t = \frac{\hat{\sigma}_l^2}{\hat{\sigma}_s^2}. \tag{3.5}$$

A benchmark variance ratio of one, is consistent with a random walk pricing process that is considered as the weak-form pricing efficiency (Campbell, Lo and MacKinlay, 1997). Intuitively, the smaller the noise in price, the closer $\hat{\sigma}_s^2$ is to $\hat{\sigma}_l^2$, and VR_t gets closer to one. In contrast, if microstructure noise is large, $\hat{\sigma}_s^2$ is larger than $\hat{\sigma}_l^2$ and the value of VR_t moves away from 1. Following Conrad, Wahal and Xiang (2015), hence, we use the absolute value of VR_t's deviation from 1, i.e. $|VR_t - 1|$ as the measure for pricing efficiency, where a smaller value of $|VR_t - 1|$ indicates more efficient prices. We exclude the intervals with less than 30 non-zero price differences to maximize the power of the statistical analysis.

3.2.3 Volatility Measure

The debate on whether AT increases volatility has been concentrated on whether certain AT practices cause extreme price changes (Kirilenko et al., 2017; Brogaard et al., 2018). Following, Hasbrouck and Saar (2013), we employ the high-low price range which has the benefit of capturing the maximum price movement in a short window as the measure for volatility. The volatility measure $HighLow$ is expressed in basis points (bps) and is defined as

$$HighLow_t = \frac{high_t - low_t}{mid_t}, \tag{3.6}$$

where $high$ and low, are the highest and lowest midquote in the time interval t, respectively. The denominator mid_t is the midpoint between $high_t$ and low_t. For consistency with the other AQ and market quality measures, t is set to be 10 minutes for corn and soybean, and 25 minutes for live cattle.

3.2.4 Liquidity and Order Execution Cost Measures

We use the effective half spread (es) to measure liquidity. For the i^{th} trade, the proportional effective half-spread, es_i is defined as

$$es_i = \frac{I_i(p_i - m_i)}{m_i}, \tag{3.7}$$

where I_i equals 1 for buyer-initiated trades and -1 for seller-initiated trades, p_i is the transaction price, and m_i is the prevailing quote midpoint calculated from the bid and ask quotes. Effective spread can be decomposed into two components: realized spread and price impact. Realized spread (rs) can be calculated as

$$rs_i = \frac{I_i(p_i - m_{i+\tau})}{m_i}, \tag{3.8}$$

where $m_{i+\tau}$ is the quote midpoint, τ periods after the transaction. The realized spread measures the revenue liquidity providers receive. The time horizon for τ needs to be short as it is intended to reflect the horizon in which liquidity providers close their position. Research has traditionally used 5 minutes for τ. However, given the frequency of quoting and trading under the current market environment, 5 minutes may be too long. Recently, O'Hara (2015) suggests using 5 or 15 seconds for τ, and Conrad, Wahal and Xiang (2015) suggest τ less than 20 seconds. We calculate realized spreads using τ equals 5 and 10 seconds, and present results for both. The adverse selection cost (ad) or price impact, which measures losses that liquidity providers incur to better informed traders, can be computed as

$$ad_i = \frac{I_i(m_{i+\tau} - m_i)}{m_i}. \tag{3.9}$$

Consistent with other measures, we calculate volume weighted averages for each measurement interval t and express es, rs and ad in bps.

Table 3.1 presents mean values for AQ and market quality measures calculated over t intervals in the sample period. The measures all have a high standard deviation as they can vary considerably within the day and under different market conditions. The live cattle market has a lower degree of pricing efficiency than the corn and soybean markets, which is expected as the live cattle market is a much thinner market than the other two markets. However, one needs to be cautious when comparing measures across markets that are normalized by dollar volume, namely, AQ and execution cost measures. The live cattle market can have a high value of *AQ* simply because it has a lower denominator (the average price for live cattle during the sample period was 112 cents/pound versus 336 cents/bushel for corn and 983 cents/bushel for soybeans). Hence, a higher *AQ* for live cattle cannot be interpreted as the live cattle market being associated to a higher level of AQ than the other two markets.

Similarly, since grain and live cattle markets use different price quotations and tick sizes, one cannot infer their relative liquidity levels by comparing the levels of their effective spreads. The average effective spreads are 3.5 basis points for corn and 1.5 basis points for soybeans. Using a simple calculation, we obtain the approximate average effective spread that equals 0.128 cent/bushel (3.5 basis points × 336 cents/bushel) for corn, and 0.147 cents/bushel (1.5 basis points × 983 cents/bushel) for soybeans, with both measures close to half of their tick size (0.125 cents). These findings are consistent with Wang, Garcia and Irwin (2013) who reveal that the average bid-ask spread in the electronically traded corn market is close to one tick size. By applying the same calculation, we get the average effective half spread equals 0.021 cents/pound (1.8 basis points × 112 cents/pound) for live cattle, which is close to one tick size (0.025 cents). Thus, these results indicate that the live cattle market is a less liquid market relative to the other two markets. In contrast to corn and soybean markets, realized spreads which represent returns to liquidity supplier are negative on average in the live cattle market. This reflects that the live cattle market is a less liquid market and liquidity supply is less profitable than in the other two markets. Our results for live cattle are consistent with Frank and Garcia (2011) assessment of live cattle BAS.

3.3 Model Specification

We identify more clearly the effects of AQ on different market quality measures using regression analysis. The selection of explanatory variables is based on research findings on the determinants of market quality (e.g., Wang, Garcia and Irwin, 2013; Hendershott, Jones and Menkveld, 2011; Chaboud et al., 2014) and is constructed as follows:

$$MQ_t = \beta_0 + \beta_1 AQ_t + \beta_3 MQ_{t-1} + \beta_4 trend_{t-19:t-1} + \beta_5 USDA_t$$
$$+\beta_6 IndexRoll_t + \boldsymbol{D}'_t\boldsymbol{\gamma} + \boldsymbol{C}'_t\boldsymbol{\delta} + e_t, \qquad (3.10)$$

where t = 1,…,T indexes 10-minute or 25-minute time intervals. Following the literature, we examine the influence of AQ on each market quality variable separately and MQ represents one of the five market quality measures described above. The variable MQ_{t-1} is the one lag term of MQ and the variable $trend_{t-19:t-1}$ is the sum of returns over the last 20 measurement intervals. These two variables aim to capture the lagged effect of MQ and informational shocks in the market, and by construction they are predetermined. The dummies $USDA_t$ and $IndexRoll_t$ control for the effects of two exogenous events, USDA reports and commodity index rolling. Adjemian and Irwin (2018) find the effect of USDA announcements dissipates within a few trading minutes after the real-time releases. Hence, the dummy variable $USDA$ equals one for the first ten-minute interval after USDA WASDE, Crop Production, and Grain Stocks announcements for corn and soybean, and zero otherwise. For live cattle, we use the Cattle on Feed report. Because Cattle on Feed reports are released after trading hours, the dummy variable $USDA$ for live cattle equals one for the first 25-minute interval on days following the report. The dummy variable $IndexRoll$ equals one for the ten-minute intervals between the fifth and tenth business days of the month preceding the expiration month, which includes the roll periods of the two largest commodity indices: S&P - Goldman Sachs and Dow Jones - UBS commodity indices. Notice that the index rolling dummy may also capture part of the effect of agency order execution algorithms.[16] The vector $\boldsymbol{D}$ is a set of dummy variables ($\boldsymbol{D}_t = (Open_t, Close_t, Mon_t, Tue_t, Wed_t, Thu_t)$**)** that controls for market opening, closing and the day-of-the-week effect. To control for market opening and closing effects, the dummies $Open_t$, and $Close_t$ equal one for the first (except for intervals following cattle on feed reports in the live cattle market) and last measurement intervals of the day, respectively. Vector $\boldsymbol{C}$ is a set of contract month dummies.

[16] Agency algorithms are widely used by "buy-side" institutions to minimize trading costs when executing large orders for portfolio rebalancing (O'Hara， 2015). Since numerous positions need to be liquidated during the short rolling window (typically 5 days), commodity index funds managed by major financial institutions depend on automated trading algorithms rather than manual trades.

3.4 Identification Strategy

AQ and market quality measures are likely to be simultaneously determined. For example, algorithmic traders would be more willing to trade in a more liquid and efficient market because it is easier for them to manage risk in such environment (Conrad, Wahal and Xiang, 2015). Previous studies typically rely on an exogenous event such as an update in the trading system (e.g., decimalization, auto-quote, and change in settlement rule) which facilitates identification of an AQ effect. However, such exogenous events do not exist in the markets we study during the sample period.[17]

Recent developments in identification through heteroskedasticity (e.g., Rigobon, 2003; Lewbel, 2012) provide a solution when typical sources of identification such as imposing parameters based on economic intuition and instrumental variables are not available. Chaboud et al. (2014) adopt Rigobon's (2003) identification through heteroskedasticity approach to identify the impact of AT in foreign exchange markets. This identification approach is based on the heteroskedasticity of the structural shocks to identify simultaneous-equation systems. However, Rigobon's (2003) approach requires that data can be split into two subsamples with the structural shocks having different degrees of variance. Since our data sample does not fulfill this requisite, we instead use Lewbel's (2012) heteroskedasticity approach which does not depend on the assumption that the variances of structural shocks have two regimes.

---

[17] Corn and soybean futures were added to the CME's messaging efficient program before the sample period. We have tried using the messaging efficiency program as an instrument for identifying the effects of AQ on market quality measures in the live cattle market. However, Stock-Yogo test results suggest the messaging efficient program is a weak instrument. Intuitively, the market efficiency program provides an upper limit to AQ rather than causing a dramatic reduction in the level of AQ.

As explained in Lewbel (2012), the identification can be achieved if there exists a vector of explanatory exogenous variables $\boldsymbol{Z}$ and error terms are heteroskedastic. Variables in $\boldsymbol{Z}$ could equal or be a subset of the exogenous variables in the regression. Excellent candidates for $\boldsymbol{Z}$ are variables that are clearly exogenous. The variables, except for AQ are all good candidates for $\boldsymbol{Z}$, as they are either predetermined lagged terms or exogenous events. Lewbel's (2012) approach requires a two-stage estimator. In the first stage, the endogenous variable AQ is regressed on all the control variables including the variables in $\boldsymbol{Z}$. Specifically, we run the following regression:

$$AQ_t = \alpha + \boldsymbol{X}'\boldsymbol{\beta} + \boldsymbol{Z}'\boldsymbol{\gamma} + \varepsilon_t, \tag{3.11}$$

Where $\boldsymbol{X}$ is a vector of exogenous variables in equation (3.10) that are not in vector $\boldsymbol{Z}$ and the other terms are as previously defined. The Lewbel's (2012) approach requires that ε_t is heteroskedastic which is a common feature in models of endogeneity. As suggested by Lewbel (2012), we use the Breusch and Pagan (1979) test for heteroskedasticity and results suggest that the null of homoscedasticity can be rejected with a *p*-value less than 0.01 in all cases.

As shown in Lewbel (2012) the identification can be achieved using the standard Generalized Method of Moments-Instrumental Variable (GMM-IV) estimator and $(\boldsymbol{Z} - \overline{\boldsymbol{Z}})\varepsilon_t$ as an instrument, where $\overline{\boldsymbol{Z}}$ is a vector of means of the variables in $\boldsymbol{Z}$. The exogeneity of $\boldsymbol{Z}$ guarantees that $(\boldsymbol{Z} - \overline{\boldsymbol{Z}})\varepsilon_t$ is not correlated with the error term e_t in equation (3.10). The restriction of heteroskedasticity, i.e. cov($\boldsymbol{Z}$,${\varepsilon_t}^2$)$\neq 0$, guarantees that $(\boldsymbol{Z} - \overline{\boldsymbol{Z}})\varepsilon_t$ is correlated with ε_t in equation (3.11) and therefore with the endogenous variable AQ. The degree of correlation between $(\boldsymbol{Z} - \overline{\boldsymbol{Z}})\varepsilon_t$ and AQ depends on the degree of heteroskedasticity of ε_t with respect to $\boldsymbol{Z}$. A low degree of cov($\boldsymbol{Z}$,${\varepsilon_t}^2$) can cause $(\boldsymbol{Z} - \overline{\boldsymbol{Z}})\varepsilon_t$ to be a weak instrument. Since then approach uses a standard GMM IV estimator, Stock and Yogo (2005) test can be used to test whether $(\boldsymbol{Z} - \overline{\boldsymbol{Z}})\varepsilon_t$ is a weak instrument. In addition, when multiple variables are included in $\boldsymbol{Z}$, Hansen's (1982) test can also be employed to test for overidentification

We use a forward stepwise procedure to select the variables included in $\boldsymbol{Z}$ for each market and regression. First, we chose one exogenous regressor at a time and add it to the current vector $\boldsymbol{Z}$ until the specification passes the weak IV test and the overidentification test (if more than one variable is used for $\boldsymbol{Z}$**)**. If not, we add two variables, and then three and so on, until the model is correctly specified for all equations and markets. Using this procedure, we select $trend_{t-19:t-1}$ for $\boldsymbol{Z}$ for corn and live cattle markets, and ($trend_{t-19:t-1}$,$Open_t, Close_t$**)** for $\boldsymbol{Z}$ for the soybean market[18]. In all cases, Stock-Yogo tests statistics suggest ($\boldsymbol{Z} - \overline{\boldsymbol{Z}})\varepsilon_t$ has enough correlation with AQ using the Stock and Yogo (2005) critical values. For the soybean market, where multiple variables are included in $\boldsymbol{Z}$, Hansen's *J* statistics indicate instrumental variables are exogenous in all cases. These test statistics are presented in the last two rows in tables 3.3 to table 3.5.

3.5 Regression Results

Before discussing the regression results, we conduct a preliminary exploratory analysis of the relationship between AQ and market quality based on the descriptive statistics. We divide the sample period into low (AQ_{Low}), medium (AQ_{Medium}) and high (AQ_{High}) AQ periods with equivalent sample size, and present summary statistics for the market quality measures for each subsample in table 3.2. With very few exceptions, sample periods with higher AQs are associated with a variance structure closer to the efficient random walk benchmark, lower short-term volatility, narrower effective spreads, and smaller adverse selection costs. The results for the realized spread are mixed across markets. The average realized spread increases

[18] As shown in Lewbel (2012), the selection of variables only affects the efficiency but not the consistency.

monotonically from AQ_{Low} to AQ_{High} in the corn market but does not show a consistent increasing or decreasing pattern in the soybean and live cattle markets. This may indicate that informed traders are losing more to AQ liquidity providers in the corn futures market. However, the cross-sectional comparisons do not provide causal relationships. To shed light on the causal relationships between the AQ and market quality, we present the regression results in tables from 3.3 to 3.5. To save space, control variables for the day-of-the-week and contract effects are omitted from the tables. Results for these dummy variables reflect market conditions in different time periods during the December 2015 to May 2017 period and are presented in supplementary result 2.

3.5.1 AQ and Pricing Efficiency

Table 3.3 reports coefficients from the pricing efficiency equations. As explained, pricing efficiency is measured using the deviations in variance ratios from the pricing efficiency benchmark value of one ($|VR_t - 1|$), using two different short sampling intervals, 500 milliseconds and 1 second. Price trend, USDA reports and weekday dummies have limited influence on the efficiency of prices in the three markets. Significant contract effects on pricing efficiency are found in the soybean and live cattle markets, but not in the corn market (supplementary result 2). In particular, more efficient prices are in contract months with lower volatility in the soybean and live cattle markets.

The results show a clear intraday pattern where pricing efficiency tends to be higher (lower $|VR_t - 1|$) at the open and significantly lower (higher $|VR_t - 1|$) at the close in all the three markets. As less news is coming into the market at the end of the trading day and intraday traders liquidate their positions, trades near the market close are more likely to be noise trades. Commodity index rolling, which initiates trades for portfolio management rather than based on fundamentals, is found to improve pricing efficiency, albeit its effect is not consistently significant across markets. The “sunshine trading effect” of commodity index rolling on liquidity has been widely captured in agricultural commodity markets. Our findings suggest that the increase in liquidity providers during these periods more than offsets any negative effects of index rolling practices on pricing efficiency. Whether more informed traders also participate in the pricing process at these moments is less clear but would be consistent with Hu et al. (2017) who find commodity index rolling improves price discovery in nearby contracts in corn and live cattle markets.

AQ significantly improves the efficiency of prices at the 10% level in the corn market but has no significant effect in the soybean market. As shown in table 3.1, a one standard deviation change in the AQ measure is 1.338 in the corn futures market. Hence, the parameter representing the variance ratio using 500 ms (1 second) short intervals in the corn futures market suggests that a one standard deviation increase in AQ narrows the deviation of the variance ratio from 1 by 1.338bps × 0.096 (0.138) = 0.128bps (0.184bps), representing a 24% (34%) decline from the mean pricing efficiency measure of 0.53bps (0.54bps) for corn. In the live cattle market, results indicate that AQ significantly improves the efficiency of prices using 1-second short sampling intervals at the 10% level.

In recent years, traditional agricultural commercials have complained that algorithmic traders cause extreme price volatility that does not reflect changes in fundamentals.[19] By focusing on price variance behavior in recent years, Couleau, Serra and Garcia (2018) find no evidence that high frequency trading is responsible for economically meaningful market noise in the live cattle futures market. Our results suggest that AQ decreases the degree of market noise. A possible

---

[19] See https://www.reuters.com/article/usa-cattle-markets-idUSL2N15003S for an example.

explanation is that trading algorithms can quickly respond to market news and other information, reducing the staleness of quotes and reflecting the fundamental information faster. These results are also consistent with previous evidence found in equity markets (Conrad, Wahal and Xiang 2015; Chaboud et al., 2014; Brogaard, Hendershott and Riordan, 2014; Carrion, 2013).

3.5.2 AQ and Volatility

Volatility, measured by the high-low price range, is heavily affected by USDA reports in all the three markets (table 3.4). Open and close dummies capture the widely documented U shape pattern of intraday volatility. The one-period lagged volatility and contract dummies show significant impacts in all the three markets as well, while commodity index rolling and weekday dummies have limited effects (supplementary result 2).

After controlling for other factors, the results show that, on average, AQ significantly reduces short-term volatility in all the three markets, particularly in the corn market. Using the standard deviations of the AQ measures presented in table 3.1 and estimated AQ coefficients in table 3.4, a one standard deviation increase in AQ reduces the high-low volatility by 1.149bps (1.338bps ×0.859), 0.529bps (0.553bps×0.956), and 1.620bps (1.904bps ×0.851) in the corn, soybean, and live cattle markets, respectively. Based on sample average prices, these can be translated into a 0.039 (366cents/bushel × 1.149bps), 0.052 (983cents/bushel × 0.529bps), and 0.018 (112cents/bushel × 1.620bps) cents/bushel decrease in the average high-low price range in the corn, soybean, and live cattle markets, respectively. These changes represent about 15% to 20% of a tick size in these markets. These results are consistent with Hasbrouck and Saar (2013) who find that proprietary algorithmic trading reduces short-term volatility.

3.5.3 AQ, Liquidity, and Order Execution Costs

Table 3.5 shows the effect of AQ on market liquidity. The effective spread is decomposed into a realized spread and an adverse selection cost component that are calculated using quote midpoint 5 and 10 seconds after the trade. IV regressions are estimated for the effective spread and each component. The results show that, in general, the effective spread is positively related to its own lagged term, USDA reports, and market open and close, while negatively related to price trend and commodity index rolling. These effects have expected signs and are consistent with the findings in Wang, Garcia and Irwin (2013). Contract effects also suggest order execution costs are higher in more volatility months (supplementary result 2).

