

The Effects of Microstructure Noise on Realized Volatility in the Live Cattle Futures Market

by

Anabelle Couleau, Teresa Serra, and Philip Garcia

Suggested citation format:

Couleau, A., T. Serra, and Philip Garcia. 2017. "The Effects of Microstructure Noise on Realized Volatility in the Live Cattle Futures Market." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO.
[<http://www.farmdoc.illinois.edu/nccc134>].

The Effects of Microstructure Noise on Realized Volatility in the Live Cattle Futures Market

Anabelle Couleau, Teresa Serra, and Philip Garcia*

*Paper presented at the NCCC-134
Conference on Applied Commodity Price Analysis,
Forecasting, and Market Risk Management
St. Louis, Missouri, April 24-25, 2017.*

Copyright 2017 by Anabelle Couleau, Teresa Serra and Philip Garcia. All rights reserved.
Readers may make verbatim copies of this document for noncommercial purposes by any means,
provided that this copyright notice appears on all such copies.

* The authors are Ph.D. research assistant, associate professor and professor in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign. The authors gratefully acknowledge economic support by the National Institute of Food and Agriculture, U.S. Department of Agriculture, under award number ILLU-470-332, and the Office of Futures and Options Research (OFOR) at UIUC. Contact: 326 Mumford Hall, 1301 W. Gregory Drive, Urbana, IL 61801. Phone: 217-333-1810. Email: couleau2@illinois.edu

The Effects of Microstructure Noise on Realized Volatility in the Live Cattle Futures Market

Abstract

Recently, U.S. live cattle futures prices have experienced high levels volatility which has raised concerns about the impact of high frequency trading. This paper identifies the market microstructure noise present in high frequency data and its implications for realized volatility of returns in live cattle futures markets from 2011 to 2015. Short- and long-term components of volatility are identified using nonparametric and semi-parametric procedures. While market microstructure noise is found to increase realized volatility when the sampling frequency is below 4-minute time intervals, the particularly high volatility in live cattle markets in 2015 is found to be strongly driven by market fundamentals, affected by supply and demand shocks. Important policy implications from the results are drawn.

Key words: live cattle, futures, microstructure noise, realized volatility, integrated variance.

Introduction

The U.S. live cattle market experienced particularly high volatility in 2015. While shocks to market fundamentals (such as weather or disease shocks) are well known sources of price volatility in agricultural commodity markets, beef producers blamed high frequency trading (HFT) activities for bearing the primary responsibility for the 2015 market events (Meyer 2016).¹ This added fuel to the ongoing debate on the impacts of high frequency trading on futures markets. High frequency trading became a possibility after the emergence of electronic trading in agricultural futures markets in 2006. By 2011, livestock futures trading on the electronic platform increased to 80% of total trading volume (Irwin and Sanders 2012) and reached about 95% in 2015 (Gousgounis and Onur 2016).² In a recent document by the Commodity Futures Trading Commission, Haynes and Roberts (2015) identified automated trading to be approximately 32.4% of total futures volumes traded in livestock markets between 2012 and 2014.

The recent structural changes and concerns warrant analysis of the characteristics of market microstructure in live cattle futures markets. While some research findings support that high frequency traders can improve market quality by acting as liquidity providers by posting limit orders at ultra-high speed (Brogaard, Hendershott, and Riordan 2014; Conrad, Wahal, and Xiang 2015), others argue that high frequency trading strategies such as high frequency quoting might harm market quality (Hasbrouck 2015; Wang 2014).³ In spite of the absence of published research results on the live cattle market, the Chicago Mercantile Exchange (CME) responded to beef producers' concerns by reducing the number of trading hours for its live cattle futures contract. Complaints by producers and the CME's response should be weighed against evidence regarding how much market noise is generated by HFT activities and the impacts of this noise on price volatility (Stebbins 2013). However, research evidence that can be directly derived is limited, as the data currently available do not identify HFT in agricultural commodity markets. Nevertheless, it is possible to shed some light on the situation by identifying market microstructure noise using

high frequency data and the extent to which this noise distorts price volatility measures. Specifically, in the context of live cattle futures market, it would be informative to know the extent to which price volatility in 2015 was due to market fundamentals, or to noise that may in part be attributable to HFT. This research contributes to this understanding.

In finance, risk is usually expressed as the volatility of returns. Return volatility is important in a number of management decisions as volatility levels have a direct impact on asset pricing, risk management, or portfolio allocation. It is thus important for market participants to understand how volatility behaves (Egelkraut, Garcia and Sherrick 2007; Dana, Gilbert and Shim 2006). While market returns can be usually measured with minimal measurement error and can be modeled using standard time series methods, the latent (unobservable) nature of volatility requires the use of more complex approaches (Andersen and Benzoni 2008).

Extensive efforts have been made in the financial economics literature to provide accurate estimates of volatility. Because of shortcomings in their theoretical or statistical assumptions, conventional measures of volatility such as historical volatility, implied volatility and stochastic volatility are limited in developing an accurate representation of volatility. In contrast, *realized variance* (RV) constitutes a different characterization; its square root, the realized volatility, is the empirical counterpart of the so-called spot or instantaneous volatility when there is no microstructure noise. It relies on the notion that if price data were continuously observed, one may obtain error free measures of realized return and return variation. Realized volatility uses nonparametric methods to estimate the ex-post return variation (Andersen and Benzoni 2008). In the absence of microstructure noise, realized volatility is a consistent estimator of latent volatility as the time between observations tends to zero. In practice, however, the microstructure noise present in high frequency data leads to non-robust realized variance estimates (Andersen et al. 2000; Kalnina and Linton 2008). As a result of the presence of noise, RV of log-price data has been shown to explode as the sampling interval approaches zero (Zhang et al. 2005). Noise is attributed to market imperfections and includes price discreteness,⁴ bid-ask bounce effects,⁵ infrequent trading, as well as high frequency quoting (Hasbrouck 2015; Wang 2014; Hagströmer and Nordén 2013; O'Hara 2015).

While other causes may exist, high frequency trading activities are conjectured to lie behind the increase in the volatility in cattle market as prices were falling in 2015 (Figure 1) (Meyer 2016). With the exception of the work by Wang (2014), empirical investigations of microstructure noise have focused on non-agricultural markets. Wang (2014) identifies microstructure noise using bid and ask quotes in corn futures markets from 2008 to 2013, and finds high frequency quoting to have a minimal role in the increase of noise volatility.

The contribution of our research article goes beyond Wang (2014) in that we not only identify microstructure noise, but also the futures market efficient price and how observed prices (transaction prices, bid and ask quotes), the efficient price and noise are related. In this regard, a more comprehensive understanding of market microstructure characteristics is provided. Our research findings show that different observed prices are not equally related to the latent efficient price. Hence, and as shown in Hansen and Lunde (2006), measures of noise variance will be sensitive to the observed price used. Consistently, and in contrast to Wang (2014), our characterization of noise, not only relies on bid and ask quotes, but also on transactions prices.

This research measures the economic magnitude of the noise by comparing the long-run variance (representing the variance caused by changes in the efficient price containing fundamental market information) with a short-run realized variance measure based on a high frequency sample. More specifically, the approach to measure noise introduced by Zhou (1996) and extended by Hansen and Lunde (2006) is used. Also, this article assesses how noise changes through time and across contracts. To the extent that HFT agents generate noise in the market, their increased market presence over time and in the more liquid contracts (Brogaard, Hendershott, and Riordan 2014) should be reflected in higher noise levels. Finally, this paper contributes to recent heated policy debates on whether high frequency trading should be regulated in agricultural futures markets.

Our findings suggest that the particularly high volatility in live cattle markets in 2015 is strongly driven by market fundamentals, captured by the long-term or integrated variance (IV). This result casts doubt on the notion that high frequency trading was responsible for the high volatility in cattle markets in 2015. The short-term variance component, also called noise variance, is found to substantially distort the measure of realized variance at 1-second sampling frequency, and to fade between 3- and 4-minute sampling frequencies. Distortions caused by noise are especially important during periods of relevant efficient price return variance. Noise is on average 1 cent per pound and represents between 0.6% and 0.9% of the transaction prices over the period studied.

Relevant Literature

Motivated by a desire to provide more accurate information to decision makers, much financial economics research has focused on finding the “true” variance in price. The availability of intraday high frequency data opened the possibility of using realized volatility to construct observable proxies for latent volatility. Due to the complex nature of high frequency data, many researchers set aside parametric approaches as a modeling instrument (Andersen et al. 2003) and focused on nonparametric methods. These methods do not impose functional forms on the stochastic volatility or specific densities on the error term. Several nonparametric variance measures have been suggested, including RV, realized bipower variation, realized kernel, among others (Bollerslev and Andersen 1998; Andersen et al. 2001; Barndorff-Nielsen et al. 2008; Hansen and Horel 2009; Angelidis and Degiannakis 2008).

Among these procedures, realized kernels have been found to moderate the impacts of microstructure noise on measures of variance (Hansen and Lunde 2006; Barndorff-Nielsen et al. 2008; 2009).⁶ Zhang et al. (2005) also mitigate the effect of microstructure noise by sampling the high-frequency data over longer horizons. In a similar vein, wavelet analyses compute long-run wavelet variances using short-run averages.

