MICROSTRUCTURE NOISE AND REALIZED VARIANCE IN THE LIVE CATTLE FUTURES MARKET

ANABELLE COULEAU, TERESA SERRA, AND PHILIP GARCIA

Recently, U.S. live cattle futures prices have experienced high levels of intraday price variance, which have raised concerns about the possible impact of microstructure noise from high frequency trading on market instability. This article identifies both the magnitude and the duration of the bias caused by market microstructure noise in measuring efficient price variance in the live cattle futures market from 2011 to 2016, with emphasis on price variance behavior in recent years. Market microstructure noise increases observed price variance, but its effects are not large and do not last more than three to four minutes in response to changing information. Intraday price variance has increased in recent years, but the findings provide little evidence that high frequency traders were responsible for economically meaningful market noise. Informatively, steps taken by the CME and cattle producers to mitigate noise have not been fruitful to date, and signal that the magnitude of noise will likely vary with the magnitude of changes in demand and cyclical supply.

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

JEL codes: C32, C36, G18, Q13.

The debate on the effects of HFT on market efficiency is active (Brogaard, Hendershott, and Riordan 2014; Wang 2014; Conrad, Wahal, and Xiang 2015; Hasbrouck 2018). A question that arises is whether bursts of HFT intensify price fluctuations, add noise to the market, and thus confound the price discovery process. This can lead observed prices to fluctuate due to both changes in the market efficient price and noise. The noise has a time dependency and thus its effects do not dissipate immediately. Our research identifies both the magnitude and the duration of market microstructure noise for the measurement of efficient price variance in the live cattle futures market. Previous literature studying other markets has found that these pervasive effects have a limited duration of 10 minutes or less (Andersen et al. 2000; Kalnina and Linton 2008). Hence, while noise may be present throughout the day as events in the market unfold, it can only be identified using high frequency data.

In 2015 and 2016, intraday live cattle futures prices experienced heightened variance (figure 1). In response, the National Cattlemen’s Beef Association (NCBA) held an internal meeting on December 2015 with...
industry traders and hedgers. The debate and the conclusions reached revolved around the idea that increased intraday variance causes difficulties in managing risk using futures contracts by complicating hedging activities. The NCBA attributed heightened variance to high frequency traders who, by means of their high-speed trading activities, may have increased intraday order execution risk. Later, the NCBA met with the Chicago Mercantile Exchange (CME) to discuss the problem and proposed changes in trading rules, and with the Commodity Futures Trading Commission (CFTC) in hopes of resolving the situation (Anderson 2016; Beltway Beef 2017).

To assess the effects of HFT on market microstructure noise, we need to focus on intraday data. While noise generated by HFT is present throughout the day, it is not likely transmitted across days. High frequency traders do not hold overnight positions (Baron, Brogaard, and Kirilenko 2012), limiting their ability to transfer intraday impacts to the next day. Additionally, noise effects have a short-term duration, and go to zero at the end of the trading session when the market closes.

Direct research evidence on the effects of HFT in agricultural commodity markets is limited, since public data currently available do not identify high frequency traders. As a result, we cannot isolate the proportion of the noise caused by HFT or by other market imperfections or frictions such as price discreteness, bid-ask bounce effects, or infrequent trading (Hagstro¨mer and Nordén 2013; Wang 2014; O’Hara 2015; Hasbrouck 2018). Nevertheless, it is possible to shed light on the situation by identifying market microstructure noise using high frequency data and the extent to which this noise distorts efficient price variance measures. Specifically, in the context of the live cattle futures market, identifying the extent to which intraday price variance in 2015 and 2016 was due to market participants incorporating information about fundamentals or to noise that may in part be attributable to HFT activities is informative to the decision makers attempting to find solutions to these pricing problems. This research contributes to this understanding.