As table 3.5 shows both realized spread and adverse selection cost are positively related to their lagged terms in all cases. Consistent with the literature, market trend and USDA report dummies, which represent time periods with volatile prices and informational shocks, significantly increase revenues to better informed traders and decrease returns to liquidity providers (Silber 1984; Shang, Mallory and Garcia, 2018; Hendershott, Jones and Menkveld, 2011). While commodity index rolling significantly reduces the effective spread in the corn futures market, it has limited influence on the realized spread and adverse selection cost in any of the markets. The realized spread and adverse selection cost in the corn and soybean markets have different intraday patterns than the live cattle market. In the corn and soybean markets, realized spreads (adverse selection costs) are significantly lower (higher) at the market open and higher (lower) at the close. This pattern is consistent with the intraday pattern of pricing efficiency presented previously, that market news is more concentrated at the market open where better informed traders earn larger profits and lose less to liquidity providers. In contrast, at the close, less news enters the market. Also, intraday traders may need liquidity to close their positions, yielding higher revenues to liquidity providers. In the live cattle market, market close and open have the same effects on the realized spread and adverse selection costs as in the other two markets, except that realized spreads are also significantly higher at market open. This is probably because trading activity in the live cattle market is highly concentrated in the morning and liquidity is limited in the afternoon (Shang, Mallory and Garcia, 2018). Hence liquidity providers earn higher profits at the market open compared to other periods of the day.

AQ does not have a significant effect on the live cattle market effective spread, nor on its realized spread and adverse selection cost components. This may be related to the lower presence of automated algorithms in the thinly traded live cattle market during the sample period (Haynes and Roberts, 2015, 2017). However, in the corn and soybean markets, AQ significantly reduces effective spreads. A one standard deviation increase in the AQ measure is associated with a 0.024bps (1.338×0.018bps) and a 0.034bps (0.553×0.062bps) decrease in effective spreads in the corn and soybean markets, respectively. These represent a 0.7% and a 2% decrease from the mean effective spread of 3.5bps and 1.5bps in the corn and soybean markets, respectively. Considering mean effective spreads in the corn and soybean markets are close to their minimum, i.e. half of a tick size, it is not surprising to find small percent reductions in effective spreads associated with a higher level of AQ. Our results are compatible with the low quoted spreads in the corn futures market reported by Wang, Garcia and Irwin (2013).

By decomposing the effective spread into the realized spread and adverse selection, we are able to identify the sources behind the decreased immediacy costs (effective spreads) in the presence of more intensive AQ. As discussed above, narrower effective spreads indicate either less revenues to liquidity providers (smaller realized spreads), smaller losses to better-informed traders (smaller adverse selection costs), or both. In the corn futures market, on average, AQ significantly increases realized spreads but also significantly reduces adverse selection costs by a larger magnitude. In the soybean market, AQ does not have a significant effect on realized spreads, but significantly reduces adverse selection costs when the 10-second horizon is used. These results suggest that the improvement in liquidity associated with higher AQ activity can be mainly attributed to reduced adverse selection costs in these markets, as liquidity providers are losing far less to informed traders when AQ is more active.

The highly significant positive effect of AQ on the realized spread in the corn market indicates AQ liquidity providers earn greater revenues from liquidity demanders than their conventional counterparts. Using the results based on the 5-second realized spread as an example, if the corn price is 400 cents/bushel a 0.1 increase in the AQ measure is associated with a 0.1×0.055 bps $\times$ 400 cents/bushel $\times$ 5000 bushel/contract $\times$ 0.01 \$/cents $\approx$ \$1.10 increase in the AQ liquidity provider revenue per contract. A large commercial trader who trades 100 contracts, in this case, pays additional \$100 to liquidity providers for immediacy.

The result suggests that algorithmic liquidity providers, on aggregate, have a competitive advantage in the corn market. As CME's agricultural commodity futures mainly use time to determine order execution priority for orders quoted at the same price,[20] high-frequency

[20] Corn and soybean futures use a split FIFO (first in, first out)/ Pro-Rata based matching algorithm, while the live cattle market only uses the FIFO. The FIFO algorithm only uses time and the Pro-Rata algorithm only uses order size to determine the priority for orders at the same price. Under a split FIFO/Pro-Rata algorithm, when large orders are submitted to the LOB, a certain percentage of each matching order gets allocated FIFO and the remainder is allocated Pro-Rata. By contacting with officials in the CME group, we were informed that about 70% - 80% of the time, FIFO is used to determine order priority in crop futures markets.

algorithmic liquidity providers who have a speed advantage can adversely select slower traders (Li, Wang and Ye, 2017). For example, if a liquidity provision algorithm observes a large buy order has been submitted and anticipates price will move up, the algorithm can quickly replace its current sell order with a higher ask price and sell to a slow trader who does not observe this order information due to the time delay.

In recent years, proprietary algorithmic trading firms have been heavily invested in developing low latency automated trading algorithms. The positive relationship between AQ and the realized spread reflects the marginal benefit of investing in developing liquidity provision algorithms. On the other side, many algorithmic trading firms have been selling high-frequency order-execution services to traditional agricultural commercial futures users to help them avoid losses to high frequency traders.[21] It is not clear why significant effects are not found in the soybean and live cattle markets. It is likely that these markets have different levels of liquidity and competitiveness among algorithmic traders.

3.6 Robustness Checks

In this section, we show that our results are robust to using different sample selection criteria. Algorithmic quoting activity is less likely to exist when the market is not active. Thus, to examine whether results are robust to using inactive market time periods, we exclude measurement intervals that either belong to the first quantile of the number of quote updates or have less than 30 trades, and then replicate the analysis. We find the results are generally the same after excluding these time intervals. To save space, results are presented in supplementary result 3.

[21] For example https://www.rcmalternatives.com/services/futures-traders-hedgers-commercials/ag-hedging-services/.

Hendershott, Jones and Menkveld (2011) show that algorithmic liquidity supplier market advantage is only temporary and declines as more competitive algorithms are developed and used in the market. It could be that our finding of liquidity provider market advantage is dominated by the early sample period and the effect is only temporary. Thus, to examine whether liquidity provider market advantage in the corn market disappeared in the later period of the sample as trading algorithms were increasingly used, we replicate the analysis for sample period before and after 2017 for the corn market (supplementary result 4). The results show AQ significantly increases the 5 and 10-second realized spreads in both periods suggesting that algorithmic liquidity providers in the corn futures market have a competitive advantage throughout the sample period.

3.7 Conclusions

Motivated by recent increasing concerns about the effects of algorithmic activity in agricultural futures markets, this paper investigates how quotations generated by algorithmic trading strategies affect pricing efficiency, short-term volatility, and liquidity in the corn, soybean, and live cattle futures markets.

Following Hendershott, Jones and Menkveld (2011), we use the number of quote updates in the LOB weighted by dollar volume as a measure for algorithmic trading activity. We show that even when overall trading is low, quotes in the LOB are updated frequently which is likely a result of active AQ activity. Our results are consistent with previous findings in equity markets (e.g., Hendershott, Jones and Menkveld, 2011; Conrad, Wahal and Xiang 2015) that, on average, more intensive AQ is not harmful to market quality. In particular, we show AQ significantly improves pricing efficiency in the corn and live cattle market which complements Couleau, Serra and Garcia (2017) who find no evidence of HFT causing meaningful market noises in the live cattle market. AQ also significantly mitigates short-term volatility in all the three markets studied. Higher AQ also significantly reduces the costs of immediacy in the more liquid corn and soybean markets, but not in the live cattle market. Lower costs of immediacy are due to reduced adverse selection costs facing these liquidity providers. The latter is suggestive that algorithmic liquidity providers are better informed than their conventional counterparts and quickly incorporate fundamental news and other information into the market, leading to an overall increase in market quality. Liquidity providers who adopt scalping strategies have long existed in futures market since the pit trading era. As shown in Silber (1984), traditional scalpers' profitability depends on their expertise in evaluating short-run bid-ask imbalances and is negatively correlated with the holding period of their positions. By operating at a higher speed and possibly possessing better market analytical capacities than conventional liquidity providers, AT liquidity providers are capable of obtaining higher profits. Hence, while algorithmic activity, overall, is beneficial to market quality, our findings suggest informed commercial hedgers pay additional costs for improvements in market quality.

Our results have important implications for regulations and market design. While algorithmic traders provide liquidity to the market and reduce the overall costs immediacy, it is important to ensure the benefits of competition among different algorithms. Commodity exchanges may consider offering lower market access and co-location service fees to attract a variety of algorithmic traders. Alternatively, exchanges may consider alternate order matching algorithms, such as frequent batch auctions, to reduce traditional commercial hedgers' disadvantage in speed.

Due to the limitations of the data, we are only able to show the aggregate effects of AQ. While we show that, on aggregate, AQ is beneficial to market quality of multiple dimensions in agricultural futures markets, there is a need for continued assessment and monitoring of AT activity. Future work may consider exploring the heterogeneity of the effects of different types of algorithms. However, this also requires regulatory agencies to provide more detailed information to the research community.

3.8 Tables and Figures

Table 3.1 Summary Statistics, December 2015-May 2017

	Corn		Soybeans		Live Cattle	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Messages per minute	514.925	382.337	1144.776	819.982	189.437	102.996
Trades per minute	39.405	33.887	53.881	46.992	19.969	11.385
Dollar Volume	11463.069	11339.354	20549.512	19756.642	1442.062	941.130
AQ	0.790	1.338	0.719	0.553	3.711	1.904
Pricing efficiency (500 ms)	0.683	0.683	0.757	0.901	1.756	1.503
Pricing efficiency (1s)	0.703	0.760	0.784	0.966	1.747	1.499
Volatility, bps	25.762	20.384	23.688	17.849	51.237	29.831
Volume-weighted effective spread, bps	3.502	0.484	1.504	0.603	1.839	0.608
Volume-weighted realized Spread (5 s), bps	0.374	1.076	0.081	0.586	-0.017	0.625
Volume-weighted realized Spread (10 s), bps	0.365	1.139	0.077	0.718	-0.066	0.686
Volume-weighted adverse selection cost (5 s), bps	3.128	1.253	1.423	1.029	1.856	0.892
Volume-weighted adverse selection cost (10 s), bps	3.137	1.338	1.427	1.138	1.905	0.925

Note: This table presents mean values for AQ and market quality measures. All measures are calculated over 10-minute intervals for corn and soybean, and 25-minute intervals for live cattle. Pricing efficiency measures are based on variance ratios calculated using 500 milliseconds and 1 second short sampling intervals. Realized spread and price impact measures are calculated using quote midpoint 5 and 10 seconds after the trade.

Table 3.2 Means of Market Quality Measures across High-Low AQ Groups, December 2015-May 2017

	Pricing Efficiency (500ms)	Pricing Efficiency (1s)	Volatility	Effective Spread	Realized Spread (5s)	Realized Spread (10s)	Adverse Selection (5s)	Adverse Selection (10s)
				Corn				
AQ_{Low}	0.77	0.79	26.02	3.56	0.31	0.28	3.25	3.29
	(0.96)	(1.24)	(22.60)	(0.65)	(1.14)	(1.29)	(1.47)	(1.66)
AQ_{Medium}	0.71	0.72	26.17	3.52	0.33	0.34	3.19	3.18
	(0.68)	(1.09)	(20.83)	(0.36)	(0.98)	(1.01)	(1.07)	(1.11)
AQ_{High}	0.65	0.65	25.66	3.42	0.46	0.46	2.97	2.97
	(0.41)	(0.72)	(17.60)	(0.38)	(1.09)	(1.09)	(1.16)	(1.16)
				Soybeans				
AQ_{Low}	0.86	0.85	25.90	1.59	0.11	0.08	1.48	1.50
	(1.10)	(1.03)	(23.17)	(0.94)	(0.83)	(1.04)	(1.61)	(1.78)
AQ_{Medium}	0.72	0.77	23.48	1.47	0.08	0.08	1.40	1.39
	(0.84)	(1.00)	(15.58)	(0.29)	(0.39)	(0.45)	(0.52)	(0.57)
AQ_{HIgh}	0.70	0.75	21.69	1.45	0.06	0.07	1.39	1.39
	(0.73)	(0.88)	(12.93)	(0.34)	(0.44)	(0.50)	(0.57)	(0.62)
				Live Cattle				
AQ_{Low}	1.71	1.83	52.08	1.84	-0.03	-0.05	1.87	1.89
	(1.50)	(1.60)	(31.27)	(0.57)	(0.62)	(0.66)	(0.87)	(0.89)
AQ_{Medium}	1.80	1.72	48.79	1.76	-0.06	-0.11	1.82	1.87
	(1.45)	(1.40)	(28.22)	(0.53)	(0.60)	(0.63)	(0.87)	(0.88)
AQ_{High}	1.75	1.56	45.99	1.76	0.00	-0.07	1.76	1.83
	(1.56)	(1.29)	(25.46)	(0.51)	(0.57)	(0.64)	(0.75)	(0.81)

Note: This table presents averages and standard deviations for market quality measures for subsamples that are associated with different levels of AQ in each market. AQ_{Low} , AQ_{Medium}, and AQ_{High} represents sample periods that are associated with the lowest, medium, and highest AQ measures, respectively. Standard deviations are presented in parenthesis. All measures are calculated over 10-minute intervals for corn and soybeans, and 25-minute intervals for live cattle. Pricing efficiency measures are based on variance ratios calculated using 500 milliseconds and 1 second short sampling intervals. Realized spread and adverse selection cost are calculated using quote midpoint 5 and 10 seconds after the trade.

Table 3.3 Effect of Algorithmic Quoting on Pricing Efficiency, December 2015-May 2017

	Corn		Soybeans		Live Cattle	
	500 ms	1 sec	500 ms	1 sec	500 ms	1 sec
AQ_t	-0.096*	-0.138*	0.061	0.080	-0.017	-0.046*
	(0.054)	(0.075)	(0.068)	(0.074)	(0.022)	(0.025)
$\lvert VR_t\text{-}1\rvert_{t-1}$	0.066*	0.074**	0.007	0.010	0.041**	0.035**
	(0.036)	(0.035)	(0.016)	(0.017)	(0.017)	(0.017)
$Trend_{t-19:t-1}$	0.146	0.229*	-0.613	-0.294	-0.018	-0.024
	(0.093)	(0.132)	(1.085)	(1.175)	(0.057)	(0.057)
$USDA_t$	-0.057	-0.081	-0.041	-0.050	0.054	0.194
	(0.167)	(0.170)	(0.109)	(0.116)	(0.407)	(0.517)
$Open_t$	-0.010	-0.005	-0.042	-0.020	-0.389*	-0.379*
	(0.043)	(0.052)	(0.042)	(0.046)	(0.219)	(0.213)
$Close_t$	0.733***	0.855***	0.430***	0.472***	0.012	0.043*
	(0.178)	(0.214)	(0.088)	(0.094)	(0.110)	(0.110)
$IndexRoll_t$	-0.043	-0.032	-0.050**	-0.052*	-0.154**	-0.146
	(0.033)	(0.042)	(0.025)	(0.027)	(0.076)	(0.077)
Intercept	0.691***	0.725***	0.741***	0.761***	1.858***	1.948***
	(0.057)	(0.071)	(0.062)	(0.067)	(0.133)	(0.140)
Over identification test:						
Hansen's *J* Statistic			1.132	1.248		
Weak IV test:						
F Statistic	16.559	18.115	48.830	47.370	27.069	51.052
Number of Obs.	2263	1973	8439	8184	3536	3563

Note: This table presents parameter estimates for the pricing efficiency equations. Equations are estimated separately for each market. Pricing efficiency measures ($\lvert VR_t\text{-}1\rvert$) are based on variance ratios calculated using 500 milliseconds and 1 second short sampling intervals. Standard errors that are robust to heteroskedasticity and autocorrelation are reported in parentheses. *,**,*** denote significance at the 90%, 95%, 99% level, respectively. Results for day-of-week and contract dummies are not reported in the table due to space limit and they can be found in supplementary result 2.

Table 3.4 Effect of Algorithmic Quoting on Volatility, December 2015-May 2017

	Corn	Soybean	Live Cattle
AQ_t	-0.859***	-0.956*	-0.851*
	(0.283)	(0.545)	(0.506)
$HighLow_{t-1}$	0.388***	0.386***	0.357***
	(0.022)	(0.021)	(0.019)
$Trend_{t-19:t-1}$	2.734**	23.632	-1.606
	(1.066)	(18.214)	(1.264)
$USDA_t$	171.092***	142.006***	56.653***
	(21.239)	(19.303)	(8.254)
$Open_t$	29.917***	23.095***	28.954***
	(1.460)	(1.068)	(5.904)
$Close_t$	11.913***	8.091***	29.182***
	(0.972)	(0.742)	(2.541)
$IndexRoll_t$	0.354	-0.528	1.047
	(0.460)	(0.366)	(1.230)
Intercept	1.799***	17.730***	36.009***
	(0.079)	(0.807)	(2.943)
Over identification test:			
Hansen's *J* Statistic		2.914	
Weak IV test:			
F Statistic	38.920	40.360	44.841
Number of Obs.	11001	10916	3574

Note: This table presents parameter estimates from the volatility equations. Equations are estimated separately for each market. Standard errors that are robust to heteroskedasticity and autocorrelation are reported in parentheses. *,**,*** denote significance at the 90%, 95%, 99% level. Results for day-of-week and contract dummies are not reported in the table due to space limit and they can be found in supplementary result 2.