Hasbrouck (2015) derives a wavelet-based short-run volatility measure and examines the high frequency quoting noise in the U.S. equity market. In contrast to the RV literature, where time may be measured in units of economic activity (Hasbrouck 2015),⁷ wavelet analysis decomposes variance on a scale by scale basis.⁸ His outcomes suggest that quotes at 50 milliseconds scale are five times more volatile than what is expected from an efficient price following a random walk

process. Short-run volatility appears to be smaller for firms with highly traded stocks. The short-to long-run variance ratio is closer to one (the more the ratio exceeds one, the more there is excess variance) for the highly traded stocks than the lowly traded stocks. Hasbrouck (2015) findings are consistent with those in Hansen and Lunde (2006), who approximate price variance using RV, and assess microstructure noise in 2001 and 2004 with quotes time stamped to the second.

Wavelet methods have also been applied to the corn futures market by Wang (2014) to evaluate the impact of high frequency quoting on volatility. Using Top-of-Book (Best-Bid-Offer - BBO) dataset time stamped to the second, Wang (2014) applies a Bayesian framework developed in Hasbrouck (2013) to sub-divide the sample into sub-seconds.⁹ Microstructure noise is measured as the net excess price variance. This measure is found by comparing the wavelet-based short-run price variance to that of the variance expected with an efficient price following a random walk process. He estimates the variances using bid and ask quotes series separately and reports the mean. Findings suggest the existence of noise with its highest value at 250 milliseconds scale. The magnitude of the noise is however found to be relatively small with a net excess volatility between 3 and 10% of a tick size and to decrease with time. This leads Wang (2014) to conclude that high frequency traders have not contributed to increase volatility between 2008 and 2013 in the corn market.

Based on the knowledge that microstructure noise present in high frequency data leads to non-robust RV estimates (Andersen et al. 2000; Kalnina and Linton 2008), methods to capture high frequency trading noise are based on specific assumptions of the properties of this noise. The assessment of the robustness of these measures is also highly dependent on specification of the noise. Hansen and Lunde (2006), Hasbrouck (2015) and Wang (2014) use the comparison between the RV and the integrated variance (IV) as a measure for market noise. The integrated volatility, the square root of IV, aims at capturing the “true” volatility caused by shocks to fundamentals such as supply and demand shocks, free of microstructure noise. Since market microstructure noise induces autocorrelation in intraday returns, Hansen and Lunde (2006) assume that noise is time dependent and propose a robust realized volatility estimator that is unbiased for a general type of noise. The time dependence property in the noise is also examined in Hasbrouck (2015). Another property of noise is that it might be correlated with efficient returns. This correlation is empirically assessed in Hansen and Lunde (2006) by deriving the efficient price through cointegration analysis and computing the correlation of efficient price returns and noise. Failure to allow for these correlations can have important implications when measuring noise volatility, creating an additional bias.

Methods

This paper uses two methods to approximate market microstructure noise. The first is fully nonparametric, based on observed prices and does not attempt to elicit an efficient price. Noise variance is identified by comparing the integrated variance of prices (which reflects the variance over longer time intervals and is robust to microstructure noise) to the variances at higher frequency intervals which contain noise. The second approach is semi-parametric and is based on the derivation of the efficient price by means of cointegration procedures. In the second approach, noise variance is estimated by comparing the realized variance of the efficient and observed prices.

Nonparametric Noise Model

This section presents a nonparametric approach to measure noise that is based on the comparison between RV and IV using high frequency data. Following Hansen and Lunde (2006) and Bandi and Russell (2003), market microstructure noise is characterized by equation (1),

$$p_t = p_t^* v_t . \quad (1)$$

where p_t is the observed price at time t , p_t^* is the latent (or "efficient" or "true") price and v_t is the microstructure noise. In this framework, microstructure noise is attributed to transactions costs (bid-ask spread), price discreteness (tick size), infrequent trading, as well as high frequency quoting which generates noise in quote prices and increases frictional variance (Hasbrouck 2015; Wang 2014). To facilitate working with price returns, a simple logarithmic transformation is applied to equation (1),

$$\ln(p_t) = \ln(p_t^*) + \ln(v_t). \quad (2)$$

The latent or efficient log-price process is assumed to follow a Brownian semi-martingale (Hamilton 1994, p. 477) solving equation (3):

$$d\ln(p_t^*) = \sigma_t dW_t , \quad (3)$$

where W_t is a standard Brownian motion, and σ_t is the (continuous) random volatility function. The IV, which reflects the long-term variance free of microstructure noise, is thus defined as follows,

$$IV \equiv \int_0^T \sigma(t)^2 dt . \quad (4)$$

The time interval $[0, T]$ can be divided into m sub-intervals. For a fixed m , the i th subinterval is given by $[t_{i-1,m}, t_{i,m}]$ with $t_{0,m} = 0 < \dots < t_{m,m} = T$. Using (2), intraday observed returns for each interval can be written as,

$$r_{i,m} = r_{i,m}^* + e_{i,m} , \quad (5)$$

where $r_{i,m} = \ln(p_{t_{i,m}}) - \ln(p_{t_{i-1,m}})$, $r_{i,m}^* = \ln(p_{t_{i,m}}^*) - \ln(p_{t_{i-1,m}}^*)$ and $e_{i,m} = \ln(v_{t_{i,m}}) - \ln(v_{t_{i-1,m}})$ for $i = 1, \dots, m$.

To derive measures of variance that lead to identification of market noise, start with the idea that RV, defined as the sum of squared returns, is a process integrated by two components. The first component, the long-run RV, assumes that returns follow a random walk, and that long-run return levels and their variances respond to market information. The RV of r^* is defined by,

$$RV_*^{(m)} = \sum_{i=1}^m (r_{i,m}^*)^2 , \quad (6)$$

for $i = 1, \dots, m$. The second component is the short-run RV that deviates from the long-run random walk variance due to the presence of noise. The RV of r , based on observable returns, is given by,

$$RV^{(m)} = \sum_{i=1}^m r_{i,m}^2 . \quad (7)$$

Equation (6) provides a consistent estimator of IV as the sampling frequency tends to infinity (Protter 2005), but p_t^* and hence r_t^* are not observable. Equation (7) provides a measure of variance which contains both the IV and noise, making it a biased and generally an inconsistent estimate of the efficient variance (Bandi and Russell 2003). In the absence of microstructure noise, RV based on observed returns has been shown to be a consistent estimator of the IV as the time between observations tends to zero. In practice, ignoring microstructure noise only seems to work well for sampling frequencies of 10 minutes or more, for which the variance seems to be free of microstructure noise and thus to reflect the IV (Kalnina and Linton 2008; Hansen and Lunde 2006).

Since equation (7) contains both the IV and the noise variance, we need a feasible measure of the IV to accurately isolate noise variance. Zhou (1996) was the first to introduce a RV estimator that attempts to estimate the IV. The estimator corrects for bias due to noise through a first-order autocorrelation term. While the standard RV is commonly measured using the previous tick, Zhou's IV estimator is based on the price sampled at a fixed time interval (e.g. every minute), resulting in a measure that is often more heteroscedastic than when sampled at every tick (Zhou 1996). Hansen and Lunde (2006) show that Zhou's estimator is not robust, and requires accounting for higher-order autocorrelations. Increasing the order of autocorrelation increases the robustness of the estimator to both noise time dependence and the correlation between the efficient intraday returns and noise. Their generalized estimator uses a Bartlett-based kernel for any $k \geq 2$ and can be expressed as,¹⁰

$$RV_{ACNW_k}^{(1\ tick)} \equiv \hat{\gamma}_0 + \sum_{j=1}^k (\hat{\gamma}_{-j} + \hat{\gamma}_j) + \sum_{j=1}^k \frac{k-j}{k} (\hat{\gamma}_{-j-k} + \hat{\gamma}_{j+k}), \quad (8)$$

where $\hat{\gamma}_j \equiv \sum_{i=1}^N r_i r_{i+j}$, $\hat{\gamma}_0 \equiv r_i^2$, and k is the order of autocorrelation. In their empirical application, autocorrelation of order 30 was selected as $RV_{ACNW_{30}}^{1\ tick}$ stabilized. An estimate of the integrated variance can be calculated for each day. For a specific sample period, it is possible to visually compare, using variance signature plots, an average of the daily estimates of the integrated variance, $\overline{RV}_{ACNW_k}^{1\ tick}$, to a sample average of the daily short-run RV (denoted by \overline{RV}_t^m) estimated for progressively longer intraday time intervals m . The difference between the short-run and the long-run RV represents the variance of noise for interval m .

Semiparametric Noise Model

An alternative method to identify and quantify noise involves estimating the efficient price and its variance using cointegration, and comparing its variance to the realized variance. The method uses a cointegration approach based on the observed prices (quotes and transaction prices) to identify the common stochastic component, which represents the efficient market price and the transitory component, which represents the noise process for each price series (Hansen and Lunde 2006). Once the efficient price is estimated, the IV is derived as the efficient price's realized variance. An advantage of using this method is that it allows us to understand how efficient prices are linked to noise and which of the observed price series (bid, ask quotes, or transaction prices) is most informative regarding the efficient price.