Despite the absence of research on the causes of the live cattle futures market price variance during 2015 and 2016, the CME
responded to beef producers’ concerns by introducing a series of changes to the trading environment. Changes include incorporating the live cattle futures contract into the Message Efficiency Program (MEP), a reduction in trading hours, and a change in the live cattle futures contract specifications. While these actions by the CME are understandable, complaints by producers and the CME’s responses need to be weighed against evidence regarding how much market noise is generated by HFT and the impacts of this noise on intraday price variance (Stebbins 2013). It is also important to assess the effectiveness of the changes introduced by CME as a response to producers’ concerns, as well as the effectiveness of other measures that were already in place such as the limit price moves (LPM). We conduct an analysis to assess their effect on noise variance.

To our knowledge, no published work has shed light on market microstructure noise in agricultural commodity markets. High-frequency price variance is composed of a permanent and a transitory component. The permanent component (sometimes called the integrated variance) reflects the efficient price variance that would prevail in a frictionless market and is driven by information flows often measured by trading volume. The transitory component is the variance due to short-term frictions arising from market microstructure noise. The literature has provided different methods to purge the high frequency price variance of its noise component and extract the efficient variance dynamics. We adopt two different approaches. The first approach is based on the well-known fact that market microstructure noise induces high-frequency return autocorrelation, which leads to biased variance estimates. Following Hansen and Lunde (2006), we use autocorrelations to “bias-correct” the variance measure. Another alternative is to identify the efficient price itself and calculate its variance subsequently (Hansen and Lunde 2006). In this case, the efficient price is assumed to be the stochastic trend common to the bid quote (reflecting the demand side of the market), the ask quote (reflecting the supply side), and the transactions price (reflecting the equilibrium reached between the two). Also, we assess how noise changes through time. To the extent that high frequency traders generate noise in the market, their increased presence (Brogaard, Hendershott, and Riordan 2014) should be reflected in higher noise levels. Finally, we contribute to recent heated policy debates on whether HFT should be regulated in agricultural futures markets.

Our findings suggest that the particularly high intraday variance in live cattle markets in 2015 and 2016 was strongly influenced by market participants incorporating information about fundamentals, captured by the integrated variance (IV). Noise is found to substantially distort the measure of price realized variance (RV) at a one-second sampling frequency, but its effect dissipates in three to four minutes. Distortions caused by noise are especially important during periods of relevant efficient price variance. Noise is, on average, $0.01 per pound and represents between 0.6% and 0.9% of the transaction prices over the period studied. CME changes in the live cattle futures market have had little effect on mitigating noise. Overall, the results cast doubt on the notion that HFT was responsible for the high variance in cattle markets in 2015 and 2016.

**Market Microstructure Noise Identification Methods**

Methods to capture intraday market microstructure noise depend on the assumptions of the properties of the noise. We consider two key properties of noise. First, noise has a mean of zero and is time dependent, which can be captured by its stationary autocovariance, $\pi(s) = E(\ln(\nu_t)\ln(\nu_{t+s}))$, where $\nu_t$ is the microstructure noise and $t$ is a time subindex. The time dependence of market microstructure noise induces autocorrelation in intraday observed price returns. Second, market microstructure noise returns can be correlated with efficient price returns. Following Hansen and Lunde (2006) and Bandi and Russell (2003), market microstructure noise is characterized by equation (1),

$$\ln(\nu_t) = \ln(p_t) - \ln(p_t^*)$$

where $p_t$ is the observed price at time $t$, and $p_t^*$ is the latent efficient price. 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 that generates noise in quote prices and increases frictional variance (Wang 2014; Hasbrouck 2018).
The latent or efficient log-price process is assumed to follow a Brownian semi-martingale (Hamilton 1994) and can be represented as equation (2),

$$d\ln(p_t^e) = \sigma_t dW_t,$$

where $W_t$ is a standard Brownian motion and $\sigma_t$ is a (continuous) random volatility function. The $IV$, which reflects the efficient price variance free of microstructure noise, is defined as follows:

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

Here, $IV$ measures the stochastic arrival of new information over time, and in an efficient market reflects how participants incorporate information and expectations about the market. For the purpose of empirical analysis, the time interval $[0, T]$ can be divided into $m$ discrete intraday sub-intervals, $[t_{i-1}, t_i ]$ with $t_0 = 0 < \ldots < t_m = T$. Using (1), intraday observed realized returns for each interval can be written as

$$r_{i,m} = r_{i,m}^* + e_{i,m}$$

where $r_{i,m} = \ln(p_{t_i,m}) - \ln(p_{t_{i-1},m})$, $r_{i,m}^* = \ln(p_{t_i,m}^e) - \ln(p_{t_{i-1},m}^e)$ and $e_{i,m} = \ln(\nu_{t_i,m}) - \ln(\nu_{t_{i-1},m})$ for $i = 1, \ldots, m$.

The $RV$ captures the total variation in prices sampled at the intraday time sequence, by summing the squares of the price changes (returns): $RV^{(m)} = \sum_{i=1}^m (\ln(p_{t_i,m}) - \ln(p_{t_{i-1},m}))^2 = \sum_{i=1}^m r_{i,m}^2$. In the absence of microstructure noise, this has been shown to provide a consistent estimator of the $IV$ as the time between observations tends to zero. However, in the presence of noise, $RV$ can be expressed as in equation (5), and provides a biased and generally an inconsistent estimate of the $IV$ (Bandi and Russell 2003):

$$RV^{(m)} = \sum_{i=1}^m (r_{i,m}^e)^2 + 2\sum_{i=1}^m e_{i} r_{i,m}^e + \sum_{i=1}^m (e_{i,m})^2.$$

The three components on the right-hand side of equation (5) represent, respectively, the efficient price $RV$, the correlation between efficient price and noise returns and the $RV$ of noise. The sum of the two last components can be referred to as noise bias ($NB$). In practice, ignoring microstructure noise only seems to work well for sampling frequencies of 10 minutes or more, for which the $RV^{(m)}$ seems to be free of microstructure noise and thus to reflect the $IV$ (Hansen and Lunde 2006; Kalnina and Linton 2008).

To identify market microstructure noise, we follow Hansen and Lunde (2006), who propose both nonparametric and semiparametric methods. Both methods capture $NB$. However, the second approach allows us to disentangle the two components (i.e., the time dependence and noise correlation with efficient price returns) of the $NB$, while the nonparametric method does not.

Nonparametric Identification of Noise

This section presents a nonparametric approach to measure $NB$ that is based on the comparison between $RV$ and $IV$ using high frequency data. Zhou (1996) was the first to introduce a $RV$ estimator that attempts to isolate the $IV$. The Zhou $RV$ estimator corrects for $NB$ through a first-order autocorrelation term. Hansen and Lunde (2006) show that Zhou’s estimator is not robust, and requires 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 and noise returns. Hansen and Lunde (2006) recommend computing the $IV$ using tick-time sampling as opposed to calendar time, to better capture the time dependence in noise. Their generalized estimator uses a Bartlett-based kernel that can be expressed as

$$RV^{(1)}_{ACNW,k} \equiv \hat{\gamma}_0 + \sum_{j=1}^k (\hat{\gamma}_{-j} + \hat{\gamma}_j)$$

$$+ \sum_{j=1}^{k} \frac{k-j}{k} (\hat{\gamma}_{-j-k} + \hat{\gamma}_{j+k}),$$

where $\hat{\gamma}_j = \sum_{i=1}^N r_{i+j}$, $\hat{\gamma}_0 \equiv r_t^2$, and $k \geq 2$ is the order of autocorrelation. A progressive increase in $k$ will change equation (6) as the amount of noise filtered increases. Hansen and Lunde (2006) choose the value of $k$ that renders equation (6) stable (i.e., a further increase in $k$ does not lead to further change in $IV$). In their empirical application, an autocorrelation of order 30 is selected. We
estimate the IV through equation (6) for each day using intraday tick data.