Table 3.5 Effect of Algorithmic Quoting on Liquidity, December 2015-May 2017

	AQ_t	$LagMQ_t$	$Trend_{t-19:t-1}$	$USDA_t$	$Open_t$	$Close_t$	$IndexRoll_t$	*Intercept*	Hansen's *J* Statistic	*F* Statistic
					Corn (*N* = 11001)					
es	-0.018**	0.192***	-0.051**	3.197***	0.155***	0.014	-0.044***	3.016***		41.865
	(0.009)	(0.034)	(0.026)	(0.524)	(0.017)	(0.013)	(0.011)	(0.129)		
rs (5 seconds)	0.055***	0.177***	-0.282***	-1.364***	-0.256***	0.456***	-0.019	0.140***		39.646
	(0.021)	(0.011)	(0.057)	(0.425)	(0.040)	(0.046)	(0.030)	(0.030)		
rs (10 seconds)	0.056**	0.141***	-0.324***	-2.438***	-0.255***	0.386***	0.000	0.153***		39.043
	(0.023)	(0.064)	(0.662)	(0.046)	(0.053)	(0.032)	(0.034)	(0.011)		
ad (5 seconds)	-0.074***	0.177***	0.230***	4.598***	0.412***	-0.443***	-0.026	2.931***		39.045
	(0.021)	(0.012)	(0.063)	(0.728)	(0.043)	(0.050)	(0.085)	(0.056)		
ad (10 seconds)	-0.076***	0.148***	0.268***	5.735***	0.413***	-0.373***	-0.047	3.029***		38.301
	(0.023)	(0.069)	(0.916)	(0.050)	(0.055)	(0.036)	(0.055)	(0.011)		
					Soybeans (*N* = 10916)					
es	-0.062***	0.096***	-0.368	2.204***	0.156***	0.062***	0.005	1.480***	0.040	40.831
	(0.020)	(0.027)	(0.660)	(0.345)	(0.019)	(0.018)	(0.013)	(0.049)		
rs (5 seconds)	-0.006	0.083**	-0.563	-1.083***	-0.067***	0.168***	-0.022	0.047**	0.196	40.438
	(0.024)	(0.038)	(0.759)	(0.349)	(0.022)	(0.026)	(0.015)	(0.022)		
rs (10 seconds)	-0.023	0.097**	-0.510	-1.226**	-0.106***	0.162***	-0.024	0.062**	1.541	40.465
	(0.033)	(0.034)	(0.984)	(0.597)	(0.026)	(0.034)	(0.021)	(0.027)		
ad (5 seconds)	-0.055*	0.066*	0.177	3.480***	0.225***	-0.105***	0.027	1.478***	0.117	40.909
	(0.031)	(0.034)	(1.171)	(0.584)	(0.032)	(0.036)	(0.022)	(0.062)		
ad (10 seconds)	-0.029	0.085**	-0.134	3.592***	0.256***	-0.099**	0.025	1.430***	1.140	40.838
	(0.039)	(0.033)	(1.330)	(0.684)	(0.035)	(0.044)	(0.027)	(0.063)		

Table 3.5 (continued)

	AQ_t	$LagMQ_t$	$Trend_{t-19:t-1}$	$USDA_t$	$Open_t$	$Close_t$	$IndexRoll_t$	*Intercept*	Hansen's *J* Statistic	*F* Statistic
					Live Cattle (*N* = 3574)					
es	0.002	0.274***	-0.048**	0.618*	0.768***	0.096***	0.037	1.503***		45.655
	(0.025)	(0.026)	(0.023)	(0.328)	(0.119)	(0.025)	(0.023)	(0.125)		
rs (5 seconds)	0.007	0.139***	0.028	0.096	0.283***	0.156***	0.020	0.013		47.729
	(0.021)	(0.028)	(0.021)	(0.130)	(0.078)	(0.032)	(0.028)	(0.087)		
rs (10 seconds)	0.000	0.115***	0.009	-0.440***	0.205**	0.175***	0.041	-0.003		48.260
	(0.028)	(0.027)	(0.028)	(0.055)	(0.089)	(0.038)	(0.031)	(0.113)		
ad (5 seconds)	-0.002	0.223***	-0.079**	0.552	0.491***	-0.062	0.021	1.595***		44.901
	(0.017)	(0.024)	(0.034)	(0.456)	(0.159)	(0.043)	(0.038)	(0.102)		
ad (10 seconds)	0.005	0.204***	-0.060	1.086***	0.559***	-0.086*	0.002	1.649***		46.220
	(0.018)	(0.025)	(0.038)	(0.371)	(0.157)	(0.048)	(0.041)	(0.011)		

Note: This table reports effects of AQ on effective spread (*es*) and its realized spread (*rs*) and adverse selection cost (*ad*) components. Equations are estimated separately for each variable and each market. Realized spread and adverse selection cost measures are calculated using quote midpoint 5 and 10 seconds after the trade. Standard errors that are robust to heteroskedasticity and autocorrelation are reported in parentheses. *,**,*** denote significance at the 90%, 95%, 99% level. Results for day-of-week and contract dummies are not reported in the table due to space limit and they can be found in supplementary result 2.

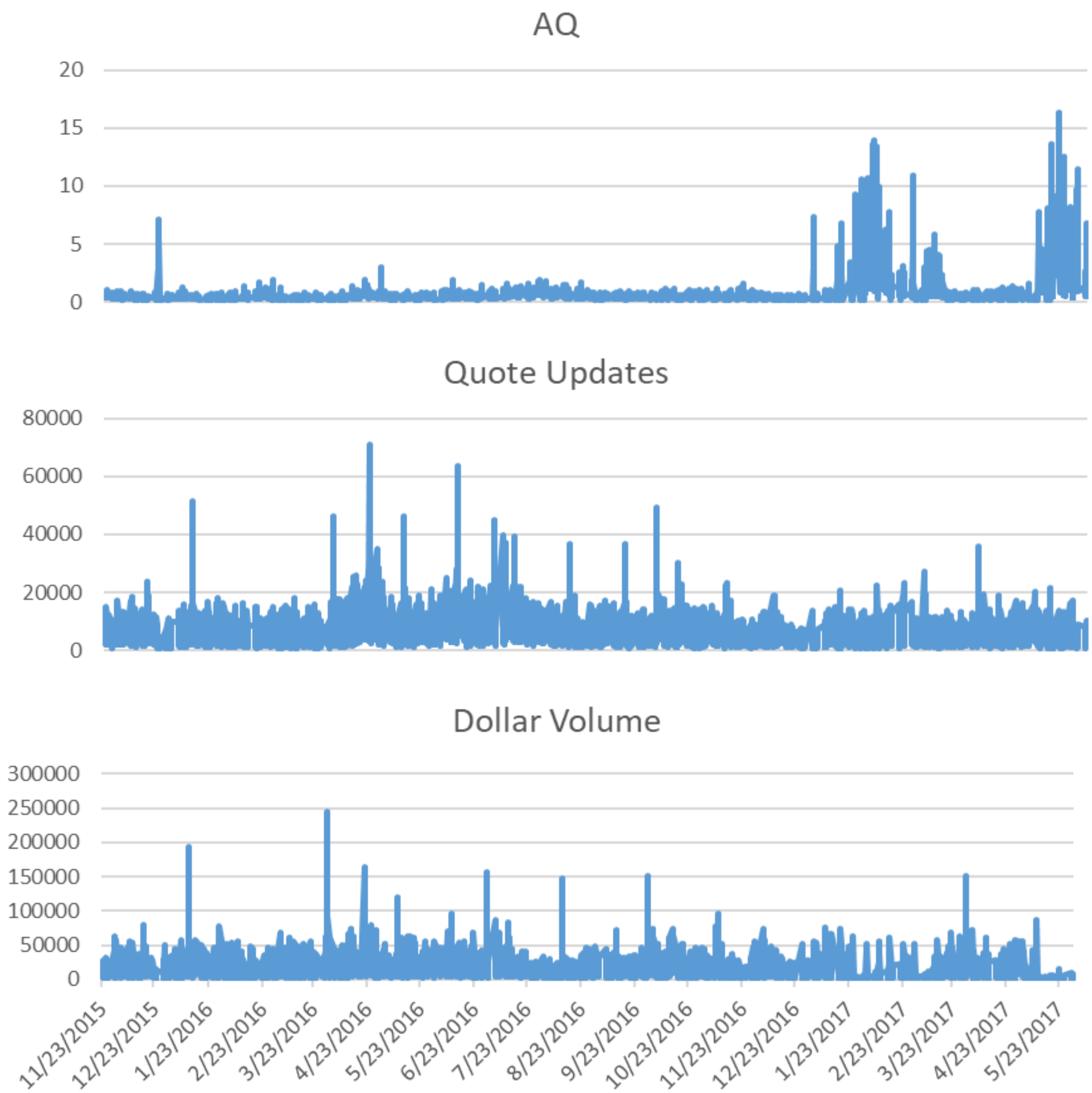


Figure 3.1 AQ, Quote Updates, and Dollar Volume in the Corn Futures Market, December 2015-May 2017

Note: AQ, quote updates, and dollar volume are measured over 10-minute intervals through the sample period. For a more readable plot of AQ measures, four observations for the 11:20-11:30, 11:30-11:40, 12:30-12:40, and 13:00-13:10 intervals on Jan 30, 2017 are clipped, as AQ reached a high of 53.

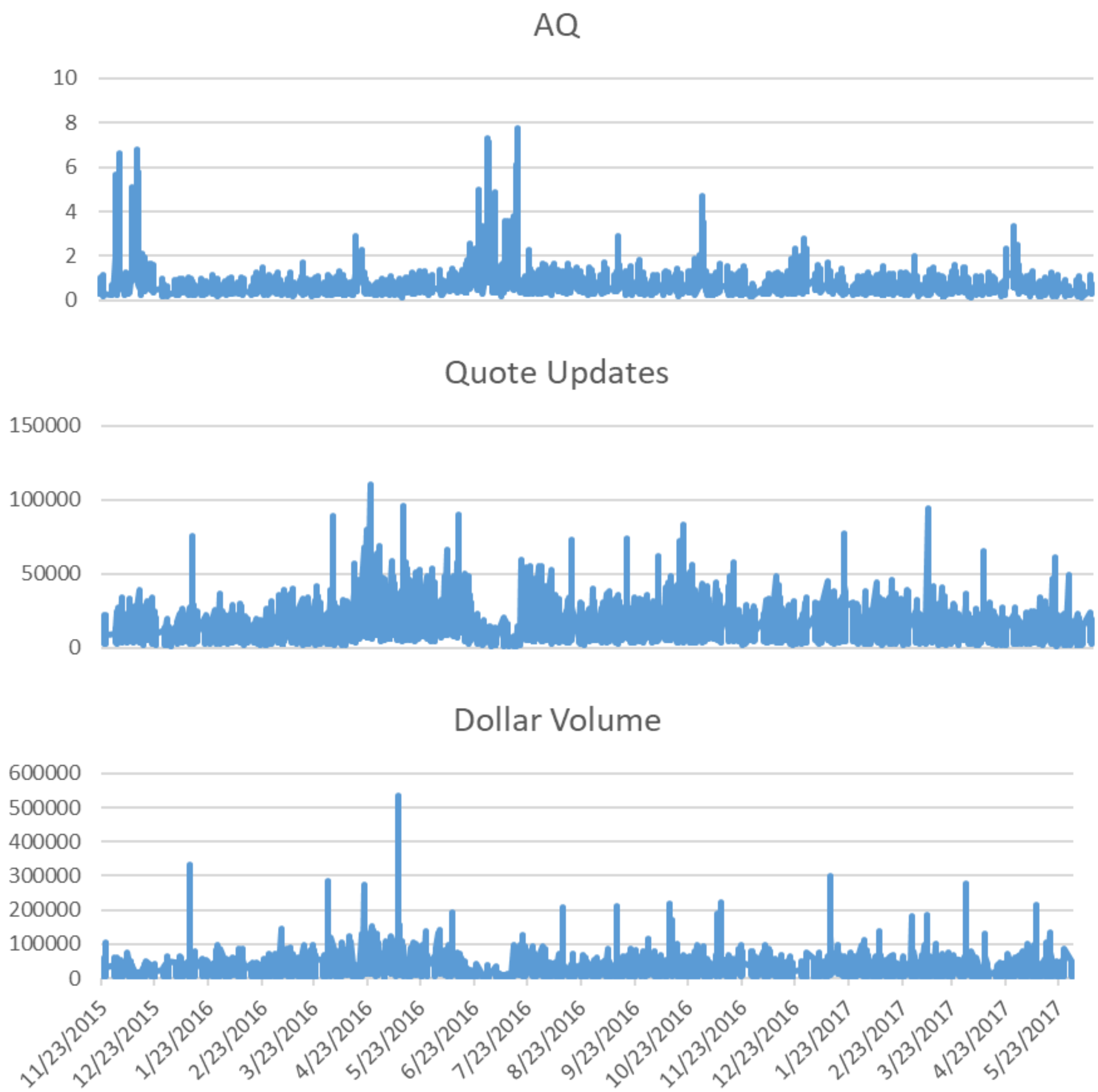


Figure 3.2 AQ, Quote Updates, and Dollar Volume in the Soybean Futures Market, December 2015-May 2017

Note: Number of quote updates, trades, and AQ are measured over 10-minute intervals through the sample period. For a more readable plot of AQ measures, two observations for the 12:30-12:40 interval on June 28, 2016 and the 13:00-13:10 interval on June 27, 2016 are clipped.

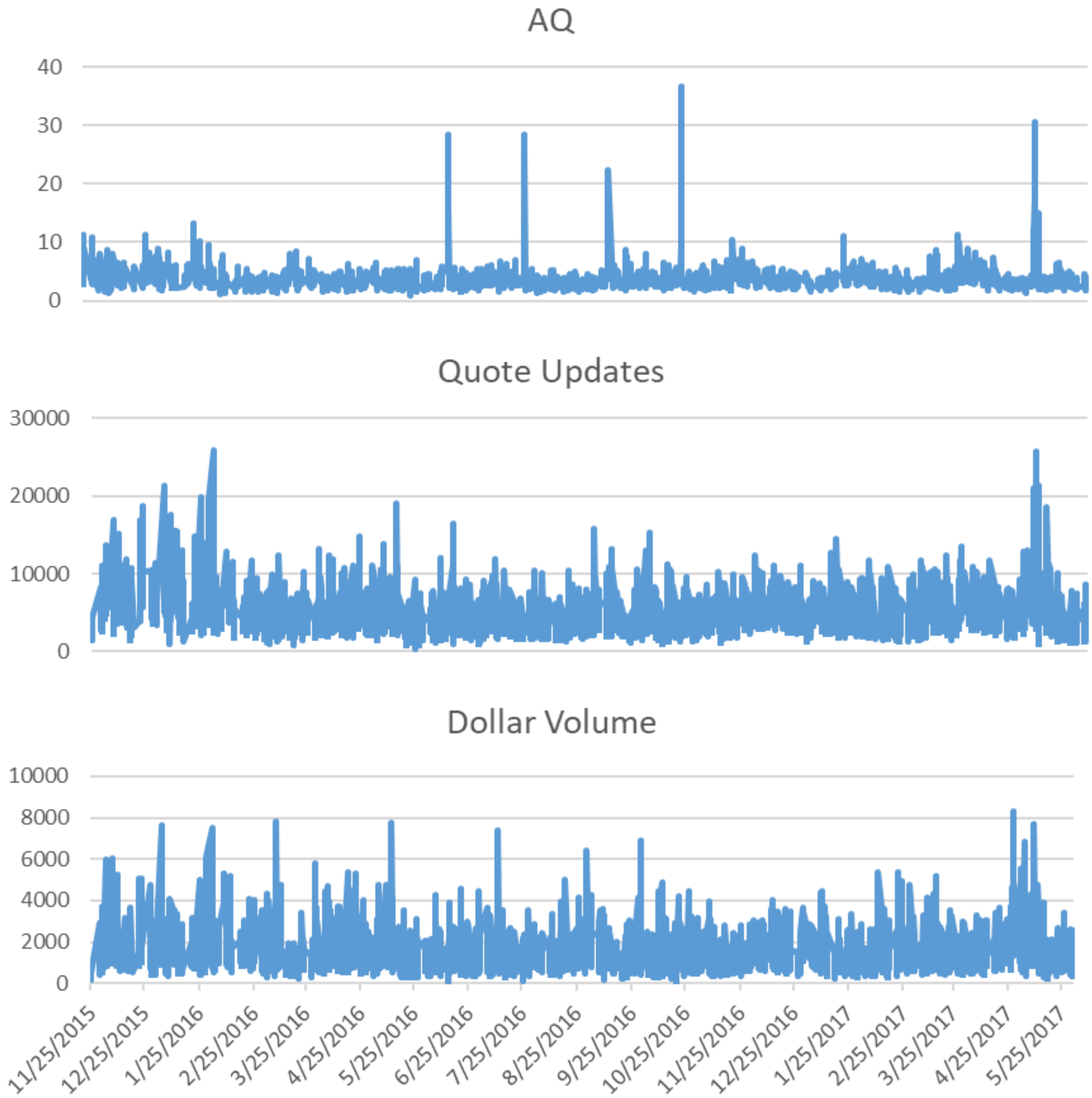


Figure 3.3 AQ, Quote Updates, and Dollar Volume in the Live Cattle Futures Market, December 2015-May 2017

Note: Number of quote updates, trades, and AQ are measured over 25-minute intervals through the sample period.

3.9 Supplementary Results

Supplementary Result 1

This part shows an example of "strategic runs" that happened between 13:06:56:69 and 13:06:56:110. To save space, the table ends at 13:06:56:82. Over this time, updates in the LOB only happened in the best and second-best bids. An order is submitted and canceled multiple times resulting in best bid size changes between 112 and 113 every time the order fleets within 1 millisecond. Meanwhile, another order is submitted and canceled at the second-best bid, resulting in the bid size changes between 281 and 280, and the update in the second best bid typically happened within 1 millisecond after the quote updated in the best bid. It is not clear the two strategic runs are created by the same algorithm or there are two algorithms competing. The two "strategic runs" in total create 135 messages in just 41 milliseconds, until the best ask size starts to decrease.

Supplemental Table 3.1 Strategic Runs

Time	Millisecond	Bid 2	Bid Size 2	Bid 1	Bid Size 1	Ask 1	AskSize 1	Ask 2	AskSize 2
1:06:05 PM	69	373.5	281	373.75	113	374	63	374.25	299
1:06:05 PM	69	373.5	280	373.75	113	374	63	374.25	299
1:06:05 PM	69	373.5	280	373.75	112	374	63	374.25	299
1:06:05 PM	69	373.5	281	373.75	112	374	63	374.25	299
1:06:05 PM	70	373.5	281	373.75	113	374	63	374.25	299
1:06:05 PM	70	373.5	280	373.75	113	374	63	374.25	299
1:06:05 PM	70	373.5	280	373.75	112	374	63	374.25	299
1:06:05 PM	70	373.5	281	373.75	112	374	63	374.25	299
1:06:05 PM	71	373.5	281	373.75	113	374	63	374.25	299
1:06:05 PM	71	373.5	280	373.75	113	374	63	374.25	299
1:06:05 PM	72	373.5	280	373.75	112	374	63	374.25	299
1:06:05 PM	72	373.5	281	373.75	112	374	63	374.25	299
1:06:05 PM	73	373.5	281	373.75	113	374	63	374.25	299
1:06:05 PM	73	373.5	280	373.75	113	374	63	374.25	299
1:06:05 PM	73	373.5	280	373.75	112	374	63	374.25	299
1:06:05 PM	73	373.5	281	373.75	112	374	63	374.25	299
1:06:05 PM	74	373.5	281	373.75	113	374	63	374.25	299
1:06:05 PM	74	373.5	280	373.75	113	374	63	374.25	299
1:06:05 PM	74	373.5	280	373.75	112	374	63	374.25	299
1:06:05 PM	74	373.5	281	373.75	112	374	63	374.25	299
1:06:05 PM	75	373.5	281	373.75	113	374	63	374.25	299
1:06:05 PM	75	373.5	280	373.75	113	374	63	374.25	299
1:06:05 PM	75	373.5	280	373.75	112	374	63	374.25	299
1:06:05 PM	75	373.5	281	373.75	112	374	63	374.25	299
1:06:05 PM	77	373.5	281	373.75	113	374	63	374.25	299
1:06:05 PM	77	373.5	280	373.75	113	374	63	374.25	299
1:06:05 PM	77	373.5	280	373.75	112	374	63	374.25	299
1:06:05 PM	77	373.5	281	373.75	112	374	63	374.25	299
1:06:05 PM	78	373.5	281	373.75	113	374	63	374.25	299
1:06:05 PM	78	373.5	280	373.75	113	374	63	374.25	299
1:06:05 PM	78	373.5	280	373.75	112	374	63	374.25	299
1:06:05 PM	78	373.5	281	373.75	112	374	63	374.25	299
1:06:05 PM	79	373.5	281	373.75	113	374	63	374.25	299
1:06:05 PM	79	373.5	280	373.75	113	374	63	374.25	299
1:06:05 PM	79	373.5	280	373.75	112	374	63	374.25	299
1:06:05 PM	79	373.5	281	373.75	112	374	63	374.25	299
1:06:05 PM	80	373.5	281	373.75	113	374	63	374.25	299
1:06:05 PM	80	373.5	280	373.75	113	374	63	374.25	299
1:06:05 PM	81	373.5	280	373.75	112	374	63	374.25	299
1:06:05 PM	81	373.5	281	373.75	112	374	63	374.25	299

Supplementary Result 2

This section shows the full estimation results for the regressions.