Let t_i for $i = 0, 1, \dots, I$ denote the time when transactions occur during trading days. The vector of log-observed prices is given by,

$$\mathbf{p}_{t_i} = \begin{pmatrix} \text{transaction price at time } t_i \\ \text{corresponding ask price at } t_i \\ \text{corresponding bid price at } t_i \end{pmatrix}, \quad (9)$$

and the vector error correction model (VECM) of these three prices is then specified as,

$$\Delta \mathbf{p}_{t_i} = \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{p}_{t_{i-1}} + \sum_{j=1}^{l-1} \boldsymbol{\Gamma}_j \Delta \mathbf{p}_{t_{i-j}} + \boldsymbol{\mu} + \boldsymbol{\epsilon}_{t_i}, \quad (10)$$

where $\boldsymbol{\epsilon}_{t_i}$ ($i = 0, 1, \dots, I$) is assumed to be an uncorrelated error vector, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are 3×2 matrices, l is the number of lags, $\boldsymbol{\Gamma}_j$ are the parameters capturing the short-run dynamics and $\boldsymbol{\mu}$ is a 3 by 3 vector of constants. The three observed prices are assumed to share the same stochastic trend (i.e. the efficient price) and the bid-ask spread is assumed to be stationary. The cointegration matrix $\boldsymbol{\beta}$ is also assumed to be known and determined by economic theory.¹¹ $\boldsymbol{\beta}'_1 \mathbf{p}_{t_i}$ is chosen to represent the difference between the transaction price and the mid-quote (average between the bid and ask quotes) and $\boldsymbol{\beta}'_2 \mathbf{p}_{t_i}$ is used to represent the bid-ask spread. As a result, $\boldsymbol{\beta}$ can be expressed as,

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1 \boldsymbol{\beta}_2) = \begin{pmatrix} 1 & 0 \\ -1/2 & 1 \\ -1/2 & -1 \end{pmatrix}. \quad (11)$$

In order to identify $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, the following normalization vectors are imposed, $\boldsymbol{\beta}_\perp = (1 \ 1 \ 1)'$ and $\boldsymbol{\alpha}'_\perp (1 \ 1 \ 1)' = 1$, where $\boldsymbol{\beta}_\perp$ and $\boldsymbol{\alpha}_\perp$ are 3×1 vectors (Hansen and Lunde 2006).

Identification of the common stochastic trend (efficient price) follows Hasbrouck (2002) and is based on the Granger representation. After estimation of the parameters, the efficient price is,

$$\hat{p}_{t_i}^* = (\hat{\boldsymbol{\alpha}}'_\perp \boldsymbol{\Gamma} \boldsymbol{\beta}_\perp)^{-1} \sum_{j=1}^i \hat{\boldsymbol{\alpha}}'_\perp \hat{\boldsymbol{\epsilon}}_{t_j} \quad (12)$$

and the corresponding efficient intraday return is given by,

$$r_{t_i}^* \equiv \frac{\hat{\boldsymbol{\alpha}}'_\perp \hat{\boldsymbol{\epsilon}}_{t_i}}{(\hat{\boldsymbol{\alpha}}'_\perp \boldsymbol{\Gamma} \boldsymbol{\beta}_\perp)}. \quad (13)$$

From (13), the RV of $\hat{p}_{t_i}^*$ is,

$$RV_{p^*} = \sum_{i=1}^I (\hat{r}_{t_i}^*)^2 \quad (14)$$

Equation (14) represents the long-run parametric variance, denoted by (RV_{p^*}) . Following Hansen and Lunde (2006), the estimated model allows us to decompose the observed RV into three components, i.e., the efficient price returns variance (14), the noise variance, and the variance due to correlations between increments of noise and the efficient prices returns,

$$RV = \sum_{i=1}^I (\hat{r}_{t_i}^*)^2 + 2 \sum_{i=1}^I \hat{e}_{t_i} \hat{r}_{t_i}^* + \sum_{i=1}^I (\hat{e}_{t_i})^2 \quad (15)$$

where \hat{e}_{t_i} is an element of $\hat{e}_{t_i} \equiv \ln(\hat{v}_{t_i}) - \ln(\hat{v}_{t_{i-1}})$ and $\hat{v}_{t_i} = \ln(p_{t_i}) - \ln(p_{t_i}^*)$. The model in equation (10) is estimated by least squares and includes a restricted constant. Specifically, the constraint $\mu = \alpha\rho$ where $\rho = (\rho_1, \rho_2)'$ is imposed to avoid a linear deterministic trend in the price series within a single day. As a result, the average difference between transaction prices and mid-quotes is represented by $-\rho_1$ and the average bid-ask spread is represented by $-\rho_2$ (Hansen and Lunde 2006).

Empirical Analysis

For purposes of comparison, the research focuses on the period from 2011 to 2015 which is characterized by a predominance of electronic platform trading in live cattle futures markets. As noted, livestock futures trading volumes on electronic platforms represented about 80% in 2011 (Irwin and Sanders 2012) and reached 100% when the CME live cattle future pit closed in July 2015.¹² Prices changed substantially during this period (see Figure 1). While prices were relatively stable from late 2012 to late 2013, they experienced significant increases from late 2013 to late 2014. Price then declines at the end of 2014 which led to the recent turmoil. The structure of the data is first offered which is followed by a description of the data cleaning procedure employed in the paper.

Data and Data Pre-processing

Data used in the analysis consist of transactions and quotes from CME Group's BBO (Best-Bid-Offer) dataset for live cattle futures contracts that are traded on the electronic trading platform. The day trading session begins at 09:05 am and ends at 1 pm. The dataset contains the best bid and ask prices (top-of-book quotes) and transactions prices, with their respective volumes. These variables are time stamped to the second. Several bid or ask quotes may take place within each second. While a sequence number allows to sort them within the second, the exact time when they hit the book is unknown. Similar data have been used by Wang (2014), who studies the electronic corn market, and Hasbrouck (2015) who focuses on the electronic stock market.

The U.S. live cattle futures contract is traded with six maturities a year: February, April, June, August, October and December. Delivery occurs on the Monday following the first Friday in the expiration month. While the debate on the impacts of HFT on market noise is not yet settled in agricultural markets, it is plausible that increased market participation of HFT agents could have increased the noise embedded in the most liquid contracts in which high frequency traders tend to operate (Brogaard, Hendershott, and Riordan 2014). The analysis focuses on the most actively traded contract in each year. Among the contracts maturing in 2011, the June contract has the highest number of daily average traded contracts: 15,545, being 13,570 for the April and 13,890 for the August 2011 contracts. The June contract is thus selected. For 2012 and 2013, the June and April contracts are chosen, respectively. The December contract is the most highly traded in 2014 and 2015. For each contract selected, the analysis focuses on the most actively traded interval: from 60 to 10 days before maturity.

The intraday distribution of quotes and trades for these most traded contracts vary. Quotes are reported in approximately 50% of the total number of seconds within the day trading hours (i.e. 14,100 seconds), resulting in an average time between two observed quotes of 1.92 seconds in 2015. In contrast, 5.82 seconds separate observed trades. Figure 2 shows the average number of bid and ask quotes per minute within a trading day in 2011 and 2015.¹³ The U-shape reveals a concentration of quotes at the beginning and the end of the day. This likely corresponds to the accumulation of information overnight, which is reflected in the first 15 minutes, or adjustments to expected overnight information, which is reflected in the 15 minutes before the futures pit trading closes (Wang 2014). The 2015 contract has the highest average number of quotes per minute (270) across all years from 2011 to 2015. The difference in the magnitude of the peak at the end of the day between 2011 and 2015, might be due to several factors. First, the cancellation of the night session by the exchange in October 2014 is likely to have shifted the overnight trading to the day trading sessions (having thus increased the average number of quotes and the height of the peak at the end of the day). Second, this end-day peak might be related to an increase in small orders generated by HFT activities. The following discussion describes the data cleaning methods used in this research.

Several issues in high frequency data have the potential to bias research results. These include: (1) misplaced decimal or abnormal zero prices, (2) several quotes or trade data being time stamped to the same second, and (3) the presence of price limit moves. Following Barndorff-Nielsen et al. (2009), pre-processing data procedures are applied to overcome these issues.¹⁴ The data selected for analysis were reviewed. First, all zero-priced bids, asks, and transactions are deleted. Second, since multiple quotes have the same time stamp, they are replaced with the median bid and ask price as proposed by Barndorff-Nielsen et al. (2009) and Hansen and Lunde (2006). Similarly, the median transaction price is selected when there are several transaction prices with the same time stamp.

Since we were concerned that large information shocks might produce price adjustments that do not normally characterize the market, USDA cattle reports release days are excluded in favor of sample homogeneity. In addition, we examined the days selected for limit moves. Limit price moves occur when the price reaches either the minimum or the maximum price change allowed by the exchange on a day. The limit price can be reached at the open or near the open of a trading session and stay at the limit with little or no trading. Prices can progress to the limit through the day and remain at the limit. Alternately, prices can reach the limit price for a period and then revert back to a trading region. When prices are at the limit, realized variance, informational variance, and noise variance are all reduced to zero for that period.