To draw conclusions from this approach, we then visually compare, using variance signature plots, an average of IV, \( \hat{RV}^{\text{tick}} \), to an average of the daily observed price RV (denoted by \( \overline{RV}^m \) and measured using calendar-time sampling to approximate the duration of noise). Further, \( \overline{RV}^m \) is estimated for progressively longer intraday time intervals, \( m \). The difference between the \( \overline{RV}^m \) and the \( \hat{RV}^{\text{ACNW}} \), represents the NB for interval \( m \). By increasing the length of \( m \), we can observe the length of time needed for NB to disappear.

**Semiparametric Identification of Noise**

The nonparametric approach does not allow us to disentangle the time dependence of noise from the noise return correlation with efficient price return correlation. An alternative method to quantify NB involves estimating the efficient price using cointegration methods. By using a vector error correction model (VECM), we identify the efficient market price, represented by the common stochastic component between the observed quotes (i.e., bid and ask representing, respectively, the demand and the supply side of the market) and transaction prices (representing the equilibrium reached between the two parts). This approach allows quotes and prices to deviate from each other in the short-run due to noise, but imposes a market equilibrium that is eventually reached and is represented by the common stochastic trend, or the efficient price free from microstructure frictions. Once the efficient price is estimated, the IV is derived as the efficient price RV and compared to the observed prices RV to approximate NB. An advantage of using this method is that it allows us to decompose NB into its two components.

Let \( t_i \) for \( i = 0, 1, \ldots, I \) denote the time when transactions occur during a trading day (i.e., tick time). The vector of log-observed nonstationary prices is given by

\[
\begin{pmatrix}
\text{transaction price at time } t_i \\
\text{corresponding ask price at } t_i \\
\text{corresponding bid price at } t_i
\end{pmatrix}
\]

and the VECM, which can be estimated by least squares, is

\[
\Delta p_{t_i} = \alpha \beta p_{t_{i-1}} + \sum_{j=1}^{l-1} \Gamma_{1,j} \Delta p_{t_{i-j}} + \sum_{j=0}^{l-1} \Gamma_{2,j} \text{vol}_{t_{i-j}} + \mu + \epsilon_{t_i}
\]

where \( \mu = \alpha \beta \) is a 3x1 restricted vector of constants, \( \rho = (\rho_1, \rho_2)' \), being \( -\rho_1 \) the average difference between transaction prices and mid-quotes, and \( -\rho_2 \) the average bid-ask spread, while \( \alpha \) is a 3x2 matrix (Hansen and Lunde 2006). Following Hasbrouck (1991), our VECM specification includes trade volumes (\( \text{vol}_{t_{i-j}} \)) which were found to be stationary, weakly exogenous, and not granger caused by quotes. As a result, volume is included as a strongly exogenous variable in equation (8). The error, \( \epsilon_{t_i} \), \( i = 0, 1, \ldots, I \), is assumed to be an uncorrelated error vector, \( \beta \) is a 3x2 matrix, \( l \) is the number of lags, and \( \Gamma_{1,j} \) and \( \Gamma_{2,j} \) are the parameters capturing the short-run dynamics. The three observed prices are assumed to share the same stochastic trend (i.e., the efficient price). The cointegration rank is assumed to be known and equal to two. The first cointegrating relationship \( (\beta_1 p_{t_i}) \) represents the long-run link between the transaction price and the quotes, and the second \( (\beta_2 p_{t_i}) \) is used to represent the long-run pattern of the bid-ask spread. As a result, \( \beta \) can be expressed as

\[
\beta = (\beta_1, \beta_2) = \begin{pmatrix} 1 & 0 \\ -1/2 & 1 \\ -1/2 & -1 \end{pmatrix}^2
\]

To identify \( \alpha \) and \( \beta \), the following normalization vectors are imposed, \( \beta_1 = (1 1 1)' \) and \( \alpha_1 (1 1 1) = 1 \), where \( \beta_1 \) and \( \alpha_1 \) are 3x1 vectors (Hansen and Lunde 2006). Identification of the common stochastic trend representing the efficient price, follows Hasbrouck (2002) and is based on the Granger representation.