Supplemental Table 3.2 Parameter Estimates for the Corn Market, December 2015-May 2017

	Price Efficiency (500 ms)	Price Efficiency (1 sec)	Volatility	Effective Spread	Realized Spread (5 sec)	Realized Spread (10 sec)	Adverse Selection (5 sec)	Adverse Selection (10 sec)
AQ	-0.096*	-0.138*	-0.859***	-0.018**	0.055***	0.056**	-0.074***	-0.076***
	(0.054)	(0.075)	(0.283)	(0.009)	(0.021)	(0.023)	(0.021)	(0.023)
LagMQ	0.066*	0.074**	3.883***	0.192***	0.177***	0.141***	0.177***	0.148***
	(0.036)	(0.035)	(0.218)	(0.034)	(0.011)	(0.064)	(0.012)	(0.069)
Trend	0.146	0.229*	2.734**	-0.051**	-0.282***	-0.324***	0.230***	0.268***
	(0.093)	(0.132)	(1.066)	(0.026)	(0.057)	(0.662)	(0.063)	(0.916)
USDA	-0.057	-0.081	171.092***	3.197***	-1.364***	-2.438***	4.598***	5.735***
	(0.167)	(0.170)	(21.239)	(0.524)	(0.425)	(0.046)	(0.728)	(0.050)
Open	-0.010	-0.005	29.917***	0.155***	-0.256***	-0.255***	0.412***	0.413***
	(0.043)	(0.052)	(1.460)	(0.017)	(0.040)	(0.053)	(0.043)	(0.055)
Close	0.733***	0.855***	11.913***	0.014	0.456***	0.386***	-0.443***	-0.373***
	(0.178)	(0.214)	(0.972)	(0.013)	(0.046)	(0.032)	(0.050)	(0.036)
Roll	-0.043	-0.032	0.354	-0.044***	-0.019	0.000	-0.026	-0.047
	(0.033)	(0.042)	(0.460)	(0.011)	(0.030)	(0.034)	(0.034)	(0.040)
Mon	-0.037	-0.042	-0.117	0.000	-0.036	-0.061*	0.036	0.061
	(0.050)	(0.057)	(0.450)	(0.016)	(0.032)	(0.036)	(0.038)	(0.042)
Tue	-0.003	0.014	0.194	-0.010	-0.039	-0.055	0.028	0.044
	(0.044)	(0.053)	(0.474)	(0.014)	(0.033)	(0.032)	(0.039)	(0.036)

Supplemental Table 3.2 (continued)

	Price Efficiency (500 ms)	Price Efficiency (1 sec)	Volatility	Effective Spread	Realized Spread (5 sec)	Realized Spread (10 sec)	Adverse Selection (5 sec)	Adverse Selection (10 sec)
Wen	-0.060	-0.041	-0.602	-0.022*	-0.019	-0.021	-0.004	-0.003
	(0.042)	(0.051)	(0.425)	(0.012)	(0.031)	(0.033)	(0.035)	(0.036)
Thu	0.004	-0.004	0.870*	-0.005	-0.043	-0.034	0.038	0.029
	(0.045)	(0.053)	(0.464)	(0.012)	(0.031)	(0.029)	(0.035)	(0.032)
H	0.003	0.013	-6.540***	-0.180***	0.219***	0.237***	-0.402***	-0.425***
	(0.049)	(0.064)	(0.433)	(0.012)	(0.027)	(0.031)	(0.030)	(0.036)
K	-0.051	-0.088	-7.178***	-0.213***	0.253***	0.253***	-0.469***	-0.475***
	(0.061)	(0.064)	(0.443)	(0.015)	(0.029)	(0.034)	(0.035)	(0.041)
N	-0.004	-0.015	-0.247	-0.346***	0.050*	0.053	-0.401***	-0.416***
	(0.031)	(0.037)	(0.512)	(0.020)	(0.030)	(0.090)	(0.036)	(0.091)
U	0.083	0.137	-4.627***	-0.253***	0.637***	0.541***	-0.890***	-0.798***
	(0.182)	(0.235)	(0.967)	(0.037)	(0.084)	(0.031)	(0.085)	(0.055)
Intercept	0.691***	0.725***	17.987***	3.016***	0.140***	0.153***	2.931***	3.029***
	(0.057)	(0.071)	(0.786)	(0.129)	(0.030)	(0.011)	(0.056)	(0.011)

Note: This table presents parameter estimates for the corn market. Equations are estimated separately. All measures are calculated over 10-minute intervals. Pricing efficiency measures are based on variance ratios calculated using 500 milliseconds and 1 second short sampling intervals. Realized spread and adverse selection cost are calculated using quote midpoint 5 and 10 seconds after the trade. Standard errors that are robust to heteroskedasticity and autocorrelation are reported in parentheses. *,**,*** denote significance at the 90%, 95%, 99% level, respectively. *H*,*K*,*N*,*U* represents March, May, July, and September, respectively.

Supplemental Table 3.3 Parameter Estimates for the Soybean Market, December 2015-May 2017

	Price Efficiency (500 ms)	Price Efficiency (1 sec)	Volatility	Effective Spread	Realized Spread (5 sec)	Realized Spread (10 sec)	Adverse Selection (5 sec)	Adverse Selection (10 sec)
AQ	0.061	0.080	-0.956*	-0.062***	-0.006	-0.023	-0.055*	-0.029
	(0.068)	(0.074)	(0.545)	(0.020)	(0.024)	(0.033)	(0.031)	(0.039)
LagMQ	0.007	0.010	0.386***	0.096***	0.083**	0.097**	0.066*	0.085**
	(0.016)	(0.017)	(0.021)	(0.027)	(0.038)	(0.034)	(0.034)	(0.033)
Trend	-0.613	-0.294	23.632	-0.368	-0.563	-0.510	0.177	-0.134
	(1.085)	(1.175)	(18.214)	(0.660)	(0.759)	(0.984)	(1.171)	(1.330)
USDA	-0.041	-0.050	142.006***	2.204***	-1.083***	-1.226**	3.480***	3.592***
	(0.109)	(0.116)	(19.303)	(0.345)	(0.349)	(0.597)	(0.584)	(0.684)
Open	-0.042	-0.020	23.095***	0.156***	-0.067***	-0.106***	0.225***	0.256***
	(0.042)	(0.046)	(1.068)	(0.019)	(0.022)	(0.026)	(0.032)	(0.035)
Close	0.430***	0.472***	8.091***	0.062***	0.168***	0.162***	-0.105***	-0.099**
	(0.088)	(0.094)	(0.742)	(0.018)	(0.026)	(0.034)	(0.036)	(0.044)
Roll	-0.050**	-0.052*	-0.528	0.005	-0.022	-0.024	0.027	0.025
	(0.025)	(0.027)	(0.366)	(0.013)	(0.015)	(0.021)	(0.022)	(0.027)
Mon	0.023	0.021	-0.357	0.019	-0.016	-0.026	0.036	0.045*
	(0.035)	(0.039)	(0.391)	(0.016)	(0.013)	(0.016)	(0.022)	(0.024)
Tue	-0.044	-0.051	-0.299	0.003	-0.014	-0.014	0.018	0.014
	(0.030)	(0.033)	(0.387)	(0.013)	(0.016)	(0.022)	(0.024)	(0.029)
Wen	0.005	0.002	-0.616*	-0.001	-0.017	-0.030*	0.017	0.030
	(0.032)	(0.035)	(0.367)	(0.012)	(0.014)	(0.016)	(0.021)	(0.023)
Thu	-0.037	-0.046	0.106	0.016	-0.001	-0.007	0.017	0.027
	(0.029)	(0.033)	(0.400)	(0.020)	(0.019)	(0.022)	(0.036)	(0.038)

Supplemental Table 3.3 (continued)

	Price Efficiency (500 ms)	Price Efficiency (1 sec)	Volatility	Effective Spread	Realized Spread (5 sec)	Realized Spread (10 sec)	Adverse Selection (5 sec)	Adverse Selection (10 sec)
F	-0.064*	-0.067*	-3.097***	-0.112***	0.007	0.016	-0.122***	-0.128***
	(0.035)	(0.038)	(0.464)	(0.019)	(0.017)	(0.020)	(0.029)	(0.031)
H	-0.094***	-0.111***	-6.393***	-0.117***	0.092***	0.093***	-0.215***	-0.213***
	(0.027)	(0.030)	(0.406)	(0.015)	(0.013)	(0.015)	(0.022)	(0.023)
K	-0.078**	-0.088*	-8.230***	-0.163***	0.103***	0.093***	-0.273***	-0.254***
	(0.033)	(0.036)	(0.465)	(0.024)	(0.021)	(0.023)	(0.040)	(0.041)
N	0.017	0.030	-1.125***	-0.153***	-0.009	0.000	-0.150***	-0.157***
	(0.032)	(0.035)	(0.428)	(0.016)	(0.014)	(0.019)	(0.024)	(0.027)
Q	0.358***	0.302***	7.402***	0.457***	0.106*	0.135*	0.365***	0.318***
	(0.106)	(0.106)	(1.607)	(0.053)	(0.057)	(0.076)	(0.078)	(0.090)
Intercept	0.741***	0.761***	17.730***	1.480***	0.047**	0.062**	1.478***	1.430***
	(0.062)	(0.067)	(0.807)	(0.049)	(0.022)	(0.027)	(0.062)	(0.063)

Note: This table presents parameter estimates for the soybean market. Equations are estimated separately. All measures are calculated over 10-minute intervals. Pricing efficiency measures are based on variance ratios calculated using 500 milliseconds and 1 second short sampling intervals. Realized spread and adverse selection cost are calculated using quote midpoint 5 and 10 seconds after the trade. Standard errors that are robust to heteroskedasticity and autocorrelation are reported in parentheses. *,**,*** denote significance at the 90%, 95%, 99% level, respectively. *F*,*H*,*K*,*N*,*Q*, represents February, March, May, July, and August, respectively.

Supplemental Table 3.4 Parameter Estimates for the Live Cattle Market, December 2015-May 2017

	Price Efficiency (500 ms)	Price Efficiency (1 sec)	Volatility	Effective Spread	Realized Spread (5 sec)	Realized Spread (10 sec)	Adverse Selection (5 sec)	Adverse Selection (10 sec)
AQ	-0.017	-0.046*	-0.851*	0.002	0.007	0.000	-0.002	0.005
	(0.022)	(0.025)	(0.506)	(0.025)	(0.021)	(0.028)	(0.017)	(0.018)
LagMQ	0.041**	0.035**	0.357***	0.274***	0.139***	0.115***	0.223***	0.204***
	(0.017)	(0.017)	(0.019)	(0.026)	(0.028)	(0.027)	(0.024)	(0.025)
Trend	-0.018	-0.024	-1.606	-0.048**	0.028	0.009	-0.079**	-0.060
	(0.057)	(0.057)	(1.264)	(0.023)	(0.021)	(0.028)	(0.034)	(0.038)
USDA	0.054	0.194	56.653***	0.618*	0.096	-0.440***	0.552	1.086***
	(0.407)	(0.517)	(8.254)	(0.328)	(0.130)	(0.055)	(0.456)	(0.371)
Open	-0.389*	-0.379*	28.954***	0.768***	0.283***	0.205**	0.491***	0.559***
	(0.219)	(0.213)	(5.904)	(0.119)	(0.078)	(0.089)	(0.159)	(0.157)
Close	0.012	0.043*	29.182***	0.096***	0.156***	0.175***	-0.062	-0.086*
	(0.110)	(0.110)	(2.541)	(0.025)	(0.032)	(0.038)	(0.043)	(0.048)
Roll	-0.154**	-0.146	1.047	0.037	0.020	0.041	0.021	0.002
	(0.076)	(0.077)	(1.230)	(0.023)	(0.028)	(0.031)	(0.038)	(0.041)
Mon	-0.015	0.010	-3.834***	-0.054*	-0.076**	-0.083**	0.013	0.020
	(0.083)	(0.083)	(1.271)	(0.029)	(0.035)	(0.038)	(0.044)	(0.046)
Tue	-0.024	-0.010	-0.251	-0.044*	-0.059*	-0.080**	0.003	0.022
	(0.079)	(0.079)	(1.356)	(0.025)	(0.030)	(0.033)	(0.041)	(0.043)
Wen	0.037	0.038	0.945	-0.003	-0.028	-0.048	0.023	0.041
	(0.081)	(0.081)	(1.283)	(0.026)	(0.027)	(0.031)	(0.039)	(0.042)
Thu	0.032	0.041	1.737	0.037	-0.046	-0.029	0.081*	0.065
	(0.081)	(0.081)	(1.375)	(0.029)	(0.028)	(0.031)	(0.043)	(0.044)

Supplemental Table 3.4 (continued)

	Price Efficiency (500 ms)	Price Efficiency (1 sec)	Volatility	Effective Spread	Realized Spread (5 sec)	Realized Spread (10 sec)	Adverse Selection (5 sec)	Adverse Selection (10 sec)
G	-0.112	-0.092	-2.734	-0.212***	-0.076*	-0.058	-0.161***	-0.180***
	(0.097)	(0.097)	(1.761)	(0.034)	(0.041)	(0.043)	(0.058)	(0.059)
J	-0.340***	-0.329***	-11.125***	-0.397***	0.055	0.045	-0.475***	-0.475***
	(0.097)	(0.096)	(1.660)	(0.035)	(0.036)	(0.039)	(0.053)	(0.056)
M	0.093	0.082	-3.341*	-0.270***	-0.064*	-0.073*	-0.232***	-0.230***
	(0.095)	(0.095)	(1.776)	(0.037)	(0.037)	(0.041)	(0.055)	(0.058)
Q	-0.133	-0.153	-1.536	-0.212***	-0.246***	-0.255***	-0.004	-0.001
	(0.102)	(0.101)	(1.943)	(0.037)	(0.045)	(0.048)	(0.062)	(0.064)
V	-0.027	-0.015	-3.854**	-0.195***	0.048	0.099**	-0.252***	-0.300***
	(0.107)	(0.107)	(1.963)	(0.040)	(0.042)	(0.047)	(0.059)	(0.062)
Intercept	1.858***	1.948***	36.009***	1.503***	0.013	-0.003	1.595***	1.649***
	(0.133)	(0.140)	(2.943)	(0.125)	(0.087)	(0.113)	(0.102)	(0.110)

Note: This table presents parameter estimates the live cattle market. Equations are estimated separately. All measures are calculated over 25-minute intervals. Pricing efficiency measures are based on variance ratios calculated using 500 milliseconds and 1 second short sampling intervals. Realized spread and adverse selection cost are calculated using quote midpoint 5 and 10 seconds after the trade. Standard errors that are robust to heteroskedasticity and autocorrelation are reported in parentheses. *,**,*** denote significance at the 90%, 95%, 99% level, respectively. *G, J, M, Q, V,* represents February, April, June, August, and October, respectively.

Supplementary Result 3

This table replicates regression results in the paper using measurement intervals that do not belong to the first quantile of the distribution of the number of quote updates and contain at least 30 trades. The replications use the same test and estimation procedures as in the paper. IVs are created using the same variables as used in the paper. To save space, only the coefficient estimates for the AQ measure are reported. Standard errors that are robust to heteroskedasticity and autocorrelation are reported in parentheses. *,**,*** denote significance at the 90%, 95%, 99% level.

Supplemental Table 3.5 Robustness Check

	Pricing efficiency (500ms)	Pricing efficiency (1s)	Volatility	Effective Spread	Realized Spread (5s)	Realized Spread (10s)	Adverse Selection (5s)	Adverse Selection (10s)
Corn	-0.096*	-0.138*	-0.955**	-0.030**	0.085***	0.080**	-0.118***	-0.115***
	(0.054)	(0.075)	(0.013)	(0.013)	(0.028)	(0.033)	(0.030)	(0.034)
Soybeans	0.050	0.061	-4.798***	-0.157***	-0.024	-0.059	-0.134**	-0.100
	(0.071)	(0.075)	(1.173)	(0.038)	(0.039)	(0.053)	(0.060)	(0.072)
Live Cattle	-0.005	-0.025	-0.942*	0.014	0.017	0.016	-0.003	-0.003
	(0.023)	(0.022)	(0.570)	(0.028)	(0.023)	(0.031)	(0.019)	(0.020)

Supplementary Result 4

This section replicates the results for the effects of AQ on the effective spread, adverse selection cost, and realized spread in the corn futures market using sample periods before and after 2017. The replications use the same test and estimation procedures as in the paper. IVs are created using the same variables as used in the paper. To save space, only the coefficient estimates for the AQ measure are reported. Standard errors that are robust to heteroskedasticity and autocorrelation are reported in parentheses. *,**,*** denote significance at the 90%, 95%, 99% level.

Supplemental Table 3.6 Subsample Results

	Pre 2017	Post 2017
Effective Spread	-0.015	-0.014
	(0.015)	(0.012)
Realized Spread (5 seconds)	0.081*	0.095***
	(0.044)	(0.034)
Realized Spread (10 seconds)	0.081**	0.097***
	(0.035)	(0.037)
Adverse Selection (5 seconds)	-0.105	-0.107***
	(0.044)	(0.037)
Adverse Selection (10 seconds)	-0.107**	-0.109***
	(0.033)	(0.039)

CHAPTER 4:

EXCHANGE RATE EFFECTS ON AGRICULTURAL EXPORT PRICES AND SALES IN HIGH-LOW STOCK REGIMES

4.1. Introduction

Since Schuh's (1974) classic article, the effects of exchange rates on U.S. agricultural commodity exports has been area of investigation. Schuh explains the agricultural commodity price boom and increased world demand for U.S. farm products in the early 1970s as a result of the two Nixon administration dollar devaluations. Chambers and Just (1981) also provide empirical evidence that the devaluation of the dollar in 1971 significantly boosted U.S. agricultural commodity exports and prices. Later, Schuh (1984), Chambers (1984) and Orden (1986) attribute depressed agricultural exports in the early 1980s to an overvalued dollar, and MacDonald (1992) and Stallings (1988) argue that recovered U.S. agricultural export activity in the late 1980s and 1990s can be explained by a decline in the value of the dollar. While major U.S. agricultural exports in the 2000s are largely driven by increased demand from emerging economies (Wright, 2011), recent studies also document linkages between exchange rates and agricultural prices (Harri, Nalley and Hudson, 2009; Hatzenbuehler, Abbott and Foster, 2016), which can impact exports.

Despite these findings, other studies fail to identify a clear relationship between exchange rates and agricultural commodity export prices and quantities. For example, Bessler and Babul (1987) find exchange rates have some impacts on wheat prices but little effect on wheat export sales and shipments. Babul, Ruppel and Sessler (1995) show the same results in the corn export market. Bradshaw and Orden (1990) reveal mixed evidence on the causality from exchange rates to prices. Frank and Garcia (2010) find exchange rates have limited influence on agricultural commodity markets during the period of 1998-2006. But linkages between exchange rates and agricultural prices appear to be stronger during the 2006 -2009 commodity boom and bust. The mixed evidence on the effects of exchange rates on export prices and quantities warrants further research.