In this research, limit price moves are defined as those lasting at least 30 minutes. Using this criterion, thirteen limit move days (almost evenly split between price increases and decreases) were identified in the sample, with the largest number, 10 occurring in 2015. Two of the remaining limit moves occurred in 2014, and the other in 2011. Inspection of the intraday transaction prices revealed a variety of price paths. Nine price paths showed a gradual trend in price which eventually hit the limit. Six of these reached the limit after more than a half a day of trading, and two of these reverted back to a trading region. In the first three weeks of November 2015, four limit moves (three down and one up) occurred quickly within the trading day and prices remained at the limit throughout the trading session. Careful examination of the USDA releases of related reports (e.g.,

cold storage, Livestock Slaughter, WASDE announcements) revealed possible effects on live cattle prices. However, several limit moves appeared unrelated to USDA information releases. Absolute measures of realized variance, informational variance, and noise variance should be reduced when the limit moves are included. However, relative measures (noise variance relative as a proportion of realized variance or noise variance as a proportion of information variance) should be less affected. In order to clarify the effect of the limit days on variance and noise, several analyses are performed with and without the limit move observations.

Two intraday sampling schemes are used in the research. First, tick time sampling (TTS) is based on the time a transaction occurs and involves unequal temporal spacing between observations. Following Hansen and Lunde (2006), TTS is employed in the estimation of the IV both in the kernel method and cointegration analysis. When using statistical methods, the use of unequally-spaced observations is preferred to filling in prices because of potential biases that may occur. Since transaction observations are spaced every 6 seconds on average in the sample, tick time sampling involves around 1/6 of the total seconds in a trading day.¹⁵ The second sampling scheme is the calendar time sampling (CTS), which implies working with observations that are equidistant in calendar time (e.g., 5-minute sampling). Following Hansen and Lunde (2006), CTS is used in the derivation of realized variance measures (equation 7). Since the raw prices have irregularly spaced observations, artificial equally spaced prices have to be built. This research uses the *previous tick* method for this purpose. This method consists of using the observation at $t-1$ if the observation at t is missing and is preferred to the linear interpolation which creates undesirable properties of the integrated variance (Hansen and Lunde 2006).

Finally, as noted above, there is a concentration of quotes during the first and the last 15 minutes of a trading session in the live cattle futures markets (figure 2). Hasbrouck (2015) and Wang (2014) discard the first and last 15 minutes of each trading day. However, discarding these periods may eliminate price adjustments to changes in fundamentals that should be reflected in the IV. Additionally, there is no a priori reason to assume that noise will not be present during the first and the last fifteen minutes of a day trading session. By computing noise variance as the difference between the RV and IV, any price adjustments responding to fundamentals will not be considered as noise.

Results

In this section, we present the findings from the nonparametric model and the cointegration approach.

Nonparametric findings

Popularized by Andersen et al. (2000) and proposed by Fang (1996), volatility signature plots allow a graphical approximation to the bias of RV due to noise. For a specific period (e.g., 50 days), the plots compare the average of IV to average realized variance for the price series at different sampling intervals. Here, similar to Hansen and Lunde's (2006) empirical application, autocorrelation of order 30 is an effective filter of the time-dependence in the noise as the IV measure stabilizes. Hence, the unbiased estimate of the IV for each series (transaction (tr), bid, and ask prices) on each day is based on $RV_{ACNW_{30},t}^{1 tick}$, expressed in equation (8). An estimate of IV for the entire market is obtained by computing an average over the number of days, n ,

$$\overline{RV}_{ACNW_{30}}^{1 tick} = \frac{1}{n} \sum_{t=1}^n (RV_{ACNW_{30},t}^{1 tick,tr} + RV_{ACNW_{30},t}^{1 tick,bid} + RV_{ACNW_{30},t}^{1 tick,ask}) \frac{1}{3} . \quad (16)$$

Hansen and Lunde (2006) argue that this process is likely to yield conditionally unbiased IV measures. Realized variance for a specific series and specific frequency, m , is calculated using equation (7) for each day. This variance is then averaged over the number of days to obtain a short-run realized variance, $\overline{RV}_t^m = \frac{1}{n} \sum_{t=1}^n RV_t^m$. The difference between the short-run and the long-run RV represents the variance of noise for interval m (which varies here from 1 second to more than 4 minutes). Note we also estimate a short-run variance of the mid-quote price which is often used to reflect the equilibrium price.

Figure 3 presents the volatility signature plots for these most traded contracts in each year from 2011 to 2015. For instance, the top left panel presents the plots for the 2011 June contract, from April 8th to May 6th. The flat line is the IV measured as $\overline{RV}_{ACNW_{30}}^{1 tick}$ (equation 16), while the declining curves are the RV measures using different series (transaction, bid, ask, and mid-quotes prices). The difference between IV and RV is the bias due to microstructure noise that causes RV to overestimate the IV in all contracts. The magnitude of the overestimation depends on the sampling frequency (1 second versus 4 minutes) and declines as the sampling frequency diminishes. At a one-second sampling frequency, the transaction price has the highest embedded noise variance (four times the IV), followed in decreasing order by the ask, the bid, and the mid-quotes. The higher realized variance of transaction prices is driven by the well-known bid-ask bounce effect, such that transaction prices are typically negatively serially correlated moving between the bid and ask quotes. At the one-second sampling frequency, bids and asks' RVs are about three times the IV, while the mid-quote, the price most used to reflect the equilibrium price in the literature, is less than twice as large as the IV. At a 4-minute sampling frequency, the RV estimates appear unbiased as they converge to the estimate of the long-run IV.¹⁶ These results indicate that noise in live cattle markets spanned 4 minutes in 2011. While the other volatility signature plots follow the same general pattern over time, they differ by year to some degree. There

appears to be a relative decline in the noise, as convergence between transaction prices' RV and IV declines from 4 minutes in 2011-2013 to 3 minutes in 2014-2015.¹⁷

Table 1 provides numerical details of the volatility signature plots. The left-hand side of the table presents, for each year, the average estimated IV (corresponding to the horizontal line in the volatility signature plots in Figure 3), the average transaction price realized variance at the one-second frequency, its associated average noise, and the average noise normalized by realized variance to make comparison more meaningful. Highest IV levels are observed in 2015, 2011, and 2012 followed by 2013, and 2014. High IV levels in 2011 and 2012 are likely driven by droughts which adversely impacted U.S. live cattle production, motivated beef producers to send a large number of beef cows to slaughter, and generated considerable volatility in market prices. IV levels declined in 2013 and 2014 before reaching 2011-2012 levels in 2015. Relatively good weather conditions and low corn prices in 2013 and 2014 led feeders to increase cattle placement. The increased placement strategy shaped the supply response in 2015. High supplies in cold storage coupled with very heavy cattle leaving feedlots and high Australian imports of beef products, pushed prices down and increased volatility (Mathews and Haley 2015). Changes in the cattle cycle may have also come into play in 2015 (Hurt 2016).

The market noise variance followed a somewhat similar pattern. Noise variance was the highest in 2011 and 2012, declined sharply in the following two years, and increased again in 2015. Noise variance thus seems to be positively correlated with long-term IV levels, which suggests that relevant market price adjustments due to fundamentals can lead to higher market noise. When the annual noise variance is normalized by the realized variance, the importance of informational variance in causing high levels of realized variance in 2015 becomes more apparent. On a percentage basis, noise variance at the one-second frequency makes up about 66% of the realized variance in 2015, while it is greater than 70% in the other years.

The right-hand side of table 1 provides the same measures for the sample without limit price moves days. For the years that experienced limit moves (i.e., 2011, 2014, and 2015), realized and noise variance generally increased, with the largest increases occurring in 2015 as would be expected because of the large number of limit moves. Information variance is now largest in 2015, and its noise variance has increased to second behind 2012. Normalized variance still is the lowest in 2015 at 66.2%, again suggesting the importance of an informational variance in overall market price changes.

Robustness analyses are conducted using the least traded contracts in each year, and an average of the three realized variances (transaction, bid, and ask prices) to reflect market volatility. For the least traded contracts, the variances are systematically lower across years when compared to the most traded contracts. For instance, the realized variances are an average 38% lower for the least traded contracts. A similar pattern emerges, with the largest variances of noise in 2011 and 2012, but the time to convergence is reduced to about two and half minutes. Here the results are suggestive of higher volume increasing information and volatility, and could be reflective of the notion that high frequency traders prefer to trade in markets with the most profitable opportunities. When using an average of the three realized variances to represent market volatility, realized variances decrease by 21.7% relative to the transaction price variance. Convergence remains

similar to the pattern discussed earlier and the noise variance for 2015 declining by 31%. (Volatility signature plots and other details are in the statistical appendix.)