\[\text{We test for the restrictions imposed on the cointegrating vectors and find them to hold on nearly 90\% of the days in the sample.}\]
From equation (10), the trade volumes from CME Group’s BBO and the intraday price variance increased substantially in 2015 (see figure 1). The Wilcoxon test for price variance in 2015 rejects the null that the average price variance was constant across the period. As noted, livestock futures trading is characterized by a predominance of electronic platform trading in live cattle futures. As of January 2011 through December 2016, which represented about 80% in 2011 (Irwin and Sanders 2012) and reached 100% when the CME live cattle futures contract was traded with six maturities a year: February, April, June, August, October, and December. The analysis focuses on a nearby contract series to reflect that most trading occurs in the current contract. We roll over to the next contract when trading volume in the nearby is below the trading volume of the next delivery contract for two consecutive days.

Several issues can emerge when working with high-frequency data that can bias research results. These include: (a) misplaced decimal or abnormal zero prices, (b) several quotes or trade data being time stamped to the same second, and (c) the presence of LPM. Following Barndorff-Nielsen et al. (2009), pre-processing data procedures are applied to the data selected for analysis to overcome these issues. First, all zero-priced bids, asks, and transactions are deleted. Second, since multiple quotes and prices can have the same time stamp, they are replaced with the median bid and ask quotes and transaction prices as proposed by Barndorff-Nielsen et al. (2009) and Hansen and Lunde (2006). LPM occur when price reaches either the minimum or the maximum price change allowed by the exchange. LPM can occur at any time in the trading session and can take various forms. For instance, 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 throughout the day. Alternately, prices can reach the limit price for a period of time and then revert back to a trading region. When prices are at the limit, RV, IV, and NB are all reduced to zero for that period of time. Here, we define the LPM days when the nearby transaction prices hit the price limit up or down and stay locked for at least 30 minutes until the end of the trading day. Using this criterion, we find five LPM days in 2011 (four limit up and one limit down).
down), two LPM days in 2013 (both up), six LPM days in 2014 (three up and three down), 14 LPM days in 2015 (evenly directionally split with nine occurring between September and December), and nine LPM days in 2016 (three up and six down). Careful examination of the LPM revealed the USDA releases of cattle-related reports (e.g., cold storage, Livestock Slaughter, WASDE announcements) as possible causes of price limits. However, several limit moves appeared unrelated to USDA information releases. Measures of RV, IV, and NB should be reduced when the limit moves are included. However, relative measures (NB as a proportion of RV or NB as a proportion of IV) should be less affected.

Another issue that arises with intraday data is the sampling scheme to use for analysis. Two intraday sampling schemes are primarily used in research. The first is tick-time sampling (TTS), which is based on the time a transaction occurs and involves unequal temporal spacing between observations. Following Hansen and Lunde (2006), to better capture the time dependence of noise, TTS is employed in the estimation of the IV in both the Bartlett-based 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 can occur when forcing unevenly-spaced observations to be evenly distributed. Throughout the sample, quotes are reported in approximately 50% of the total number of seconds (i.e., 14,100 seconds) within the day trading session, resulting in an average time between observed quotes of 2 seconds. In contrast, 6 seconds separate observed trades. Since transaction observations are spaced every 6 seconds on average in the sample, tick-time sampling involves around one-sixth of the total seconds in a trading day. 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). Calendar-time sample is used in the derivation of observed price RV ($RV^{(m)} = \sum_{i=1}^{m} i^2 m^2$). This allows us to easily approximate the duration of noise in the variance signature plots. Since the raw prices have irregularly-spaced observations, artificial equally-spaced prices have to be built. This research uses the previous-tick method. The 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 IV (Hansen and Lunde 2006).