Classic commodity excess supply-demand theory (Kost, 1976; Chambers and Just, 1979) suggests that the domestic demand along with available domestic supply shape the excess supply function for agricultural exports. In the large exporter case, its excess supply function and the excess demand from importers interact to determine quantity traded and prices that in a competitive environment and in absence of transaction costs are equal in the exporting and importing countries when adjusted by exchange rates. Increases in the value of importers' exchange rates increase excess demand and raise prices and quantity traded. Conversely, a decrease in the value of importers exchange rates decreases excess demand and reduces prices and quantity traded. In these situations, the magnitudes of the change in price and quantity will be influenced by the magnitude of the change in the exchange rate, and by the shape of the excess supply function. The more inelastic (elastic) the excess supply function the larger (smaller) the price effect and the smaller (larger) the quantity effect.

In storable commodity markets, the theory of competitive storage shows that the overall elasticity of domestic demand for grains varies with the level of inventories and is highly inelastic when the level of stocks falls below a certain level (Wright, 2011). Since the excess supply function is determined by domestic demand and available domestic supply, it can become inelastic as well and results in large price changes with changing exchange rates. As a consequence, the degree of exchange rate impacts on export prices and quantities may differ depending on stock conditions. More specifically, one would anticipate that responses of export prices to exchange rate changes are greater when stocks are low. Similarly, when stocks are low one would expect sales to vary little (Chambers and Just, 1979). However, this may not be the case in a dynamic situation. Chambers and Just (1981) illustrate that export responses to changes in exchange rates are more elastic than static theory would imply, and that shorter responses are often more pronounced since markets adjust to changing conditions.

In this paper, we investigate how export prices and sales respond to exchange rate movements allowing for changes in stocks-to-use ratio in several important grain markets for the period of 1990-2019. The analysis is performed in the corn, soybean, and wheat export markets using Threshold Vector Autoregressive (TVAR) models and monthly data. We extend the linear VAR framework that has been widely adopted in previous literature to study linkages among the exchange rate, commodity price, and export sales, by introducing the stocks-to-use ratio as an endogenous threshold variable. This allows us to relate the variation in the impact of the exchange rate on export price and volume to changes in exporter's market stocks-to-use ratio.

This study contributes to the literature by providing an economic explanation for changes in the exchange rate-exports relationship. Previous studies typically address this issue statistically by identifying a structural break and estimating the impact of the exchange rate in different sample periods (e.g., Babula, Ruppel and Bessler, 1995). However, structural breaks that are relevant to storable commodities should be reflected in the stocks-to-use ratio. This is the case of the 2006-2009 food commodity price boom and bust that are mostly explained by low stocks of major grains (Wright, 2011; Janzen et al., 2014; Bruno, Büyükşahin and Robe, 2017). Empirically framing the changes in market behavior in the context of a continuous variable like the stocks-to-use ratio may be preferable to using a one-time structural phenomenon. These changes are likely to recur regularly, though with different intensities, and seem more realistically represented based on changing market stock conditions.

Perhaps the closest paper to ours is Hatzenbuehler, Abbott and Foster (2016), who show that supply-use factors such as low stocks and policy shifts can affect the responsiveness of soybean and corn prices to exchange rates. We extend their study in several ways. First, in addition to the price responsiveness, we explore for the first time the impact of the level of stocks on the responsiveness of export sales to exchange rates. In recent years, major U.S. grain stocks-to-use ratios are above the historical averages and prices are depressed. In an era where a weak dollar has been advocated to boost exports, understanding the impact of the exchange rate on exports allowing for differences in stock conditions has important policy implications.[22]

Second, in addition to USDA trade-weighted exchange rates that reflect changes in the dollar value relative to importing countries' currencies, we investigate for the first time how U.S. export competitors' exchange rates affect U.S. export prices and sales by constructing exporter weighted exchange rate indices for each commodity using major exporters' share of global exports as weights. A strong dollar tends to make U.S. agricultural exports less competitive as it raises export prices relative to other competitors. However, considering the differences in marketing seasons and grain quality between the U.S. and other exporting countries, the degree to which export sales are affected by competitors' currencies is not clear. Our paper provides the first empirical evidence.

Third, although Hatzenbuehler, Abbott and Foster (2016) do not ignore stocks, they treat them as an exogenous variable in a static model. We, instead, use the TVAR model that allows for the

[22] For example, see Steven Mnuchin who endorsed the weakening of the dollar as good for U.S. exports at the World Economic Forum in Davos.

dynamics of the system of export prices, quantities and exchange rates to adjust to different endogenous stock conditions. In addition, the TVAR model also allows us to generate Generalized Impulse Response Functions to show the time-variant responses of export prices and sales to exchange rate changes given different levels of stocks-to-use ratio.

Overall, our results suggest that the influence of the exchange rate on U.S. agricultural exports is rather complex. The impacts of the exchange rate on export prices and sales differ across markets, and the importance of importer and exporter exchange rate also differ within each market. Primarily, the extent to which exchange rate changes affect export prices and sales is determined by the degree of market dependence on exports. Exchange rate effects are more pronounced in the more export-dependent soybean and wheat markets. In contrast, the effects of both importer and exporter exchange rates in the corn market are either not significant or small in economic value due to the relatively small export share of production.

The results for the soybean and wheat export markets indicate that exporter exchange rate effects are limited by the low substitutability of exports between the U.S. and other exporting countries. Specifically, major soybean exporters have different marketing seasons than the U.S., and U.S. wheat classes and uses are different from other wheat exporters. Our results indicate that while a dollar appreciation relative to other export competitors has expected negative impacts on export prices in soybean and wheat markets, the magnitudes of these effects are much smaller compared to the effect of changes in importer exchange rates. More importantly, exporter exchange rates effects on both soybean and wheat export sales are not significant.

Expected threshold effects are found in the importer exchange rate-exports relationship in soybean and wheat markets, indicating the important role of stock conditions in determining the effects of importer exchange rates in these markets. In both markets, an increase in the value of the dollar relative to importing countries' currencies has significant and negative impacts on export prices and sales in both low and high stocks-to-use regimes. More importantly, responses of export prices and sales to exporter exchange rate changes are higher in the low stocks-to-use regime in soybean and wheat markets. The corn exporter exchange rate also presents similar threshold effects, albeit price and export sales responses are not largely different in their economic value across regimes. Our results provide important implications for both policymakers and market participants that stocks-to-use conditions need to be considered for accurate evaluations and forecasts on exchange rate effects in agricultural markets.

4.2 Data

We examine effects of exchange rates on U.S. agricultural commodity exports in the corn, soybean, and wheat markets for the January 1990 to December 2019 period. The period is chosen because fixed exchange rates were used by several countries (e.g., Russia, China, Ukraine) before 1990.

4.2.1 Exchange Rate Indices

Previous studies typically prefer exchange rate indices to bilateral exchange rates because the former can better capture multilateral trade and balance of payments. In this analysis, we consider two exchange rate indices. The first one is the USDA real monthly importer-weighted exchange rate index compiled by the USDA Economic Research Service (USDA ERS). This exchange rate index is calculated for each commodity using real trade-weighted currencies of U.S. major importers of the commodity relative to the U.S. dollar. Trade weights are the import shares of importing countries, which are calculated based on average shares of U.S. exports during the 2014-2016 marketing years. In the corn market, the top three importers—Mexico, Japan and Korea—account for nearly 60% of the total share. In the soybean market, the top three importers—China, Mexico and Japan—account for more than 70% of the total share. In the wheat market, the top five importers are Japan, Mexico, Philippines, Nigeria, and Brazil. They account for 50% of the total share. While the USDA importer exchange rate index is the most commonly used exchange rate index in the literature, it has the limitation that it only captures the effects of exchange rate movements in countries that import agricultural commodities from the U.S. However, U.S. agricultural commodity exports can also be affected by currencies of other competitors in export markets.

To study effects of exchange rates of major U.S. agricultural commodity export competitors on U.S. agricultural exports, we construct a real monthly export-weighted exchange rate index by using real exchange rates of major exporting countries and their export shares as weights. We construct the exporter-weighted exchange rate index using procedures similar to those USDA employs to generate the importer-weighted exchange rate index for each commodity. First, we obtain real exchange rates of major exporters in each market from the WASDE reports for the 17/18 marketing year, the most recent finalized numbers.[23] Then, we calculate their weights based on their total exports in the 17/18 marketing year. Brazil, Argentina and Paraguay are used for the soybean market and their weights are 90.3%, 2.5%, and 7.2%, respectively. Brazil, Argentina, Ukraine, Russia and South Africa are used for the corn market and their weights are 33.4%, 31.1%, 25.0%, 7.7%, and 2.8%, respectively. Russia, European Union, Canada, Ukraine

[23] USDA ERS typically uses weights based on more recent periods. However, trade weights are not updated annually. Here, we use the most recent exports instead of exports during same time period used by the USDA ERS as several new competitors have emerged in recent years (e.g. Russia and Ukraine in the global wheat market). Including new competitors' exchange rates will provide more relevant implications for the future.

Australia and Argentina are used for the wheat market and their weights are 31.6%, 17.8%, 16.8%, 13.6%, 10.6%, and 9.6%, respectively.[24] Then, the exporter exchange rate index for each commodity can be derived by multiplying the export weights by the respective real exchange rates.

The two upper panels in figure 4.1 plot the two exchange rate indices over time with 1990 as the base year. Both exchange rates represent foreign currencies relative to the U.S. dollar and therefore an increase in each of the exchange rate indices indicates an appreciation of the dollar. The three importer exchange rates follow a similar pattern over the sample period. An exception occurs in the soybean importer exchange rate which increased sharply in 1994 when China officially devaluated its currency by 33% as part of a tightly managed floating exchange rate policy.

Exporter exchange rates also follow the same general pattern across markets as they are all largely affected by South American grain exporters like Brazil and Argentina, and/or countries in the Black Sea region like Russia and Ukraine. However, the exporter exchange rate for corn seems to be more volatile compared to the other two exporter exchange rates after 2005. Particularly, in the corn market exporter exchange rate index, where Ukraine Hryvnia accounts for 25% of the weight, the sharp increase in the exporter exchange rate in early 2015 is related to the escalation of the war in eastern Ukraine. However, the magnitude of the increase in the wheat exporter exchange rate in the same period is smaller as the weight of Ukraine Hryvnia is only 13%. No similar sharp increase is found in the soybean exporter exchange rate as it is not affected by the Ukraine currency.

[24] Former Soviet Union countries and European Union countries use different currencies in early periods of the sample; see USDA ERS for how continuous exchange rate series are created for those countries.

4.2.2 Prices, Export Sales, and Stocks-to-Use Ratios

Export prices obtained from the USDA Agricultural Marketing Services, are the average monthly FOB prices for No. 2 yellow corn, No. 2 yellow soybeans, and No. 1 hard red winter wheat, all delivered at the Gulf of Mexico.[25] Consistent with previous studies, these prices are deflated using the U.S. Consumer Price Index (CPI) from the Federal Reserve to create real prices.

Monthly net export sales for the three commodities are computed from weekly data prepared by the Export Sales Reporting Division, USDA Foreign Agricultural Service (USDA FAS). Changes in net export sales reflect new foreign purchases as well as cancelled or adjusted purchases for a commodity.[26] Another candidate for the export quantity variable is export shipments. We use sales rather than shipments because export sales are more likely to respond to exchange rate changes. As argued by Bradshaw and Orden (1990) and Ruppel (1987), shipments are better characterized as a logistic variable depending on logistic factors such as transportation costs, the availability of freight space and shipping schedule at the port.

Monthly stocks-to-use ratios are taken from the World Agricultural Supply and Demand Estimates (WASDE) reports. The new marketing year begins in September for corn and soybeans and June for wheat. However, traders typically start focusing on the supply-demand

---

[25] Although hard red winter wheat prices are used in our analysis, prices for different wheat classes in the U.S. move in tandem during most of the time as shown in Janzen et al. (2014). See Bradshaw and Orden (1990) and Bessler and Babula (1987) who also use hard red winter wheat prices.

[26] We use monthly rather than weekly data because weekly sales are frequently affected by revisions and cancelations that introduce noise to the measure. Wheat sales are sales for all wheat. Stocks-to-use ratios and CPI are only available monthly.

conditions for the next marketing year one month before the next marketing year begins (Hu et al., 2017). Hence, we use projected ending stocks and total use for the next marketing year starting August for corn and soybeans, and May for wheat. Otherwise, estimated/projected U.S. ending stocks and total use for the current marketing year are used to calculate the stocks-to-use ratio.

Real commodity prices and export sales are plotted in the two lower panels in figure 4.1. The three real commodity prices seem to follow a similar general pattern with important differences during the period. By comparing commodity prices and the two exchange rate indices, it appears that the value of the dollar is negatively correlated with the level of real commodity price in each market.[27] For example, both exchange rate indices show an increasing trend in the 1995-2000 period as well as after 2015, while commodity prices present a decreasing trend in all markets in the same periods. Export sales mostly reflect a seasonal pattern with occasional negative numbers due to cancelled sales in all the three markets and don’t appear to correlate much with exchange rates.[28]

Figure 4.2 presents stocks-to-use ratios. The stocks-to-use ratio move within the range of 0.05 to 0.25 in the corn and soybean markets, and 0.1 to 0.5 in the wheat market. Low stocks-to-use ratios appear around the 1996 drought and 2007 commodity price boom periods in all the three markets. Further details on the time-series properties of these variables and how the relationship between exchange rates and export prices and sales is affected by the level of stocks-to-use ratio in each market are discussed in the following sections. Summary statistics for all the variables are presented in table 4.1.

[27] Correlation coefficients are provided in supplementary result 1.

[28] Again, correlation coefficients are presented in supplementary result 1.

4.3 Method

Economic theory indicates that the magnitudes of the responses of agricultural export prices and quantities to exchange rates are determined by the excess supply elasticity (Kost, 1976; Chambers and Just, 1979). In storable commodity markets, the theory of competitive storage suggests that the supply curve is elastic when stocks are high and inelastic when stocks are low (Wright, 2011). Hence, the relationship between exchange rates, agricultural export prices and quantities is likely to be non-linear and depend on the level of stocks. To capture the non-linear impact of exchange rate changes on agricultural export prices and sales, we adopt the TVAR model which has several attractive features that suit our analysis. First, the threshold variable, which is the stocks-to-use ratio in our case, is endogenous in the TVAR model. This allows us to attribute non-linear dynamics among the variables to regime switches of the stocks-to-use ratio. Second, changes in parameters across regimes capture the nonlinearities and allow deriving regime-dependent impulse response functions.

Without losing generality, consider a TVAR model with two regimes (j =1,2). Given the vector of endogenous variables (Y) and the threshold variable $w \in Y$, the model can be expressed as follows:

$$Y_t = C_j + \sum_{i=0}^{p} A_{j,i} Y_{t-i} + \varepsilon_{t,j} \qquad (4.1)$$

where j =1 if $w_{t-d} < r$ and j=2, otherwise; r is the value of the threshold; p is the autoregressive lag order; d is the lag of the threshold variable and $1 \leq d \leq p$; the matrix $A_{j,i}$ is a matrix of coefficients for regime j and lag i, and C_j is a vector of constants specific for each regime. Within each regime, the TVAR model is linear, and changes in parameters across regimes allow for nonlinearities.

In this analysis, we select and estimate the TVAR model using the following steps. First, we specify a linear VAR with p autoregressive lags selected using the Bayesian Information Criterion (BIC). Then the sup-LR test proposed by Hansen (1999) and modified by Lo and Zivot (2001) is employed to test for linearity for different values of the delay parameter d. Because there is no prior information about the threshold value r, the nonlinearity test involves estimating the threshold model using all possible threshold values. To avoid overfitting, we restrict the possible threshold value so that at least 20% of the observations are in each regime (i.e. the trimming parameter = 0.2). If the null hypothesis of linearity is rejected, then the TVAR model can be estimated using conditional least squares and the Generalized Impulse Response Functions (GIRFs) are used to study the impulse responses in each regime. Otherwise, a linear VAR is estimated, and results are presented using linear Impulse Response Functions (IRFs).

We use the GIRF developed by Koop, Pesaran and Potter (1996) and Pesaran and Shin (1998). A GIRF represents the difference between the conditional expectation of Y_{t+m} with and without a shock to a variable of interest which can expressed as:

$$GIRF = E[Y_{t+m}|\varepsilon_t, \omega_{t-1}] - E[Y_{t+m}|\omega_{t-1}], \tag{4.2}$$

where *m* is the time horizon, ω_{t-1} represents the history of the series and ε_t is a vector of specific shocks. Typically, the effect of a single shock is examined at a time, so that value of the *i*th element is set to a specific value. The conditional expectation in equation (4.2) and confidence intervals are calculated following the algorithm in Baum and Koester (2011):

1. For each regime with R observations, we select a history ω_{t-1}^r for each possible starting point $r = 1, \dots, R$.
2. Generate sequences of shocks ε_{t+m}^* by randomly drawing bootstrap samples from the estimated residuals of the TVAR model.

3. Simulate the evolution of a variable $y \in Y$ over m periods by iterating the model using ω_{t-1}^{r}, ε_{t+m}^{*} and the estimated coefficients of the TVAR. This yields the baseline path $y_{t+m}(\omega_{t-1}^{r}|\varepsilon_{t+m}^{*})$.
4. Add a one standard deviation positive shock ε_0 to the first residual of the randomly drawn errors to change the path of y. Simulate the evolution of y over m periods to derive the shock path $y_{t+m}(\omega_{t-1}^{r}|\varepsilon_0\,, \varepsilon_{t+m}^{*})$.
5. Repeat step 2 to step 4 500 times to get 500 estimates of the difference between the baseline path $y(\omega_{t-1}^{r}|\varepsilon_{t+m}^{*})$ and the shocked path $y_{t+m}(\omega_{t-1}^{r}|\varepsilon_0\,, \varepsilon_{t+m}^{*})$. Then, calculate the average difference between the baseline path and the shocked path.
6. Repeat step 1 to step 5 using all possible starting points (i.e. R observations) for each regime.
7. The average GIRF for y_{t+m} in a given regime with R observations can be expressed as

$$y_{t+m}(\varepsilon_0\,) = \frac{1}{R}\sum_{r=1}^{R}\frac{y_{t+m}\left(\omega_{t-1}^{r}\middle|\varepsilon_0\,, \varepsilon_{t+m}^{*}\right) - y_{t+m}(\omega_{t-1}^{r}|\varepsilon_{t+m}^{*})}{500}\,, \tag{4.3}$$

and the 95% confidence intervals can be obtained using the quantiles from the same sample.

4.4 Model Specifications and Threshold Tests

We begin by testing for the stationarity of the data. Previous studies typically conclude that real exchange rates and real commodity prices are $I(1)$ and export sales are $I(0)$ series (e.g., Bradshaw and Orden, 1990; Babula, Ruppel and Bessler, 1995). Table 4.2 presents the Augmented Dicky-Fuller (ADF) unit root test results. Consistent with previous findings, ADF tests conclude exchange rates and commodity prices are nonstationary at the 5% significance level. For export sales and stocks-to-use ratios, ADF tests reject the null hypothesis of a unit root in all cases, albeit the significance level is only 10% for the wheat stocks-to-use ratio. Hence, we use levels for export sales and stocks-to-use ratios, and first differences for exchange rates and real prices for all the three markets, which is in line with Babula, Ruppel and Bessler (1995) and Bessler and Babula (1987).[29]

We investigate the effects of importer and exporter exchange rates separately in each market. Four endogenous variables are considered in the VAR/TVAR model: first-differenced exchange rate (either the importer or exporter exchange rate), first-differenced real commodity price, export sales, and the stocks-to-use ratio. When using the TVAR, ordering of the variables is irrelevant since GIRF analysis does not require the orthogonalization of shocks and is invariant to ordering (Pesaran and Shin, 1998). For linear VAR models, we order exchange rate changes first as our objective is to investigate the impact of exchange rates on export prices and sales. As stocks-to-use ratios reflect the overall market supply-demand conditions, they are likely affected by all the other variables and therefore appear last in the ordering. We tried using different orders for sales and price and found results are visibly the same. Hence, we present IRFs with real commodity price ordered before export sales.[30]

---

[29] We tested for the cointegration relationship between the exchange rates and the real price series and found they are cointegrated in soybean and wheat markets. However, first differences are used for these variables, since estimation of the VAR and TVAR using their levels leads to a nonstationary model as reflected by the impulse response functions. To save space, these results are not presented but are available.