Semi-parametric findings

The daily cointegration analysis permits estimation of the efficient price, computed as the common stochastic component of observed prices. Estimation of the efficient price allows us to identify the relative loss in price due to the noise. The parameters in model (10) are estimated for each day. The optimum number of lags, l , is chosen in the range from 0 to 10, as the value that makes the Ljung-Box test insignificant at the 5% significance level. Annual averages of daily lag lengths are 3.54, 3.52, 3.72, 3.51, and 3.35 for years 2011, 2012, 2013, 2014, and 2015, respectively. Since the null hypothesis of no ARCH effects is systematically rejected for each day, the wild bootstrap method (1000 iterations for each day) is used in the estimation process. An alternative measure of the daily IV from equation (14), RV_{p^*} , is derived from the cointegration approach as the square returns of the efficient price for each day (table 2, column 5). As expected, the latter is found to be close to the IV from the volatility signature plots; similar findings provide a robustness check of the estimates.

In table 2, the averages of the alpha matrix components identify the instantaneous correlation between the efficient price and innovations in the observed prices. From 2011 to 2013, the estimates are very close to $(\bar{\alpha}_1^t, \bar{\alpha}_1^b, \bar{\alpha}_1^a) = (1/2, 1/4, 1/4)$, suggesting that transaction prices are more informative of the efficient price than quotes. This pattern becomes more accentuated in 2014 and 2015, when a substantial increase in the transaction price's coefficient has taken place (table 2, columns 2-4). Hence, in recent years, the live cattle market's transaction price has been more closely aligned to the efficient price, which seems to correspond to the rapid increases in prices that began in mid-2014 followed by the sharp decline in price that began at the end of 2014. Hansen and Lunde (2006) find that transaction prices are closely related to efficient prices in the NYSE, while the quotes prices are the closest to efficient prices for NASDAQ stocks in year 2004.

Volatility signature plots from the nonparametric analysis suggest that the bias caused by noise is positive, i.e., RV is systematically above IV. However, they do not identify the sign of the correlation between the efficient price and noise. Using equation (15), it is possible to show the components of RV. A positive bias can be obtained when the noise process is uncorrelated or positively correlated with efficient intraday returns. A positive bias can also be obtained when there is a negative correlation between observed returns and noise, if the downward bias caused by the negative correlation does not exceed the upward bias due to the RV of noise. The cointegration analysis allows us to identify the sign of this correlation (last column in table 2). The findings suggest that the correlation between the increments of noise and the efficient price returns is negative. This refines the results that noise variance increases when IV increases. If changes to fundamentals imply a decline in returns as occurred in 2015, noise will increase.

The cointegration framework also permits us to provide an indication of the economic measure (EM) of the noise. The log-noise at one tick sampling frequency can be derived by subtracting the log-efficient price obtained from the cointegration analysis (equation 12) from the observed transaction log-prices (equation 2). By applying the exponential function to the log-noise, an estimate in cents per pound is derived and put on a percentage price basis,

$$EM = \frac{e^{\ln(p_t) - \ln(p_t^*)} * 100}{p_t} \quad (17)$$

where p_t is the transaction price at time t , and p_t^* is the efficient price at time t . For each day, the median EM is identified, and for sample periods a box plot of the daily median EM s is developed (figure 4). When prices are transformed to cents per pound, the difference in the numerator of the EM is small, on average, 1 cent per pound during 2011-2015. For the most traded contracts, the median EM is 0.87% of the transaction prices in 2011 and 0.86%, 0.77%, 0.60%, and 0.73% in 2012, 2013, 2014, and 2015, respectively. When limit price moves for these contracts are excluded, the change is negligible, with the only change arising in 2012 when the EM increases to 0.87%.

Conclusions

Agricultural commodity futures markets' have experienced change with the emergence of electronic trading platforms in 2006. These platforms allow for automated trading based on computerized trading algorithms that are used for decision making, order entry, cancelation or modification. In recent years, financial markets have been affected by several events related to high speed trading activities such as the 2010 'flash crash' event, which cast doubt on the effect of high frequency traders on volatility and market efficiency. U.S. beef producers have blamed high frequency traders (HFT) for the high volatility observed in live cattle futures market in 2015. This paper identifies the market microstructure noise present in high frequency data and its implications for realized variance of returns in the U.S. live cattle futures markets from 2011 to 2015.

Using nonparametric and semiparametric methods and high frequency data for the most traded contracts, we identify volatility patterns in the live cattle futures market. The nonparametric procedures allow us to identify noise variance as the difference between short-run realized variance and long-run variance, which reflects changes in fundamental market conditions. The noise variance is viewed as a function of market frictions or imperfections which can include price discreteness, bid-ask bounce effects, infrequent trading, and HF trading. For the most traded contracts, where HFT are supposed to be more active, noise variance was high in 2011, reached its maximum in 2012, declined in the following two years, and increased again in 2015. Informational variance also varied by year and appeared to be generally consistent with changes in market fundamentals, such as droughts, changes in cold storage, and changes in beef imports. The magnitude of the noise depends on sampling frequency (1 second to 4 minutes). Transaction price variance has the highest embedded noise variance (four times the informational variance (IV)), followed by the bid, ask, and mid-quote variances. Informatively, all realized variances converge to the IV in a span of 3 to 4 minutes, indicating the effect of noise disappears quickly. At the one-second interval when the noise variance is at its highest, it reaches only 66% of the realized variance in 2015, while in other years it is greater than 70%. Removing limit price moves, which primarily occurred in 2015, from the sample increases the absolute magnitude of the noise in 2015, but the ratio of noise to realized variance changes only modestly. Other robustness checks on the least traded contracts showed about 40% less realized and noise variance, but similar noise to realized variance ratios. Convergence between realized variances and the IV also shortened to about two and half minutes. Changing the definition of market variance from the transaction price to an average of the bid, ask, and transaction prices (as was done in calculating IV), reduced most

of the variances by about 20%, with little change in convergence, and almost no change in the noise to realized variance ratio in 2015.

The parametric findings add insights on the behavior of prices, and the interaction between prices and the efficient prices. They also provide insights into the variance and the importance of noise. Transaction prices appear to be the important driver of efficient price, a relationship that is steadily strengthening. Noise varies negatively with movements in the efficient returns and prices, so that recent declines in price in the live cattle futures market led to higher levels of noise. Finally, noise is on average one cent per pound which represents between 0.6% and 0.9% of transaction prices over the 2011 to 2015 period.

On balance, the analysis provides limited if any evidence that noise increased in 2015 relative to other years. Noise variance can be large at short sampling frequencies, but it converges to informational variance in less than 4 minutes. Informational variance increased in 2015 which is consistent with the decline in the general level of prices due to fundamental factors in the market. As we observed in the parametric analysis, this decline in prices in 2015 may have strengthened to a modest degree the noise in the market (due to the negative covariance). However, in almost all contracts, the absolute value of the difference between the efficient price and the observed price is small and less than 1% of price.

Noise is found to be relevant and to mainly affect transactions prices which are the most strongly linked to daily changes in the efficient price. However, noise is short-lived and should not affect hedgers who don't usually change positions in 4-minute time intervals. It should not affect index funds either, as their business model does not imply trading at high frequency. Noise, however, will affect liquidity providers, as the increased variance in the bid and ask prices due to noise and the nonsynchronous revisions of the bid and ask quotes are likely to result in an increased variance in the BAS.

Due to data constraints, this research is limited regarding the conclusions that can be obtained on the market impacts of HFT, because available data do not classify traders into a high frequency trading category. Nevertheless, analysis of the realized variance suggests that a systematic problem is not evident. This appears to be consistent with the recent CME communication (CME 2016) that reports that HF traders make up about only 10% of the current market participants in the live cattle futures markets. This is consistent with the notion that market fundamentals were a driving factor of the increased market variance in 2015.

Figures



Figure 1. CME live cattle nearby futures prices in cents per pound, 2011-2016

Note: Source, Bloomberg Terminal. The intervals depicted by vertical dashed lines represent the period selected in each year of 60 – 10 days to expiration date of the most traded contract.

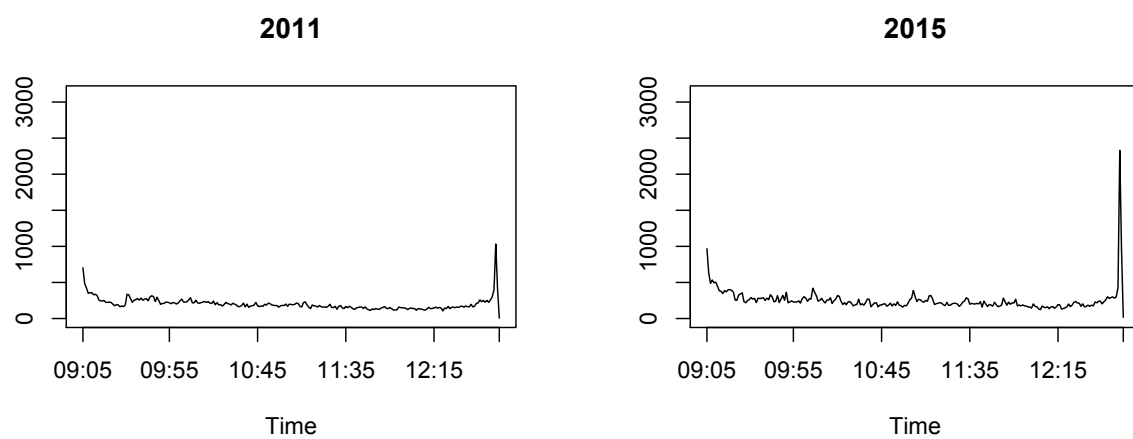


Figure 2. Average number of bid and ask quotes per minute for the most traded contracts and during the 60 - 10 days before maturity, 2011 and 2015