Nonparametric Variance Findings

Proposed by Fang (1996) and popularized by Andersen et al. (2000), variance signature plots allow a graphical approximation to the bias of RV due to noise. For a number of days, the plots compare the average of IV to average RV at different sampling intervals. Since our IV (as well as RV) measures are derived for each observed price, an unbiased IV estimate is obtained by computing an average over the number of days, $n$,

$$\overline{RV}^1_{ACNW_{30}} = \frac{1}{3} \left( RV^{1 \text{ tick}, \text{ tr}}_{ACNW_{30}} + RV^{1 \text{ tick}, \text{ bid}}_{ACNW_{30}} + RV^{1 \text{ tick}, \text{ ask}}_{ACNW_{30}} \right) \tag{13}$$

The RV for a specific series and specific frequency, $m$, is calculated for each day. This variance is then averaged over the number of days to obtain $\overline{RV}^m_t = \frac{1}{n} \sum_{t=1}^{n} RV^{(m)}$. The difference between this estimate of the RV and the IV represents the NB for interval $m$ (which varies here from 1 second to more than 4 minutes). Note that we also estimate a RV of the mid-quote price, which is often used to reflect the equilibrium price. The results are organized and presented in three periods. The first period, 2011–2014 reflects relatively low intraday variance, while 2015 and 2016 with high intraday variance are examined separately.

Figure 2 presents the variance signature plots. The horizontal line in each panel is the IV for that period measured as $\overline{RV}^1_{ACNW_{30}}$. The declining curves are the RV measures using different series (transaction, bid, ask, and mid-quote prices) at different sampling frequencies. The difference between IV and RV is the bias due to microstructure noise that causes RV to overestimate the IV. 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 NB (four times the IV), followed in decreasing order by the ask, bid, and mid-quotes. The higher RV of transaction prices is driven in part by the well-known bid-
ask bounce effect that creates a negative serial correlation in transactions prices as they move between bid and ask quotes. Also, transaction prices tend to respond more and more quickly to information. At the one-second sampling frequency, bids and asks’ RVs are about three times the IV, while the mid-quote RV, the price most used to reflect the equilibrium price in the literature, is less than twice as large as the IV. At a three-four minute sampling frequency, RV estimates appear unbiased as they converge to the estimate of the IV. In 2011–2014, the noise in live cattle markets spanned four minutes. While the other plots follow the same general pattern over time, they differ to some degree by year. In 2015 and 2016, the RVs are higher as are the IVs. However, in both years RVs converge to the IVs more quickly, reaching the IV in only three minutes.

Table 1 presents selected details of the signature plots. For each period, the table provides estimated IV (which corresponds to the horizontal line in figure 2), and for one-second frequency sampling the transaction price RV, NB, and the NB normalized by RV. When excluding LPM days, highest IV levels are observed in 2016 and 2015 followed by 2011–2014. While 2011–2014 had the lowest IV levels, its confidence interval was the largest due to the heterogeneity in annual estimates. Annual IV estimates in 2011 and 2012, which were likely driven by droughts that motivated producers to send large numbers of beef cows to slaughter, were high, nearly reaching 2015 levels. In contrast, 2013 and 2014 were the least volatile years. 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 variance (Mathews and Haley 2015). Changes in the cattle cycle appear to have come into play in 2015 and 2016 (Hurt 2016).

The NB followed a similar pattern; NB was the highest in 2016, followed by 2015, and

---

5 Convergence is defined when the difference between the RV and IV is less than 1%.

---

**Figure 2. Variance signature plots for the live cattle nearby futures contracts, 2011–2016**

Note: Realized variance (RV$^m$) for transaction prices, mid-quotes, bid and ask quotes, where m is the sample frequency (1 second - 4 minutes). The horizontal line is the IV estimate. Days with limit-price moves longer than 30 minutes are excluded.
2011–2014. NB thus seems to be positively correlated with long-term IV levels, which suggests that relevant market price adjustments due to the inflow of information about fundamentals can lead to higher market noise. When NB is normalized by the RV at the one-second frequency, NB makes up about 71% to 75% of the RV. While absolute NB is the highest in 2015 and 2016, it represents a smaller portion of the RV than in 2011–2014.

Table 1 also provides the same variance measures including the LPM days. For the periods that experienced a high number of LPM (e.g., 2015), RV and NB generally decreased as expected; IV is still the largest in 2016. Normalized variances do not change, which was verified using a Wilcoxon signed rank nonparametric test.