[30] The ordering we use is in line with Babula, Ruppel and Bessler (1995) and Bessler and Babula (1987). We also considered using GIRFs for VAR models. As shown in Pesaran and Shin (1998), when shocks occur in the first variable in a VAR system which is the exchange rate in our case, generalized impulse responses and orthogonalized impulse responses are the same.

The number of lags (q) in the VAR model is determined by the Bayesian Information Criterion (BIC) which strongly penalizes the number of coefficients estimated in the model relative alternative information criteria such as AIC. The BIC selects 1 lag in all cases and the delay parameter d is set equal to 1 as it cannot exceed q. The sup-LR test proposed by Hansen (1999) and modified by Lo and Zivot (2001) is employed to test for the threshold. To avoid overfitting, we require that at least 20% of the observations (72 observations) are in each regime. We consider the possibility of the existence of two and three regimes, and the test results are presented in table 4.3.

In the corn market, the sup-LR test fails to reject the null hypothesis of linearity when importer exchange rates are used. However, the linearity hypothesis is rejected when using corn exporter exchange rates, albeit the significance level is only 10% when testing against the presence of one threshold. In contrast, significant threshold effects are found using importer exchange rates in the soybean market, but the null hypothesis of linearity is not rejected using soybean exporter exchange rates. In the wheat market, significant threshold effects are found using both importer and exporter exchange rates.

Based on the sup-LR test results, we estimate a linear VAR model if the null hypothesis of linearity is not rejected. Otherwise, a TVAR model is employed to study the impact of exchange rates on export prices and sales in different stocks-to-use ratio regimes. One exception is the wheat market using exporter exchange rates. While exporter exchange rates have the expected negative influence on wheat export prices, unexpected positive effects on export sales were found in certain regimes when a TVAR model was used. A possible explanation for the unexpected positive impact on wheat export sales is that wheat quality and end use characteristics are different between the U.S. and other major wheat exporting countries

(O'Brien and Olson, 2014). As wheat varieties produced in other major exporting countries are not perfect substitutes for U.S. wheat, their currencies may not have expected effects on U.S. wheat export sales. Therefore, we present results from a linear VAR model for wheat when exporter exchange rates are used.

Later, we discuss with more details that the absence of the threshold effect of the importer exchange rate in the corn market is likely because only a small proportion of corn is exported during the sample period and effects of exchange rates (both importer and exporter exchange rates) on corn export prices and sales are either small or non-significant. For exporter exchange rates, the absence of the threshold effect in soybean and wheat markets is due to the low substitutability between commodities produced in the U.S. and other exporting countries, not only in terms of quality but also time.

To check the sensitivity of our results to the number of lags used in the VAR and TVAR, longer lags (up to 3) were tried. We found that the results are generally the same when using longer lags, except for the VAR model using soybean exporter change rates. When one lag is used, an increase in the soybean exporter exchange rate has an unexpected positive impact on soybean export sales. However, the impact becomes statistically indistinguishable from 0 when longer lags are used. As a result, we include 2 lags in the VAR model using soybean exporter exchange rates. Model specifications adopted for each market using importer and exporter exchange rates are summarized in table 4.4.

Although more significant sup-LR test statistics are found when the presence of two thresholds is tested, we present the GIRFs from TVAR models with one threshold. When multiple thresholds are used, responses to exchange rate shocks in the middle regime are often undistinguishable from responses in one of the two extreme regimes. The threshold estimates and the number of observations in high and low stocks-to-use ratio regimes are presented in the last two columns in table 4.4. Figure 4.2 shows, for each market, the stocks-to-use ratio together with the threshold value indicated by a continuous horizontal line and the sample mean indicated by a dashed horizontal line. All the threshold values are well below their sample means, particularly in soybean and wheat markets. Overall, at least 25% of the observations (94 months) are in the low stocks-to-use ratio regime in the markets studied. The 1996-1997 severe drought and the 2006-2009 period which includes the commodity price boom/bust, financial crisis and biofuel expansion (i.e. Renewable Fuel Standard) are in the low stocks-to-use ratio regime in all the three markets. While the wheat stocks-to-use ratio remains above its threshold value in the recent decade, stocks-to-use ratios below the threshold also appear during the 2012-13 North American drought in the soybean and corn markets.

4.5 Dynamic Responses

Dynamic responses of real commodity price changes to importer and exporter exchange rate shocks are presented in figure 4.3. Dynamic responses of export sales to exchange rate shocks are presented in figure 4.4. The exchange rate shocks are positive and one-standard deviation in magnitude. Importer and exporter exchange rate shocks are presented in the left and right panels, respectively. Linear IRFs are presented using black lines and GIRFs for high and low stocks-to-use ratio regimes are presented in blue and red lines, respectively.

4.5.1 Price Responses

As shown in figure 4.3, real commodity prices are not only affected by importer exchange rates, but also influenced by currencies of major competitors in export markets. In all cases, an increase in the value of the dollar leads to a significant decrease in the export price. However, the negative impacts of both importer and exporter exchange rates on real commodity prices are short lived and usually become indistinguishable from zero after 2 to 3 months. As expected, in cases where threshold effects are found, namely the corn market using exporter exchange rates and soybean and wheat markets using importer exchange rates, the responsiveness of the real commodity price change is higher in the low stocks-to-use ratio regime than in the high stocks-to-use ratio regime.

To better quantify the economic values of these responses, we present on a percent basis the dynamic responses of export prices to a change in exchange rate in table 4.5 for the first 3 months where significant responses are found. Since real prices enter the model in first differences, the percentage change in the real price is computed by dividing the cumulative impulse response of the real commodity price by its sample mean. Similarly, the percentage change in the exchange rate is computed diving the cumulative exchange rate change by its sample mean. Specifically, the dynamic percentage change of the export price (m_h^{PE}) at time horizon h is computed as

$$m_h^{PE} = \frac{\Delta P_h}{\Delta E_h} \cdot \frac{\bar{E}}{\bar{P}} \tag{4.4}$$

where ΔP_h and ΔE_h are computed by adding up the corresponding impulse responses of real commodity price changes (figure 4.3) and the impulse responses of exchange rate changes[31] from horizon 1 to h (h=1,2,3), respectively; $\bar{P}$ and $\bar{E}$ are the real commodity price and exchange rate sample means presented in table 4.1. In addition, we compute the 3-month cumulative economic impact (in cents per metric ton) by multiplying the percentage change for the third period by the sample mean of the real commodity price (i.e. $\frac{\Delta P_3 \cdot \bar{E}}{\Delta E_3}$).

[31] To save space, these are not presented.

As shown in table 4.5, effects of exchange rates are smaller in the corn market than in the other two markets, which is likely because the U.S. corn market is less dependent on exports. Specifically, U.S. corn exports have been declining and only account for about 7% - 20% of the production during the sample period. While the U.S. corn production is mostly for domestic uses, particularly for ethanol production, U.S. soybean and wheat markets are more dependent on exports as more than 40% of U.S. soybeans and wheat are exported during most of the time in the sample period.

Consistent with the differences in impulse responses across regimes, dynamic percentage changes in export prices are higher in the low stocks-to-use ratio regime. Within each market, on balance, real export prices are more responsive to importer exchange rates than exporter exchange rates. In particular, while exporter exchange rates are associated with percentage changes that are less one in all cases, greater-than-unit percentage changes in response to importer exchange rates are found in the low stocks-to-use ratio regime in soybean and wheat markets.

4.5.2 Export Sales Responses

While both importer and exporter exchange rates have significant impacts on real commodity prices in all markets, they do not always affect export sales. Responses of export sales to one standard deviation positive shocks to the importer and exporter exchange rate changes are presented in the left and right panels in figure 4.4, respectively.

As shown in the right panels in figure 4.4, importer exchange rate changes have statistically significant effects on export sales in soybean and wheat markets, but not in the corn market. Specifically, in both soybean and wheat markets, an increase in the importer exchange rate leads to a small initial decline in export sales in both regimes. While initially the decline has a small magnitude, in the second month the magnitude is larger, especially in the low stocks-to-use ratio regime. Then, the response of export sales declines and becomes insignificant after about 4 months in both regimes.

Higher responses of export sales for future delivery are found in the low stocks-to-use ratio regime at longer horizons are likely due to the differences in market participants' expectations under different stocks-to-use and price conditions, and the ability of importers to cancel export contracts. As shown in figure 4.1, importer exchange rates present strong cyclic behavior. When stocks-to-use ratios are low and prices are more responsive to exchange rate changes, importers may anticipate higher import costs in the future and search for alternative sources. Export contracts provide importers this flexibility. In contrast, when stocks-to-use ratios are high, export prices (import costs) are low. Export sales are less affected by exchange rate changes because exchange rates have limited impacts on export prices at longer horizons as we show in figure 4.3.

In contrast to importer exchange rate effects, exporter exchange rate shocks only show significant influence on export sales in the corn market, but not in soybean and wheat markets. In the soybean market, major exporters, namely Brazil, Argentina, and Paraguay, have different marketing seasons than the U.S. During the U.S. soybean export seasons, U.S. soybean export sales are not likely to be affected by an appreciation of the dollar because south American soybean supplies are limited. For wheat, as we discussed, the effects of exporter exchange rates are limited by the relatively low degree of substitutability between the U.S. wheat and wheat produced in other exporting countries. For example, lower quality feed grade wheat typically makes up a large proportion of export sales for Russia and Ukraine which account for about 45%

of the trade weights in the exporter exchange index, while U.S. wheat exports are mostly high-quality milling grade (O'Brien and Olson 2014).

Compared to soybeans and wheat, corn exports are more substitutable between the U.S. and other major corn exporters in terms of quality and marketing seasons. The two Northern Hemisphere corn exporters Russia and Ukraine have similar marketing seasons with the U.S. Also, the largest U.S. corn export competitor – Brazil –has second-crop corn exported from September to January which overlaps with the U.S. marketing seasons. As shown in Allen and Valdes (2016), corn exports from Brazil are greater in September-January than in the prior May-July period in recent years because of increased second-crop corn. Hence, as shown in the upper right panel in figure 4.4, an increase in the value of the dollar relative to currencies of other major corn exporting countries will result in a significant decrease in the U.S. corn export sales. In addition, the responsiveness of U.S. corn export sales is higher in the low stocks-to-use ratio regime compared to the high stocks-to-use ratio regime, which is consistent with the threshold effects found in the corn export price-exchange rate relationship.

Again, to quantify the economic values of these responses, we also present a dynamic percentage change of export sales to a shock to exchange rates as well as their 3-month cumulative economic impacts (in million metric tons) in table 4.6. The dynamic percentage change of export sales (m_t^{SE}) at time horizon *t* is computed as

$$m_t^{SE} = \frac{\Delta S_h}{\Delta E_h} \cdot \frac{\bar{E}}{\bar{S}}. \tag{4.5}$$

As export sales are in levels, ΔS_h is the corresponding impulse response of export sales (figure 4.5) at horizon *h* and $\bar{S}$ is the sample mean of export sales. The cumulative economic impact of the exchange rate change is computed multiplying the 3-month percentage change by the sample mean of export sales (i.e. $\frac{\Delta S_3 \cdot \bar{E}}{\Delta E_3}$).

Consistent with findings in table 4.5, dynamic percentage changes in export sales are smaller in the corn market than the other two markets. In particular, the cumulative economic impacts of both importer and exporter exchange rates are only marginal in the corn market. The limited impacts of exchange rates on corn export sales reflect the relatively small size of U.S. corn exports. In soybean and wheat markets where impulse responses of export sales to exporter exchange rate shocks are not significant, the 3-month cumulative economic values of exporter exchange rate impacts are close to zero. However, importer exchange rates in soybean and wheat markets are associated with greater dynamic percentage changes and cumulative economic impacts, particularly in the low stocks-to-use ratio regime. Consistent with impulse responses, in both soybean and wheat markets, the dynamic importer exchange rate-export sales effect has a small initial value that peaks in the second period. However, the impact quickly dies out after the peak. This pronounced shorter-term importer exchange rate impact on export sales may be related to large order cancelations as the cost to importers is heighted with an increase in the value of the dollar. Notice that an exchange rate shock in the low-stock period can often occur when prices are already high and availability is limited. The three-month cumulative economic impact indicates that on average a 1% increase in the importer exchange rate causes a reduction of 0.17 and 0.07 million metric tons (or 6.25 and 2.57 million bushels) in soybean and wheat export sales under low stocks-to-use conditions, respectively.

4.6 Conclusions

This paper studies for the first time the impact of the level of grain stocks on the responsiveness of export sales and prices to changes in the exchange rates in the U.S. corn, soybean and wheat markets. The theory of agricultural commodity excess supply-demand model (Kost, 1976; Chambers and Just, 1979; Hatzenbuehler, Abbott and Foster, 2016) predicts that the elasticity of price with respect to the exchange rate increases when the domestic demand becomes more inelastic. Given that market demand is more inelastic in low stocks-to-use conditions in storable commodity markets (Wright, 2011), we expect that the responsiveness of real agricultural export prices to exchange rates is greater when the stock-to-use ratio is low, particularly in the short run. When the stocks-to-use ratio is low, export sales which reflect future delivery are more likely affected by exchange rates at longer horizons as importers adjust their purchases in response to higher costs.

For the first time, we investigate and show nonlinear responses of real commodity export prices and sales to exchange rate shocks using TVARs where the stocks-to-use ratio is used as the endogenous threshold variable. In addition to the import-weighted exchange rates that are commonly considered in the literature, we investigate the effects of major global exporters' currencies on U.S. export prices and sales by building export-weighted exchange rate indices based on U.S. export competitors' export shares for each commodity.

Overall, consistent with Chamber and Just (1981), our results suggest that the dynamic exchange rate impacts on agricultural exports are complex. In particular, we show that the effects of exchange rates differ across markets as well as between the import and export exchange rates within each market. Primarily, the magnitudes of exchange rate effects on real agricultural commodity export prices and sales are determined by the market dependence on exports. During the sample period of 1990 to 2019, effects of both importer and exporter exchange rates on real price and export sales are either insignificant or small in their economic values in the corn market where export share of production is relatively small.

In soybean and wheat markets that are more exports-dependent, both importer and exporter exchange rates have significant and negative impacts on real export prices. However, export sales are only significantly affected by importer exchange rates. The effects of exporting countries' currencies on U.S. export sales in soybean and wheat markets are likely limited by the differences in marketing seasons and crop classes between the U.S. and other major exporters.

This paper demonstrates that changes in the responsiveness of export prices and sales to exchange rates can be explained by the changes in the underlying market fundamental conditions. The results help to explain the mixed evidence on the effects of exchange rates on export prices and quantities in the literature. Failure to account for nonlinear threshold effects may be part of the explanation. We show that when the stocks-to-use ratio is low exchange rate effects on export prices and sales is be more pronounced in exports-orientated grain markets. Specifically, the responsiveness of real export prices and sales to importer exchange rate changes are greater in the low stocks-to-use ratio regime in soybean and wheat markets. Similar threshold effects are also present in the effects of corn exporter exchange rates on real corn prices and export sales, albeit the economic values in different regimes of stocks-to-use ratio do not seem to be largely different. In contrast, in periods of high stocks-to-use ratios the effects of changing exchange rates are non-existent or sharply muted. Failure to consider these changing responses likely masked the relevant exchange rate effect.

While our analysis shows the importance of storage in grain markets in determining price and export sales responses to exchange rates fluctuations, it is also possible that other factors like the Renewable Fuel Standard (RFS), China's Soybean reserve policy, economic recession and quantitative easing also affect the exchange rate-export relationship. However, much of these events overlap with periods of low stocks. Additionally, storable commodity markets are more vulnerable to issues like extreme weather, increased biofuel and export demand when stocks are low (Wright, 2011). Due to multicollinearity and degrees of freedoms, it is always difficult to incorporate multiple factors to model exchange rate effects on agricultural exports (Chambers and Just, 1981). Further research may consider modeling these factors in a structural framework to clearly disentangle their effects. Alternatively, one may consider using time-varying parameter models to capture the exchange rate influence on agricultural exports under different market conditions.

The results provide important implications for policy makers who intend to use a weak dollar policy to boost agricultural exports and increase farmers' welfare. Considering stocks-to-use ratios for major U.S. grain and oil seed markets are at a relatively high level in recent years, a weak dollar policy may not be able to dramatically boost export prices and sales, at least for the markets studied in this article.

4.7 Tables and Figures

Table 4.1 Monthly Summary Statistics, January 1990-December 2019

	Importer Exchange Rate	Exporter Exchange Rate	Price ($/Metric Ton)	Export Sales (Million Metric Tons)	Stocks-to-Use Ratio
			Corn		
Min	90.47	53.00	43.46	-0.75	0.04
Max	120.98	282.57	145.68	10.03	0.27
Median	103.66	172.39	70.42	3.27	0.14
Mean	104.03	159.00	75.70	3.33	0.14
Standard Deviation	6.85	57.56	22.74	1.74	0.05
			Soybeans		
Min	91.92	78.95	103.78	-0.61	0.04
Max	130.51	266.49	297.50	9.51	0.25
Median	106.93	124.95	168.87	1.60	0.11
Mean	107.59	127.07	173.35	2.08	0.11
Standard Deviation	9.06	34.11	40.61	1.77	0.05
			Wheat		
Min	87.33	69.97	50.61	-0.19	0.10
Max	121.90	174.55	189.19	6.67	0.56
Median	104.15	117.69	85.74	1.94	0.29
Mean	102.95	116.85	89.24	2.02	0.31
Standard Deviation	9.41	21.69	26.14	1.03	0.10

Note: The importer (exporter) exchange rate is the trade-weighted exchange rate. Prices and export sales are measured at NOLA. Stocks-to-use Ratio are USDA's national estimate.

Table 4.2 Augmented Dickey–Fuller Tests, January 1990-December 2019

	Importer Exchange Rate	Exporter Exchange Rate	Price	Export Sales	Stocks-to-Use Ratio
Corn	-1.76	-2.70	-2.51	-6.44**	-4.21**
	(0.68)	(0.28)	(0.36)	(0.00)	(0.00)
Soybeans	-2.42	-2.11	-2.52	-12.16**	-2.98**
	(0.40)	(0.53)	(0.36)	(0.00)	(0.04)
Wheat	-1.70	-3.21*	-2.84	-9.22**	-3.27*
	(0.70)	(0.08)	(0.22)	(0.00)	(0.06)

Note: ADF tests are all specified with a constant and lags are selected based on the Akaike information criterion (AIC). *P*-values for the ADF tests are presented in parentheses. * and ** indicate significance at the 10% and 5% level, respectively.