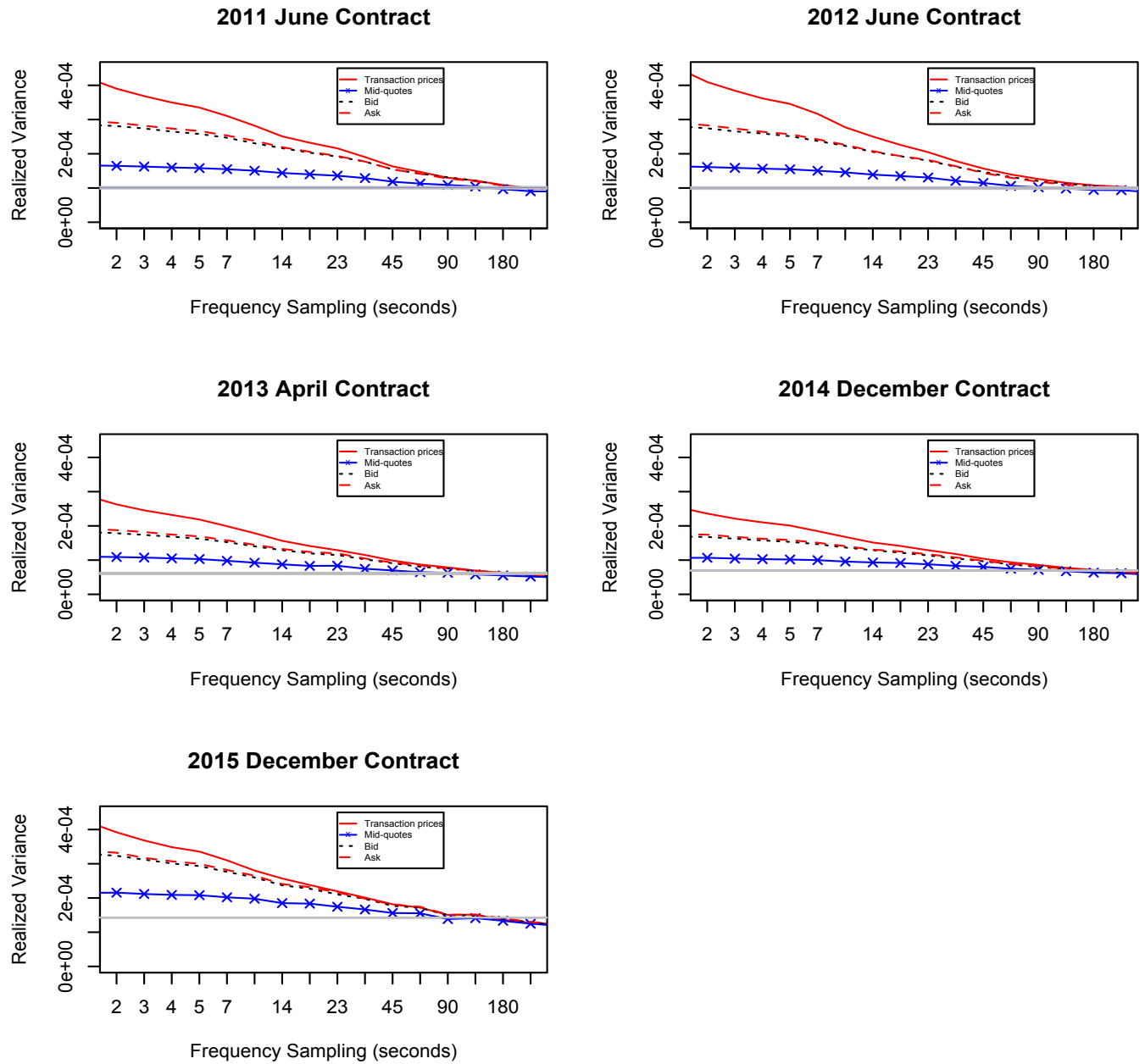


Figure 3. Volatility signature plots of the most traded contracts by contract, 2011-2015

Note: $RV^{(m)}$ is the realized volatility for transaction prices, mid-quotes, ask quotes, and bid quotes, where m is the frequency sampling from 1 second to 4 minutes. The horizontal line is the estimate of the IV. The sample period is from 60 to 10 days before maturity.

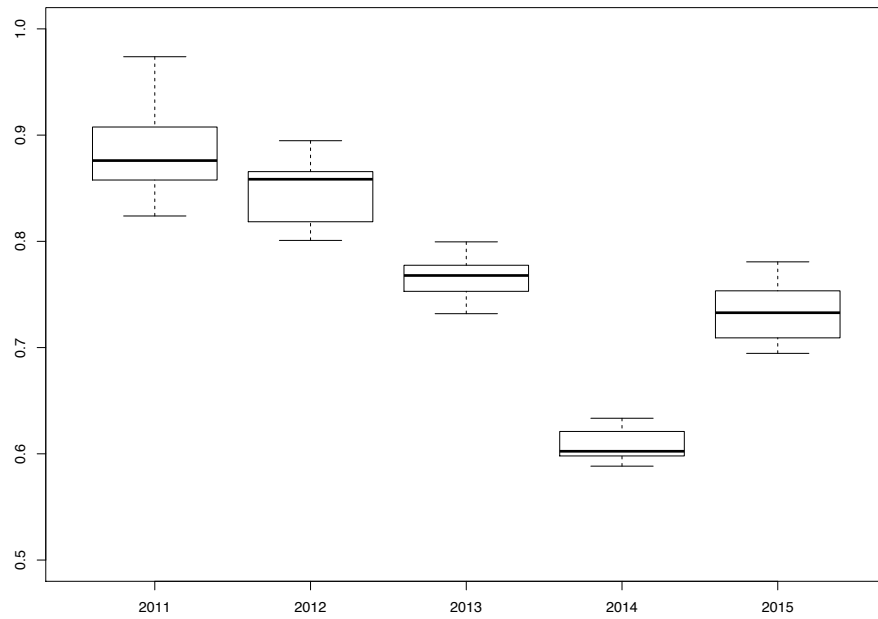


Figure 4. Daily median of noise as a percent of transaction price by contract, 2011-2015

Note: The observations by contract are the daily median noise as a percent of transaction price.

Tables

Table 1. Estimated IV, RV of Transaction Prices at One-second Frequency and the Noise Variance for the Most Traded Contracts with and without Limit Price Move Days, 2011-2015

Year	With limit price moves days				Without limit price moves days			
	IV	1-second RV on transaction prices	Noise Variance on transaction prices	Noise variance/1-second RV on transaction prices)	IV	1-second RV on transaction prices	Noise Variance on transaction prices	Noise variance/1-second RV on transaction prices)
	\widehat{IV}	$RV_t^{(1 \text{ sec})}$	$RV_t^{(1 \text{ sec})} - \widehat{IV}$	$\frac{RV_t^{(1 \text{ sec})} - \widehat{IV}}{RV_t^{(1 \text{ sec})}}$	\widehat{IV}	$RV_t^{(1 \text{ sec})}$	$RV_t^{(1 \text{ sec})} - \widehat{IV}$	$\frac{RV_t^{(1 \text{ sec})} - \widehat{IV}}{RV_t^{(1 \text{ sec})}}$
2011	1.01E-04 (0.98E-04,1.03E-04)	4.20E-04	3.19E-04	7.60E-01	1.02E-04 (0.94E-04,1.04E-04)	4.23E-04	3.21E-04	7.59E-01
2012	9.70E-05 (9.45E-05,9.95E-05)	4.41E-04	3.44E-04	7.80E-01	9.70E-05 (9.45E-05,9.95E-05)	4.41E-04	3.44E-04	7.80E-01
2013	6.10E-05 (5.89E-05,6.34E-05)	2.86E-04	2.25E-04	7.87E-01	6.10E-05 (5.89E-05,6.34E-05)	2.86E-04	2.25E-04	7.87E-01
2014	6.94E-05 (6.84E-05,7.04E-05)	2.54E-04	1.84E-04	7.24E-01	6.83E-05 (6.75E-05,6.91E-05)	2.54E-04	1.86E-04	7.31E-01
2015	1.43E-04 (1.42E-04,1.43E-04)	4.21E-04	2.78E-04	6.60E-01	1.68E-04 (1.67E-04,1.69E-04)	4.97E-04	3.29E-04	6.62E-01

Note: The $RV_t^{(1 \text{ sec})}$ stands for the RV of transaction prices. The confidence intervals for IV are constructed using $CI(\bar{\sigma}^2) \equiv \exp\left(\log(\hat{\sigma}^2) \pm \hat{\sigma}\left(c_{1-\frac{\alpha}{2}}\right)\right)$ where $c_{1-\frac{\alpha}{2}}$ are the 5 and 95 quantiles of the standard normal distribution, see Hansen and Lunde (2006).