**Semiparametric Findings**

The 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 noise. The parameters in model (8) are estimated for each day. The optimum lag length, $l$, is chosen in a 0 to 10 range, as the value that makes the Ljung-Box test insignificant at the 5% significance level. Averages of daily lag lengths are 1.90, 1.71, and 1.87 in 2011–2014, 2015, and 2016, respectively. Since the null hypothesis of no ARCH effects is systematically rejected, the wild bootstrap method (1,000 iterations for each day) is used in the estimation process.

Table 2 presents average estimates from the cointegration model for the three periods identified earlier. Daily IV is derived from equation (12), as the square returns of the efficient price (table 2, column 5). These semiparametric measures are close to the IV from the variance signature plots, providing a robustness check. Table 2 also provides averages of the alpha matrix components that identify the instantaneous correlation between the efficient price and innovations in the observed prices. For 2011–2014, the estimates are very close to $(\hat{\zeta}_{1}, \hat{\zeta}_{b}, \hat{\zeta}_{a}) = (1/2, 1/4, 1/4)$, where $\hat{\zeta}_{1}$, and $\hat{\zeta}_{a}$ are the 5% and 95% quantiles of the standard normal distribution (Hansen and Lunde 2006). Days with limit-price moves longer than 30 minutes were excluded in the left panel.
takes place (table 2, column 2) with the coefficient decreasing slightly in 2016. Hence, in recent years, the live cattle market’s transaction price has been more closely aligned with the efficient price, which seems to correspond to the rapid decline in price that began at the end of 2014 and continued through most of 2016.6

Recall that the variance signature plots suggested that the NB is positive, that is, RV is systematically above IV. However, these plots do not identify the sign of the correlation between the efficient price and noise. Using equation (5), it is possible to show the components of RV. A positive bias can be obtained when the noise return 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 NB increases when IV increases. If changes to fundamentals imply a decline in returns, as appeared to have occurred in 2015 and 2016, noise will increase.

The cointegration framework also permits us to provide an economic measure (EM) of the noise by comparing observed and efficient price. Specifically, the log-noise at a 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 1). By applying the exponential function to the log-noise, an estimate in cents per pound is derived which can be put on a percentage price basis,

\[ EM = \frac{e^{\ln(p_t) - \ln(p^*_t)} - 100}{p_t} \]

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 our sample periods a box plot of the daily median \( EM \)s is developed (figure 3); \( EM \) is small, on average, $0.01 per pound during 2011–2016. In 2011–2014, the median \( EM \) is 0.79% of transaction prices, and 0.67% and 0.86% in 2015 and 2016, respectively.7

**Noise Bias Analysis, 2015–2016**

As noted, due to increased variance cattle producers have raised concerns, and met with the CME to resolve the variance issues in the live cattle futures contracts. The CME has responded by taking steps to stabilize the live cattle market, but recent releases from the Cattlemen Association indicate that concerns continue. Here, we examine the characteristics of noise and the effects of the steps taken by CME to reduce unnecessary noise. We focus on the period of September 2015 to December 2016, which starts just before the dramatic drop in prices and increased variance.

---

6 Live cattle futures prices declined from $170 to nearly 95 cents per lb. in this period.

7 When LPM days are included, there is no change at 2 decimal places.
Figure 4 presents the daily RV, IV, and NB, derived from nonparametric procedures. Vertical lines are inserted to identify LPM days (black) and CME changes (red). A supply change is also included (green) to reflect the sharp increases in slaughter numbers identified by Hurt (2016) in an outlook publication. The day when the USDA agricultural marketing service introduced online auctions is also included (blue; USDA-Agricultural Marketing Service 2016). Visual examination substantiates the statements that variance has increased. Recall from our earlier analysis that variance was higher in 2015 than in earlier years, and here it can be seen that all three measures increased around mid-March in 2016 and have remained high. Since the NB increased, it should be clear that RV increased more than IV. While the timing of the events in the market and CME can be identified, it is difficult to identify the effect of these on market microstructure noise. For this purpose, we use a straightforward regression framework, which we modify for statistical reasons.