Table 4.3 Sup-LR Linearity Tests, January 1990-December 2019

	Lags	Delay Parameter	Sup-LR test	
	(q)	(d)	1 Threshold	2 Thresholds
Importer Exchange Rate				
Corn	1	1	28.544	57.338
			(0.370)	(0.410)
Soybeans	1	1	36.591*	71.374**
			(0.096)	(0.046)
Wheat	1	1	39.669*	71.572**
			(0.068)	(0.046)
Exporter Exchange Rate				
Corn	1	1	35.442*	69.764***
			(0.100)	(0.000)
Soybeans	1	1	39.473	76.108
			(0.112)	(0.260)
Wheat	1	1	40.208*	82.947***
			(0.073)	(0.013)

Note: The null hypothesis states that the relationship is linear. *P*-values are presented in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively.

Table 4.4. VAR and TVAR Model Specifications, January 1990-December 2019

	Model	Lags (q)	Threshold	Number of Observations (Low/High Regime for TVAR)
Importer Exchange Rate				
Corn	VAR	1	none	359
Soybeans	TVAR	1	0.07	94/265
Wheat	TVAR	1	0.22	122/237
Exporter Exchange Rate				
Corn	TVAR	1	0.13	150/209
Soybeans	VAR	2	none	359
Wheat	VAR	1	none	359

Note: none indicates no threshold and a linear VAR is estimated.

Table 4.5 Dynamic Price-Exchange Rate Responses, January 1990-December 2019

Model	Exchange Rate	Dynamic Response 1-Month	2-Month	3-Month	Cumulative Impact (cents/metric ton)
		Corn			
VAR	Importer Exchange Rate	-0.45	-0.54†	-0.61†	-0.46
TVAR	Exporter Exchange Rate	-0.06/-0.00	-0.16/-0.04	-0.20/-0.05	-0.15/-0.04
		Soybeans			
TVAR	Importer Exchange Rate	-1.75/0.32	-2.13/-0.40	-2.35†/-0.39†	-4.08/-0.67
VAR	Exporter Exchange Rate	-0.08	-0.09†	-0.10†	-0.18
		Wheat			
TVAR	Importer Exchange Rate	-0.38/-0.55	-1.25/-0.4†	-1.76/-0.37†	-1.57/-0.33
VAR	Exporter Exchange Rate	-0.44	-0.83†	-0.97†	-0.86

Note: The dynamic responses are in percentages, except for the cumulative impact. The dynamic responses and cumulative impacts for low and high regimes are presented before and after the slash, respectively. † indicates the computation involves using insignificant impulse response(s) of real price changes.

Table 4.6 Dynamic Export Sales-Exchange Rate Responses, January 1990-December 2019

Model	Exchange Rate	Dynamic Responses 1-Month	2-Month	3-Month	Cumulative Impact (million metric tons)
		Corn			
VAR	Importer Exchange Rate	1.83†	-0.52†	-0.51†	-0.02
TVAR	Exporter Exchange Rate	-2.78/-1.26	-0.53/-0.34	-0.15/-0.08	-0.00/0.00
		Soybeans			
TVAR	Importer Exchange Rate	-0.93/-1.08	-10.11/-3.75	-8.35/-2.07	-0.17/-0.04
VAR	Exporter Exchange Rate	0.57†	0.580†	0.60†	0.01
		Wheat			
TVAR	Importer Exchange Rate	-1.23†/0.41	-7.54/-3.84	-3.47/-1.40	-0.07/-0.03
VAR	Exporter Exchange Rate	1.40†	1.18†	1.14†	0.02

Note: The dynamic responses are in percentages, except for the cumulative impact. The dynamic responses and cumulative impacts for low and high regimes are presented before and after the slash, respectively. † indicates the computation involves using insignificant impulse response(s) of export sales.

Figure 4.1 Importer and Exporter Real Exchange Rates, Real Commodity Prices, and Net Export Sales for Corn, Soybeans, and Wheat, January 1990-December 2019

Note: Horizontal dash lines indicate sample means.

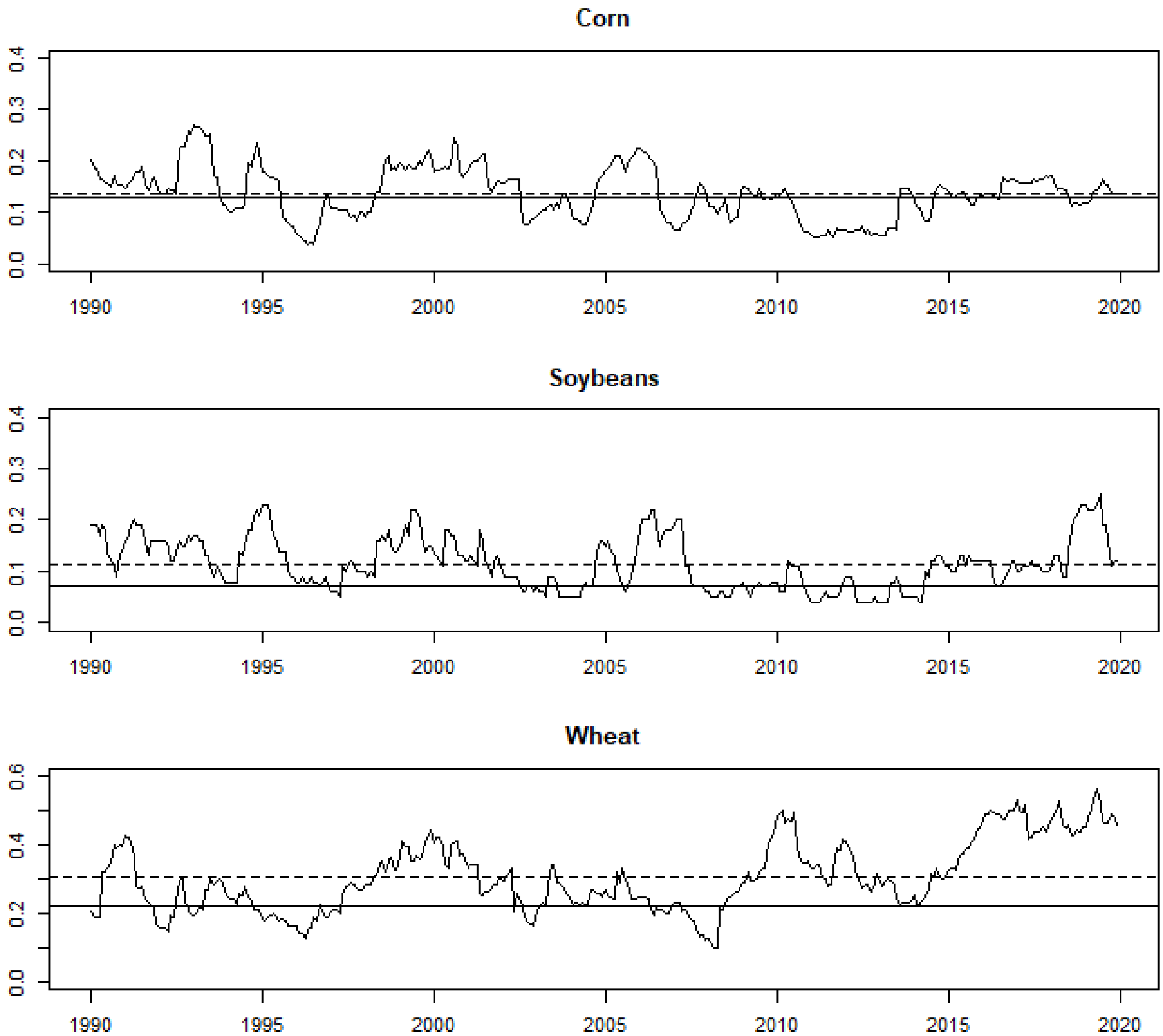


Figure 4.2 Stocks-to-Use Ratios, January 1990-December 2019

Note: Horizontal solid lines indicate threshold values identified using a threshold vector autoregression (TVAR) model. The threshold value in the corn market is determined by the corn exporter exchange rate, and threshold values in soybean and wheat markets are determined by the corresponding importer exchange rates. Horizontal dashed lines indicate sample means.

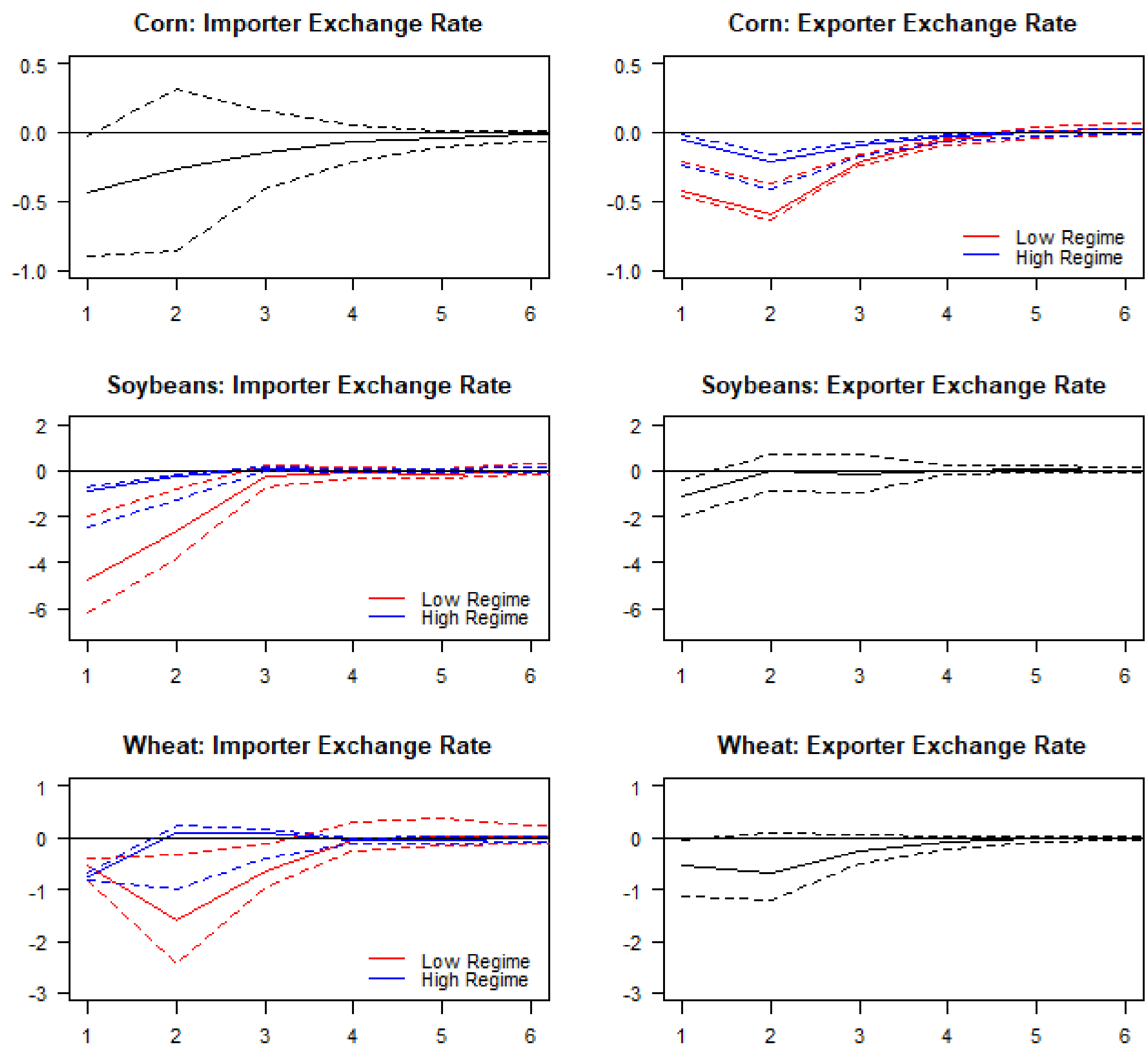


Figure 4.3 Responses of Real Commodity Price Changes to Importer and Exporter Exchange Rate Shocks

Note: Linear IRFs are represented by black lines. GIRFs for high and low stocks-to-use regimes are represented by the blue and red lines, respectively. 95% confidence bands are indicated by dashed lines.

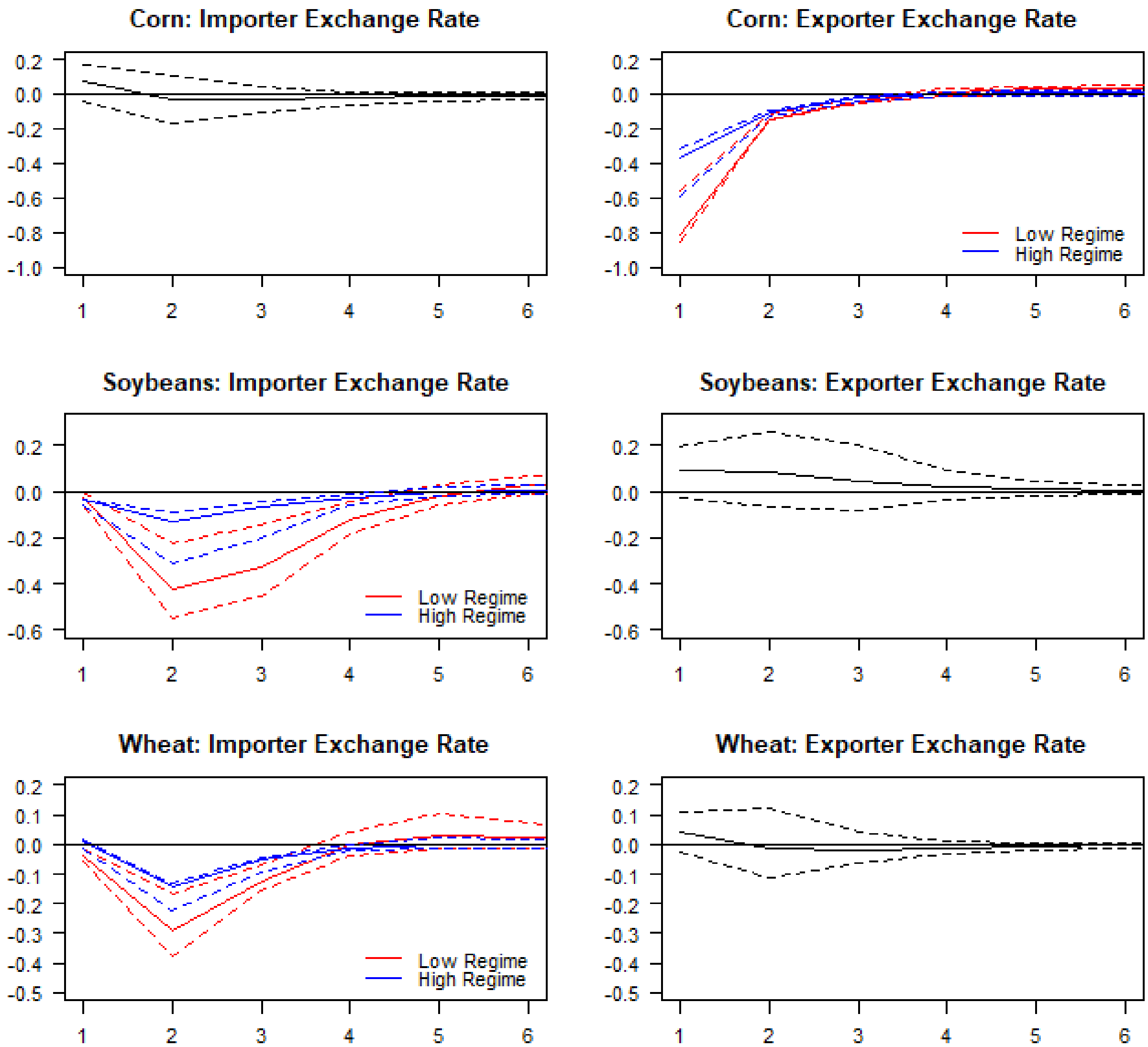


Figure 4.4 Responses of Export Sales to Importer and Exporter Exchange Rate Shocks

Note: Linear IRFs are represented by black lines. GIRFs for high and low stocks-to-use regimes are represented by blue and red lines, respectively. 95% confidence bands are indicated by dashed lines.

4.8 Supplementary Results

Supplementary Result 1

Pearson Correlations, January 1990-December 2019

Supplemental Table 4.1 Corn Market

	Importer Exchange Rate	Exporter Exchange Rate	Price	Sales	Stocks-to-Use Ratio
Importer Exchange Rate	1.000	0.401	-0.562	0.112	0.205
Exporter Exchange Rate	0.401	1.000	-0.003	-0.051	-0.255
Price	-0.562	-0.003	1.000	-0.253	-0.673
Sales	0.112	-0.051	-0.253	1.000	0.197
Stocks-to-Use Ratio	0.205	-0.255	-0.673	0.197	1.000

Supplemental Table 4.2 Soybean Market

	Importer Exchange Rate	Exporter Exchange Rate	Price	Sales	Stocks-to-Use Ratio
Importer Exchange Rate	1.000	0.077	-0.058	-0.011	-0.013
Exporter Exchange Rate	0.077	1.000	-0.127	0.024	-0.002
Price	-0.058	-0.127	1.000	0.068	-0.018
Sales	-0.011	0.024	0.068	1.000	-0.017
Stocks-to-Use Ratio	-0.013	-0.002	-0.018	-0.017	1.000

Supplemental Table 4.3 Wheat Market

	Importer Exchange Rate	Exporter Exchange Rate	Price	Sales	Stocks-to-Use Ratio
Importer Exchange Rate	1.000	0.326	-0.649	-0.018	0.155
Exporter Exchange Rate	0.326	1.000	-0.450	-0.104	0.526
Price	-0.649	-0.450	1.000	0.085	-0.613
Sales	-0.018	-0.104	0.085	1.000	-0.099
Stocks-to-Use Ratio	0.155	0.526	-0.613	-0.099	1.000

CHAPTER 5:

CONCLUSIONS

In recent years, agricultural commodity markets have been affected by several large changes. In agricultural commodity futures markets, the move to electronic trading has reshaped the market in many aspects including the way that fundamental information is being reflected, new trading technologies being used by high frequency traders, and changes in the interactions between traditional commercial traders and high frequency traders. While the transition to electronic trading has caused a dramatic change in agricultural commodity markets, the important role of market fundamentals is not weakened. In this context, we investigate three aspects of agricultural commodity markets to improve our understanding of how modern agricultural commodity markets are affected by market microstructure and market fundamentals.

In the first essay, we measure the relative importance of nearby and deferred contracts in price discovery. The analysis is performed in the corn and live cattle futures markets using intraday data for the 2008-2015 period. The results show that price discovery is dominated by the nearby contract in the storable corn market than the non-storable live cattle market, which provides the empirical support for Working's theory of price of storage. Nevertheless, deferred contracts still play a significant role in the price discovery process. This demonstrates the importance of futures' forward pricing role in price discovery as argued by Tomek (1997). In both the corn and live cattle markets, the nearby contract loses its leadership in price discovery when its volume share dips below 50%. In the corn market, this typically occurs one week before the maturity month. In the live cattle market, it is about two weeks before the maturity month. Based on these findings, we recommend rolling to the next nearby contract when it achieves more than 50% of the volume share. The regression analysis shows that the share of price discovery along the forward curve is strongly correlated with trading volume and nonlinearly correlated with time to expiration in both markets. In the corn market, USDA announcements, inverted markets, price declines and commodity index rolls also are statistically related to the relative importance that nearby and deferred contracts play in price discovery. However, they do not have significant effects in the live cattle market, which likely reflects differences in liquidity and commodity storability between corn and live cattle futures.