Table 2. Cointegration Results, 2011-2015

Year	$\hat{\alpha}_{\perp}^t$	$\hat{\alpha}_{\perp}^b$	$\hat{\alpha}_{\perp}^a$	\widehat{RV}_{p^*}	\widehat{IV}	$\sum_t \mathbf{e}_{t_i} \mathbf{r}_{t_i}^*$
2011	0.55 (0.09)	0.23 (0.07)	0.22 (0.06)	1.09e-04	1.01e-04 (1.03e-04,0.98e-04)	- 6.56e-05
2012	0.53 (0.07)	0.24 (0.05)	0.23 (0.05)	9.91e-05	9.70e-05 (9.95e-05,9.45e-05)	-6.58e-05
2013	0.53 (0.08)	0.23 (0.07)	0.23 (0.06)	6.93e-05	6.10e-05 (6.34e-05,5.87e-05)	-4.11e-05
2014	0.62 (0.07)	0.19 (0.05)	0.19 (0.06)	7.03e-05	6.94e-05 (7.04e-05,6.84e-05)	-3.76e-05
2015	0.68 (0.15)	0.16 (0.07)	0.16 (0.13)	1.49e-04	1.43e-04 (1.43e-04,1.42e-04)	-6.72e-05

Note: This table presents the daily sample average from 2011 to 2015 of the $\hat{\alpha}_{\perp}$, RV of \hat{p}_t^* , the measure of IV using the kernel estimator from table 1, and the covariance between noise and return. The cointegration specification is estimated for each day. The parameters presented in this table are averaged over days. The average number of observations in one day were 2211, 2492, 2408, 2158, and 2360 in 2011, 2012, 2013, 2014, and 2015 respectively. The numbers in parentheses for the orthogonal alphas are standard deviations while the confidence intervals are reported for \widehat{IV} . The last column refers to the middle component of equation (14), where \mathbf{e}_{t_i} is the increment in intraday log-noise level and $\mathbf{r}_{t_i}^*$ is the intraday efficient price return.

Appendix for paper entitled “The Effects of Microstructure Noise on Realized Volatility in the Live Cattle Futures Market.”

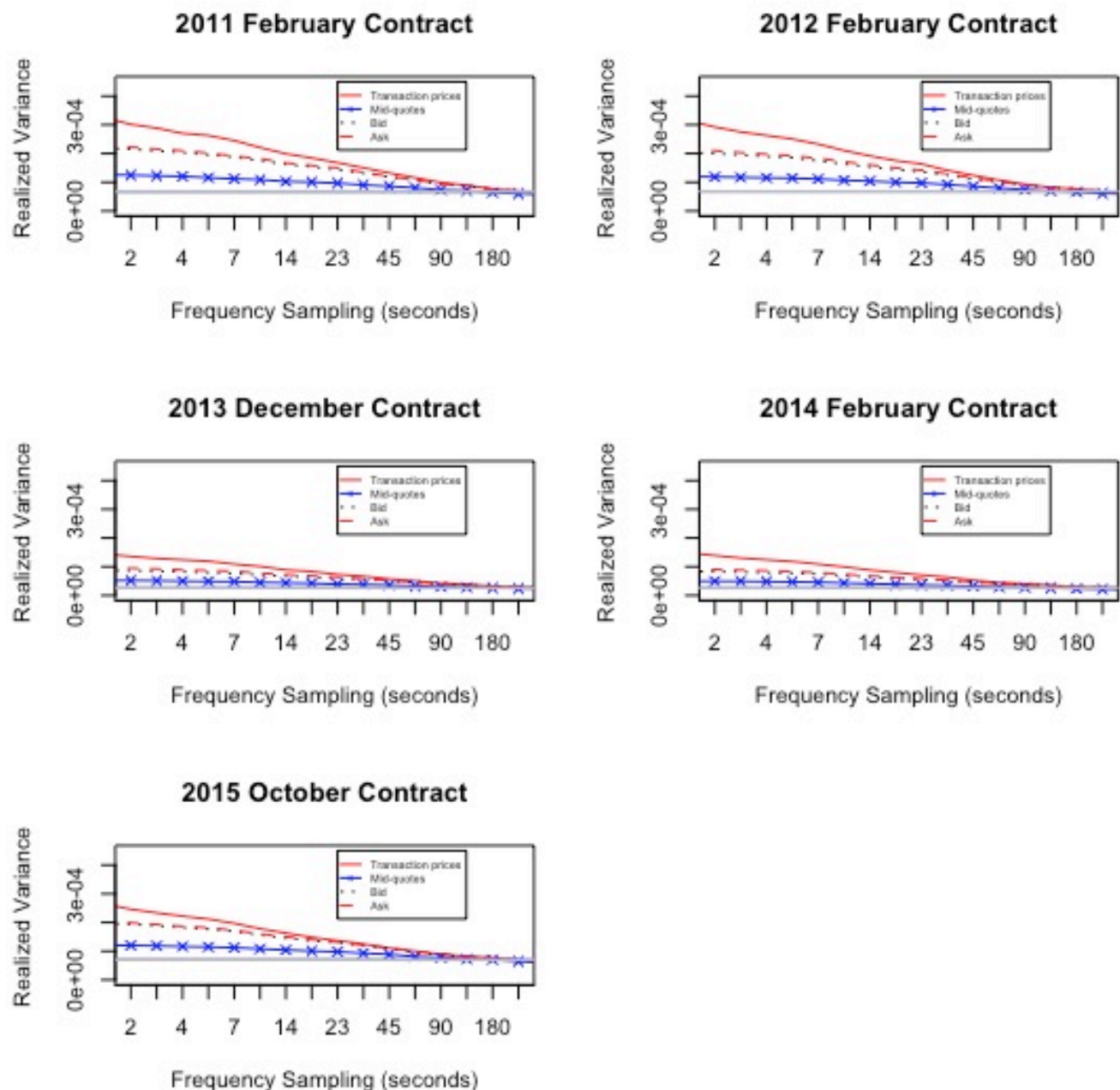


Figure A1. Volatility signature plots of the least traded contracts by contract, 2011-2015

Note: $RV^{(m)}$ is the realized volatility for transaction prices, mid-quotes, ask quotes, and bid quotes, where m is the frequency sampling from 1 second to 4 minutes. The horizontal line is the estimate of the IV. The sample period is from 60 to 10 days before maturity.

Table A1. Estimated IV, RV of Transaction Prices at One-second Frequency and the Noise Variance for the Least Traded Contracts, 2011-2015

	IV	1-second RV on transaction prices	Noise Variance on transaction prices	Noise variance/1-second RV on transaction prices)
Year	\widehat{IV}	$RV_t^{(1 \text{ sec})}$	$RV_t^{(1 \text{ sec})} - \widehat{IV}$	$\frac{RV_t^{(1 \text{ sec})} - \widehat{IV}}{RV_t^{(1 \text{ sec})}}$
2011	6.75e-05 (6.55e-05, 6.95e-05)	3.24e-04	2.57e-04	7.93E-01
2012	6.81e-05 (6.69e-05, 6.94e-05)	3.14e-04	2.46e-04	7.83E-01
2013	2.64e-05 (2.54e-05, 2.72e-05)	1.44e-04	1.17e-04	8.13E-01
2014	2.61e-05 (2.46e-05, 2.77e-05)	1.48e-04	1.21e-04	8.18E-01
2015	7.05e-05 (7.20e-05, 7.90e-05)	2.63e-04	1.93e-04	7.34E-01

Note: The $RV_t^{(1 \text{ sec})}$ stands for the RV of transaction prices. IV confidence intervals are constructed using $CI(\bar{\sigma}^2) \equiv \exp\left(\log(\hat{\sigma}^2) \pm \hat{\sigma}\left(c_{1-\frac{\alpha}{2}}\right)\right)$ where $c_{1-\frac{\alpha}{2}}$ are the 5 and 95 quantiles of the standard normal distribution, see Hansen and Lunde (2006).

Table A2. Cointegration Results for the Least Traded Contracts, 2011-2015

Year	$\hat{\alpha}_{\perp}^t$	$\hat{\alpha}_{\perp}^b$	$\hat{\alpha}_{\perp}^a$	\widehat{RV}_{p^*}	\widehat{IV} (from table 1)	$\sum_t \mathbf{e}_{t_i} \mathbf{r}_{t_i}^*$
2011	0.51 (0.07)	0.25 (0.06)	0.24 (0.07)	7.31e-05	6.75e-05 (6.55e-05, 6.95e-05)	-4.72e-05
2012	0.52 (0.08)	0.24 (0.05)	0.24 (0.06)	7.18e-05	6.81e-05 (6.69e-05, 6.94e-05)	-4.74e-05
2013	0.53 (0.07)	0.23 (0.06)	0.24 (0.06)	2.95e-05	2.64e-05 (2.54e-05, 2.72e-05)	-1.67e-05
2014	0.47 (0.08)	0.26 (0.06)	0.27 (0.06)	3.28e-05	2.61e-05 (2.46e-05, 2.77e-05)	-2.17e-05
2015	0.61 (0.13)	0.20 (0.08)	0.19 (0.10)	8.21e-05	7.05e-05 (7.20e-05, 7.90e-05)	-4.10e-05

Note: This table presents the daily sample average from 2011 to 2015 of the $\hat{\alpha}_{\perp}$, RV of \hat{p}_t^* , the measure of IV using the kernel estimator from table 1, and the covariance between noise and return. The cointegration specification is estimated for each day. The parameters presented in this table are averaged over days. The average number of observations in one day were 1698, 1872, 1561, 1597, and 2194 in 2011, 2012, 2013, 2014, and 2015 respectively. The numbers in parentheses for the orthogonal alphas are standard deviations while the confidence intervals are reported for \widehat{IV} . The last column refers to the middle component of equation (14), where \mathbf{e}_{t_i} is the increment in intraday log-noise level and $\mathbf{r}_{t_i}^*$ is the intraday efficient price return.