In the second essay, we investigate how algorithmic quotations (AQ) affect pricing efficiency, short-term volatility, and liquidity in corn, soybean, and live cattle futures markets using CME's limit order book data. Overall, results show that more intensive AQ is beneficial to market quality as AQ improves the efficiency of prices, mitigates short-term volatility and reduces the costs of immediacy, although the influence can vary across the markets. In addition, by decomposing effective spreads into the realized spread and price impact components, we show lower costs of immediacy is mainly a result of reduced adverse selection costs. Additionally, in the corn market, liquidity provider revenues increase with heightened AQ activity. The increased liquidity provider revenue effect points to a tradeoff between the dimensions of market quality, and the need for continued monitoring of algorithmic trading activity in agricultural commodity futures markets.

In the third essay, we use Threshold Vector Autoregressive (TVAR) models to investigate how the responses of export prices and sales to exchange rate movements are influenced by the level of the stocks-to-use ratio in the corn, soybean, and wheat export markets for the period of January 1990-December 2019. The results show that the dynamic exchange rate impacts on agricultural exports are complex as exchange rate effects differ across markets as well as between the import and export exchange rates within each market. The effects of both importer and exporter exchange rates on corn export prices and sales are either insignificant or small in their economic value. In the more export-oriented soybean and wheat markets, both increases in the value of importer and exporter exchange rates have significant and negative impacts on export prices. However, soybean and wheat export sales are only significantly affected by importer exchange rates but not by exporter exchange rates. The effects of exporter exchange rates in soybean and wheat markets are likely limited by the differences in marketing seasons and crop classes between the U.S. and other exporters. In addition, the effects of importer exchange rates in soybean and wheat markets on real prices and export sales also differ across stocks-to-use ratio regimes. In both markets, the responses of real export prices and sales to importer exchange rate changes are greater in the low stocks-to-use ratio regime. Similar threshold effects are also found in the effects of corn exporter exchange rates on real corn prices and export sales, albeit the effects do not seem to be largely different in economic value. The results in this essay provide important implications for policymakers and market participants that the underlying market fundamental conditions need to be considered for accurate evaluations and forecasts on exchange rate effects in agricultural export markets.

Overall, the three studies reveal the important message that the structural change caused by electronic high frequency trading has not changed the underlying economic logic of how agricultural commodity markets function, as market fundamentals still play a determinate role. In particular, the first essay shows price discovery between intraday prices along the forward curve can be affected by inverse carry charges, supply-demand information in USDA reports, and severe droughts. The second essay shows algorithmic trading activity improves market quality, but results also show pricing efficiency, volatility, and liquidity are affected by the underlying market conditions. The third essay indicates market fundamentals in a nonlinear manner affect the exchange rate-export relationship, which has limited empirical evidence in the literature.

This dissertation contributes to the understanding of agricultural commodity markets in several ways. First, it provides the first empirical evidence of the dynamic price discovery relationship along the futures forward curve. Second, it is the first study that provides directly identification of the effects of algorithmic quoting activity on agricultural commodity market quality. Third, it assesses for the first time how the level of the stocks-to-use ratio affect the impacts of exchange rates on both export prices and sales in major U.S. agricultural export markets.

Future research may consider to extent the analysis in this dissertation in several directions. First, this dissertation has focused on the aggregate effects of algorithmic activity. Future work could consider exploring the heterogeneity of the effects of different types of algorithms. Second, while the dissertation has explored the role of stocks-to-use ratio in the exchange rate effects on agricultural export prices and sales, other factors can influence the impacts of exchange rates as well. Further efforts could investigate how monetary and agricultural policy changes affect the influence of exchange rate movements on agricultural export markets.

REFERENCES

Adjemian, M.K. (2012). Quantifying the WASDE announcement effect. *American Journal of Agricultural Economics*, *94*(1), 238-256.

Ahumada, H., & Cornejo, M. (2016). Out-of-sample testing price discovery in commodity markets: the case of soybeans. *Agricultural Economics*, *47*(6), 709-718.

Allen, E., & Valdes, C. (2016). Brazil's corn industry and the effect on the seasonal pattern of US corn exports. United States Department of Agriculture, Economic Research Service

Arnade, C., & Hoffman, L. (2019). Determination of the international maize price: What is the growing role of Brazil and Ukraine? *Agricultural Economics*, *50*(6), 735-747.

Aulerich, N. M., Irwin, S. H., & Garcia, P. (2014). Bubbles, food prices, and speculation: Evidence from the CFTC's daily large trader data files. In: Chavas J.P., Hummels D. L., Wright B. (Eds.), *Economics of Food Price Volatility*, University of Chicago Press, Chicago, IL. pp. 211-253.

Babula, R. A., Ruppel, F. J., & Bessler, D. A. (1995). US corn exports: The role of the exchange rate. *Agricultural Economics*, *13*(2), 75-88.

Baillie, R.T., Booth, G.G., Tse, Y., & Zabotina, T. (2002). Price discovery and common factor models. *Journal of Financial Markets*, *5*(3), 309-321.

Bandi, F. M., & Russell, J. R. (2008). Microstructure noise, realized variance, and optimal sampling. *Review of Economic Studies*, *75*(2), 339-369.

Baum A, Koester GB. (2011). The impact of fiscal policy on economic activity over the business cycle: Evidence from a threshold VAR analysis. Discussion Paper Series 1: Economic Studies No. 03/2011, Deutsche Bundesbank.

Bessembinder, H. (2003). Trade execution costs and market quality after decimalization. *Journal of Financial and Quantitative Analysis*, *38*(4), 747-777.

Bessler, D. A., & Babula, R. A. (1987). Forecasting wheat exports: Do exchange rates matter? *Journal of Business & Economic Statistics*, *5*(3), 397-406.

Biais, B., Declerck, F., & Moinas, S. (2016). Who supplies liquidity, how and when? Working Paper.

Bradshaw, G. W., & Orden, D. (1990). Granger causality from the exchange rate to agricultural prices and export sales. *Western Journal of Agricultural Economics*, *15*(1), 100-110.

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, *47*(5), 1287-1294.

Brogaard, J., Carrion, A., Moyaert, T., Riordan, R., Shkilko, A., & Sokolov, K. (2018). High frequency trading and extreme price movements. *Journal of Financial Economics*, *128*(2), 253-265.

Brogaard, J., Hendershott, T., & Riordan, R. (2014). High-frequency trading and price discovery. *Review of Financial Studies*, *27*(8), 2267-2306.

Bruno, V. G., Büyükşahin, B., & Robe, M. A. (2017). The financialization of food? *American Journal of Agricultural Economics*, *99*(1), 243-264.

Cafiero, C., & Wright, B. D. (2011). The empirical relevance of the competitive storage model. *Journal of Econometrics*, *162*(1), 44-54.

Campbell, J. Y., Lo, A., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.

Carrion, A. (2013). Very fast money: High-frequency trading on the NASDAQ. *Journal of Financial Markets*, *16*(4), 680-711.

Chaboud, A. P., Chiquoine, B., Hjalmarsson, E., & Vega, C. (2014). Rise of the machines: Algorithmic trading in the foreign exchange market. *Journal of Finance*, *69*(5), 2045-2084.

Chambers, R. G. (1984). Agricultural and financial market interdependence in the short run. *American Journal of Agricultural Economics*, *66*(1), 12-24.

Chambers, R. G., & Just, R. E. (1979). A critique of exchange rate treatment in agricultural trade models. *American Journal of Agricultural Economics*, *61*(2), 249-257.

Chambers, R. G., & Just, R. E. (1981). Effects of exchange rate changes on US agriculture: A dynamic analysis. *American Journal of Agricultural Economics*, *63*(1), 32-46.

Chambers, R. G., & Just, R. E. (1982). An investigation of the effect of monetary factors on agriculture. *Journal of Monetary Economics*, *9*(2), 235-247.

Charles, A., & Darné, O. (2009). Variance-ratio tests of random walk: An overview. *Journal of Economic Surveys*, *23*(3), 503-527.

Chen, Y.L., & Tsai, W.C. (2017). Determinants of price discovery in the VIX futures market. *Journal of Empirical Finance*, *43*(2017), 59-73.

Conrad, J., Wahal, S., & Xiang, J. (2015). High-frequency quoting, trading, and the efficiency of prices. *Journal of Financial Economics*, *116*(2), 271-291.

Couleau, A., Serra, T., & Garcia, P. (2019). Microstructure noise and realized variance in the live cattle futures market. *American Journal of Agricultural Economics*, *101*(2), 563-578.

Covey, T., & Bessler, D.A. (1995). Asset storability and the informational content of inter-temporal prices. *Journal of Empirical Finance*, *2*(2), 103-115.

deB. Harris, F. H., McInish, T.H., & Wood, R.A. (2002). Security price adjustment across exchanges: An investigation of common factor components for Dow stocks. *Journal of Financial Markets*, *5*(3), 277-308.

Dorfman, J.H., & Karali, B. (2015). A nonparametric search for information effects from USDA reports. *Journal of Agricultural and Resource Economics*, *40*(1), 124-143.

Engle, R.F., & Granger, C.W. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251-276.

Evans, M. D., & Lyons, R. K. (2008). How is macro news transmitted to exchange rates? *Journal of Financial Economics*, *88*(1), 26–50.

Ferraresi, T., Roventini, A., & Fagiolo, G. (2015). Fiscal policies and credit regimes: A TVAR approach. *Journal of Applied Econometrics*, *30*(7), 1047-1072.

Frank, J., & Garcia, P. (2010). How strong are the linkages among agricultural, oil, and exchange rate markets? Proceedings of the NCCC134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO.

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

Fricke, C., & Menkhoff, L. (2011). Does the "bund" dominate price discovery in euro bond futures? Examining information shares. *Journal of Banking & Finance*, *35*(5), 1057-1072.

Garbade, K., & Silber, W. (1983). Price movements and price discovery in futures and cash markets. *Review of Economics and Statistics*, *65*(2), 289-297.

Garcia, P., Irwin, S.H., & Smith, A. (2015). Futures market failure? *American Journal of Agricultural Economics*, *97*(1), 40-64.

Gousgounis, E., & Onur, E. (2018). The effect of pit closure on futures trading. *Journal of Commodity Markets*, *10*(2018), 69-90.

Hagströmer, B., & Nordén, L. (2013). The diversity of high-frequency traders. *Journal of Financial Markets*, *16*(4), 741-770.

Hansen, B. E. (1999). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics*, *93*(2), 345-368.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, *50*(4),1029-1054.

Hansen, P. R., & Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, *24*(2), 127-161.

Harri, A., Nalley, L., & Hudson, D. (2009). The relationship between oil, exchange rates, and commodity prices. *Journal of Agricultural and Applied Economics*, *41*(2), 501-510.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *Journal of Finance*, *50*(4), 1175-1199.

Hasbrouck, J. (2018). High-frequency quoting: Short-term volatility in bids and offers. *Journal of Financial and Quantitative Analysis*, *53*(2), 613-641.

Hasbrouck, J. (2018). Price discovery in high resolution. Working Paper.

Hasbrouck, J. and Saar, G. (2013). Low-latency trading. *Journal of Financial Markets*, *16*(4), 646-679.

Hasbrouck, J., & Saar, G. (2009). Technology and liquidity provision: The blurring of traditional definitions. *Journal of Financial Markets*, *12*(2), 143-172.

Hatzenbuehler, P. L., Abbott, P. C., & Foster, K. A. (2016). Agricultural commodity prices and exchange rates under structural change. *Journal of Agricultural and Resource Economics*, *41*(2), 204-224.

Hauptfleisch, M., Putniņš, T. J., & Lucey, B. (2016). Who sets the price of gold? London or New York. *Journal of Futures Markets*, *36*(6), 564-586.

Haynes, R., & Roberts, J. S. (2015). Automated trading in futures markets. CFTC White Paper.

Haynes, R., & Roberts, J. S. (2017). Automated trading in futures markets. CFTC White Paper.

Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *The Journal of Finance*, *66*(1), 1-33.

Hu, Z., Mallory, M., Serra, T., & Garcia, P. (2017). Measuring price discovery between nearby and deferred contracts in storable and non-storable commodity futures markets. Working Paper.

Hubbs, T. 2017 a. USDA reports provide support to soybean prices. *farmdoc daily*, *7*(7), Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.

Hubbs, T. 2017 b. Summer pricing for corn and soybeans. *farmdoc daily*, *7*(94), Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.

Irwin, S., & Good, D. (2016). The relationship between stocks-to-use and corn and soybean prices: An alternative view. *farmdoc daily*, *6*(66), Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.

Irwin, S.H., & Sanders, D.R. (2012). Financialization and structural change in commodity futures markets. *Journal of Agricultural and Applied Economics*, *44*(3), 371-396.

Isengildina, O., Irwin, S.H., & Good, D.L. (2006). The value of USDA situation and outlook information in hog and cattle markets. *Journal of Agricultural and Resource Economics*, *31*(2), 262-282.

Ivanov, V., & Kilian, L. (2005). A practitioner's guide to lag order selection for VAR impulse response analysis. *Studies in Nonlinear Dynamics & Econometrics*, *9*(1).

Janzen, J., Carter, C. A., Smith, A., & Adjemian, M. (2014). Deconstructing wheat price spikes: A model of supply and demand, financial speculation, and commodity price comovement. USDA-ERS Economic Research Report.

Janzen, J.P., & Adjemian, M.K. (2017). Estimating the location of world wheat price discovery. *American Journal of Agricultural Economics*, *99*(5), 1188-1207.

Kirilenko, A., Kyle, A.S., Samadi, M., & Tuzun, T. (2017). The flash crash: High-frequency trading in an electronic market. *Journal of Finance*, *72*(3), 967-998.

Koontz, S., Garcia, P., & Hudson, M. (1990). Dominant-satellite relationships in live cattle markets. *Journal of Futures Markets*, *10*(2), 123-136.

Koop, G., Pesaran, M. H., & Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, *74*(1), 119-147.

Kost, W. E. (1976). Effects of an exchange rate change on agricultural trade. *Agricultural Economics Research*, *28*(3), 99-106.

Kyle, A. S., & Obizhaeva, A. A. (2016). Market microstructure invariance: Empirical hypotheses. *Econometrica*, *84*(4), 1345–1404.

Leath, M.N., & Garcia, P. (1983). The efficiency of the corn futures market in establishing forward prices. *North Central Journal of Agricultural Economics*, *5*(2), 91-101.

Lee, C. M., & Ready, M. J. (1991). Inferring trade direction from intraday data. *Journal of Finance*, *46*(2), 733-746.

Leuthold, R.M. (1972). Random walk and price trends: The live cattle futures market. *Journal of Finance*, *27*(4), 879-889.

Leuthold, R.M., Junkus, J.C., & Cordier, J.E. (1989). *The Theory and Practice of Futures Markets*. Champaign: Stipes Publishing.

Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, *30*(1), 67-80.

Li, S, Wang, X and Ye, M. (2017). Who provides liquidity and when? Working Paper.

Lo, A. W., & MacKinlay, A. C. (1989). The size and power of the variance ratio test in finite samples: A Monte Carlo investigation. *Journal of Econometrics*, *40*(2), 203-238.

McKenzie, A. M. (2008). Pre-harvest price expectations for corn: The information content of USDA reports and new crop futures. *American Journal of Agricultural Economics*, *90*(2), 351-366.

Mizrach, B., & Neely, C.J. (2008). Information shares in the US Treasury market. *Journal of Banking & Finance*, *32*(7), 1221-1233.

Naik, G., & Leuthold, R.M. (1988). Cash and futures price relationships for nonstorable commodities: An empirical analysis using a general theory. *Western Journal of Agricultural Economics*, *13*(2), 327-338.

Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, *61*(4), 631-653.

O'Brien, D. M., & Olson, F. (2014). The Competitive position of the black sea region in world wheat export markets. Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO.

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics*, *116*(2), 257–70.

Orden, D. (1986). A critique of exchange rate treatment in agricultural trade models: comment. *American Journal of Agricultural Economics*, *68*(4), 990-993.

Pesaran, H. H., & Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics letters*, *58*(1), 17-29.

Putniņš, T.J. (2013). What do price discovery metrics really measure? *Journal of Empirical Finance*, *23*(2013), 68-83.

Reed, W.R., & Smith, A. (2017). A time series paradox: Unit root tests perform poorly when data are cointegrated, *Economics Letters*, *151*(2017), 71-74.

Rigobon, R. (2003). Identification through heteroskedasticity. *Review of Economics and Statistics*, *85*(4), 777-792.

Ruppel, F. J. (1987). Agricultural commodity export data: Sales and shipments contrasted. *Journal of* A*gricultural and Resource Economics*, *39*(2), 22-38.

Sanders, D.R., Garcia, P. & Manfredo, M.R. (2008). Information content in deferred futures prices: Live cattle and hogs. *Journal of Agricultural and Resource Economics*, *33*(1), 87-98.

Schnake, K.N., Karali, B. & Dorfman, J.H. (2012). The informational content of distant-delivery futures contracts. *Journal of Agricultural and Resource Economics*, *37*(2), 213-227.

Scholtus, M., Van Dijk, D., & Frijns, B. (2014). Speed, algorithmic trading, and market quality around macroeconomic news announcements. *Journal of Banking & Finance*, *38*(2014), 89-105.

Schroeder, T.C., & Goodwin, B.K. (1991). Price discovery and cointegration for live hogs. *Journal of Futures Markets*, *11*(6), 685-696.

Schuh, G. E. (1974). The exchange rate and US agriculture. *American Journal of Agricultural Economics*, *56*(1), 1-13.

Shang, Q., Mallory, M., & Garcia, P. (2018). The components of the bid-ask spread: Evidence from the corn futures market. *Agricultural Economics*, *49*(3), 381-393.

Silber, W. L. (1984). Market maker behavior in an auction market: An analysis of scalpers in futures markets. *Journal of Finance*, *39*(4), 937-953.

Smith, A. (2005). Partially overlapping time series: A new model for volatility dynamics in commodity futures. *Journal of Applied Econometrics*, *20*(3), 405-422.

Stock, J. H., & Yogo, M. (2002). *Testing for weak instruments in linear IV regression* (No. t0284). National Bureau of Economic Research.

Tomek, W.G. (1997). Commodity futures prices as forecasts. *Review of Agricultural Economics*, *19*(1), 23-44.

Tomek, W.G., & Gray, R.W. (1970). Temporal relationships among prices on commodity futures markets: Their allocative and stabilizing roles. *American Journal of Agricultural Economics*, *52*(3), 372-380.

Tsay, R. S. (1998). Testing and modeling multivariate threshold models. *Journal of the American Statistical Association*, *93*(443), 1188-1202.

U.S. Securities and Exchange Commission. (2014). Equity Market Structure Literature Review Part II: High Frequency Trading. (March), 1–37.

Wang, X., Garcia, P., & Irwin, S. H. (2013). The behavior of bid-ask spreads in the electronically-traded corn futures market. *American Journal of Agricultural Economics*, *96*(2), 557-577.

Working, H. (1948). Theory of the inverse carrying charge in futures markets. *Journal of Farm Economics*, *30*(1), 1-28.

Working, H. (1949). The theory of price of storage. *The American Economic Review*, *39*(6), 1254-1262.

Wright, B. D. (2011). The economics of grain price volatility. *Applied Economic Perspectives and Policy*, *33*(1), 32-58.

Yan, B., & Zivot, E. (2010). A structural analysis of price discovery measures. *Journal of Financial Markets*, *13*(1), 1-19.

Yang, J., Bessler, D., & Leatham, D. (2001). Asset storability and price discovery in commodity futures markets: A new look. *Journal of Futures Markets*, *21*(2), 279-300.

Zhang, F. (2010). High-frequency trading, stock volatility, and price discovery. Working Paper