References

- Andersen, T.G., and L. Benzoni. 2008. "Realized Volatility." Working Paper No. 2008–14.
- Andersen, T.G., T. Bollerslev, F.X. Diebold, H. Ebens, D. Backus, M. Brandt, R. Deo, R. Engle, C. Granger, L. Hansen, L. Hentschel, C. Hurvich, P. De Lima, B. Schwert, and R. Stambaugh. 2001. "The Distribution of Realized Stock Return Volatility." *Journal of Financial Economics* 61:43–76.
- Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys. 2000. "Great Realizations." *Risk* 13:105–108.
- Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys. 2003. "Modeling and Forecasting Realized Volatility." *Econometrica* 71(2):579–625.
- Angelidis, T., and S. Degiannakis. 2008. "Volatility Forecasting: Intra-day versus Inter-day Models." *Journal of International Financial Markets, Institutions and Money* 18:449–465.
- Bandi, F.M., and J.R. Russell. 2003. "Microstructure Noise, Realized Volatility, and Optimal Sampling." Unpublished paper
- Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde, and N. Shephard. 2008. "Designing Realized Kernels to Measure the Ex-Post Variation of Equity Prices in the Presence of Noise." *Econometrica* 76(6):1481–1536.
- Barndorff-Nielsen, O.E., R.P. Hansen, A. Lunde, and N. Shephard. 2009. "Realized Kernels in Practice: Trades and Quotes." *Econometrics Journal* 12:C1–C32.
- Bollerslev, T., and T.G. Andersen. 1998. "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts." *International Economic Review* 39(4):885–905.
- Brogaard, J., T. Hendershott, and R. Riordan. 2014. "High-Frequency Trading and Price Discovery." *The Review of Financial Studies* 27(8):2267–2306.
- CME. 2016. "FAQ: CME Cattle Market Volatility." Available at: <http://www.cmegroup.com/trading/agricultural/cattle-volatility-faq.html>.
- Conrad, J., S. Wahal, and J. Xiang. 2015. "High-Frequency Quoting, Trading, and the Efficiency of Prices." *Journal of Financial Economics* 116(2):271–291.
- Dana, J., C.L. Gilbert, and E. Shim. 2006. "Hedging Grain Price Risk in the SADC: Case Studies of Malawi and Zambia." *Food Policy* 31(4):357–371.
- Egelkraut, T.M., P. Garcia, and B.J. Sherrick. 2007. "The Term Structure of Implied Volatility: Recovery and Informational Content in the Corn Options Market." *American Journal of Agricultural Economics* 89(1):1–11.

- Fang, Y. 1996. *Volatility Modeling and Estimation of High-Frequency Data with Gaussian Noise*. Ph.D. Thesis, Massachusetts Institute of Technology.
- Gousgounis, E., and E. Onur. 2016. “The Effect of Pit Closure on Futures Trading.”
- Hagströmer, B., and L. Nordén. 2013. “Granger’s Representation Theorem: A Closed-Form Expression for I(1) Processes.” *Journal of Financial Markets* 16(4):741–770.
- Hamilton, J.D. 1994. *Time Series Analysis* Volume 2. Princeton University Press, ed. Princeton.
- Hansen, P.R., and G. Horel. 2009. “Quadratic Variation by Markov Chains.” Research Paper No. 2009–13.
- Hansen, P.R., and A. Lunde. 2006. “Realized Variance and Market Microstructure Noise.” *Journal of Business & Economic Statistics* 24(2):127–161.
- Hasbrouck, J. 2015. “High Frequency Quoting: Short-Term Volatility in Bids and Offers.”
- Hasbrouck, J. 2013. “High Frequency Quoting: Short-Term Volatility in Bids and Offers.”
- Hasbrouck, J. 2002. “Stalking the ‘Efficient Price’ in Market Microstructure Specifications: An Overview.” *Journal of Financial Markets* 5(3):329–339.
- Haynes, R., and J.S. Roberts. 2015. “Automated Trading in Futures Markets.” CFTC White Papers
- Hurt, C. 2016. “Weekly Outlook: Falling Cattle Prices, Where Is the Bottom?” *farmdoc daily* 83(6).
- Irwin, S.H., and D.R. Sanders. 2012. “Financialization and Structural Change in Commodity Futures Markets.” *Journal of Agricultural and Applied Economics* 44(3):371–396.
- Kalnina, I., and O. Linton. 2008. “Estimating Quadratic Variation Consistently in the Presence of Endogenous and Diurnal Measurement Error.” *Journal of Econometrics journal* 147(1):47–59.
- Mathews, K., and M. Haley. 2015. “Livestock, Dairy, and Poultry Outlook.” *Situation and Outlook ERS, USDA*.
- Meyer, G. 2016. “Cattlemen Lock Horns with Futures Exchange over Market Volatility.” *Financial Times*.
- Newey, W.K., and K.D. West. 1987. “Hypothesis Testing with Efficient Method of Moments Estimation.” *International Economic Review*:777–787.
- O’Hara, M. 2015. “High Frequency Market Microstructure.” *Journal of Financial Economics* 116(2):257–270.

- Protter, P.E. 2005. “Stochastic Differential Equations.” In *Stochastic Integration and Differential Equations*. Springer Berlin Heidelberg, pp. 249–361.
- Stebbins, C. 2013. “Analysis: High Speed Trading a Stiff Challenge for U.S. Regulators.” *Reuters*.
- Wang, X. 2014. *Price Volatility and Liquidity Cost in Grain Futures Markets*. Ph.D. Thesis, University of Illinois.
- Zhang, L., P.A. Mykland, and Y. Aït-Sahalia. 2005. “A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data.” *Journal of the American Statistical Association* 100(472):1394–1411.
- Zhou, B. 1996. “High-Frequency Data and Volatility in Foreign-Exchange Rates.” *Journal of Business & Economic Statistics* 14(1):45–52.

Endnotes

¹ While no official definition exists, CME characterizes high frequency traders as “proprietary firms that submit more than two automated order messages per second per instrument measured over the course of a given time period.” (CME 2016).

² The rest of the contracts were traded via open outcry in the pit. CME closed pit trading in almost all futures markets, including live cattle market in July 2015.

³ High frequency traders might cause uncertainty by canceling, revising, and resubmitting orders. This uncertainty may result in increased price volatility.

⁴ Price discreteness refers to tick size. A tick is the minimum increment allowed for a price change. Tick size, which did not change during the period studied, is 0.025 cent/pound for live cattle futures contract (CME Group’s website).

⁵ Bid-ask bounces can arise from trades that randomly occur at the bid or ask prices, but it more commonly refers to the spread that liquidity traders require to make the market liquid. These effects introduce dependencies into the price dynamics and the random walk hypothesis may no longer be a reasonable assumption (Hasbrouck 2002).

⁶ The realized kernel is a nonparametric method which requires the selection of a bandwidth, here the autocorrelation order.

⁷ Hansen and Lunde (2006), for example, discuss sampling when using realized kernel estimator based on calendar time, tick time (i.e. trade event), as well as business time (intervals of constant IV).

⁸ In this case, time is sampled at intervals defined by a geometric sequence; for instance, 1, 2, 4, 8, ... seconds.

⁹ The time stamp at the millisecond scale is simulated using the Bayesian Markov Chain Monte Carlo.

¹⁰ The derivation of this estimator is found in Hansen and Lunde (2006), page 141. The notation follows Newey and West (1987).

¹¹ This approach is preferred to actual estimation of the cointegration vectors on a daily basis, as intra-daily data may not properly reflect the market equilibrium.

¹² One could question the effect of closure of the pit trading on the liquidity costs in the live cattle futures market in 2015. This question has been examined by Gousgounis and Onur (2016). They do not find any significant influence of pit closure on price or effective spread in the electronic market.

¹³ To save space, only two years are presented in figure 2. The figure for 2012 is similar to the one in 2011, whereas the graphs for 2013 and 2014 depict a higher peak at the end of the day than in 2011 and 2012 but lower than in 2015.

¹⁴ Barndorff-Nielsen et al. (2009) also delete entries when transaction prices are higher than the ask plus bid-ask spread or lower than the bid minus the bid-ask spread. These outliers represent less than 0.9% for each year were deleted from the analyzed sample.

¹⁵ This translates into 2211, 2492, 2408, 2158, and 2360 observations (daily average in most traded contracts) in 2011, 2012, 2013, 2014, and 2015, respectively.

¹⁶ Convergence is defined when the difference between the RV of transactions prices and the estimated IV is less than 1%.

¹⁷ Results are similar when TTS is used instead of CTS.