We estimate the following model:

\[
NB_t = \beta_0 + \beta_1 NB_{t-1} + \beta_2 IV_t + \beta_3 VOL_t + \beta_4 LPM_t + \beta_5 D1_t + \beta_6 D2_t + \beta_7 D3_t + \beta_8 D4_t + \beta_9 D5_t + \beta_{10} SPR_t + \beta_{11} SUM_t + \beta_{12} AUT_t + u_t
\]

where \(NB_t\) is the daily noise bias at a one-second sampling frequency and \(IV_t\) is the daily estimated integrated variance, both computed from the nonparametric approach. Further, \(NB_t\) and \(IV_t\) are divided by the corresponding last transaction price of the day session to standardize them for the change in the level of prices that occurred in the period; \(IV\) is included to control for the level of fundamental information entering the market, but also to assess the degree to which noise is influenced by information arrival. Lagged daily noise bias \((NB_{t-1})\) is included to assess the extent to which noise is transmitted across days, while \(VOL_t\) is the daily volume traded in the nearby contract. Volume, which is an indication of the information arrival, can also reflect liquidity in the market, so that higher levels may reduce \(NB\); finally, \(u_t\) is the error term.

Dummy variables are added to represent changes in CME policies. These dates correspond to the vertical lines in figure 4. CME implemented three recent regulation changes in the live cattle futures market to reduce the high variance and improve the reliability of the live cattle futures prices. The first change involved

---

\(^8\) Note that these measures represent variance for a particular day, in contrast to the variances in the signature plots that reflect averages for specific intervals over a year. As a result, they appear to be more volatile.
the addition of live cattle futures contracts to the Message Efficiency Program (MEP) on February 1, 2016. The MEP is designed to reduce potentially harmful high frequency activity in the market. High frequency traders can use different messaging strategies (e.g., spoofing) to gain advantage. Essentially, a firm’s messages are counted and used to monitor daily trading activity. If a firm’s message to volume traded ratio exceeds a limit established by CME or the total daily message count exceeds 20,000, it is subject to a $1,000 fine per product per day.\footnote{More details on MEP and further discussion of its limited success are found on their website (CME 2017).}

Also, $D_1$ is a dummy variable equal to one on February 1, 2016 and afterwards, and zero otherwise. Less than a month later, the CME reduced the trading hours in live cattle futures contracts to more closely align trading hours (now, 8:30 a.m. to 1:05 p.m.) with the period of greatest contract activity. Matching trading hours and demand can focus the liquidity in the market to when it is most needed and reduce variance; $D_2$ equals one on February 29, 2016 and afterwards, zero otherwise.

Finally, the CME implemented changes directed at the cash live cattle market. Futures contracts rely on cash market information to function well. A lack of transparency in the cash market can hinder the efficiency of futures market in reflecting fundamental information and can create additional basis risk when hedging. After discussions with the cattle industry, the CME modified the specifications of the contract to permit a seasonal discount in the South Dakota delivery location. This was done to align more effectively delivery values with cash market prices, and maintain compliance with CFTC’s policy on location price differentials (CME 2016). In addition, it revised grading quality, and delayed listing additional contracts beyond October 2017, giving time for the cattle industry and CME to find solutions to improve cash market transparency; further, $D_3$ equals one on August 5, 2016 and afterwards, and zero otherwise.

Figure 4. Daily realized variance ($RV$), integrated variance ($IV$), and noise bias ($NB$), September, 2015–December, 2016

Note: The $RV_t$ is the realized variance computed on intraday transaction prices at one-second sampling frequency, the $IV_t$ is the daily integrated variance from the nonparametric approach, and the daily $NB_t$ is the difference between $RV_t$ and $IV_t$. In the top panel, vertical black dashed lines represent limit-price move (LPM) days. In the bottom panel, vertical lines represent CME market regulations and supply shocks events. The policy and supply shocks from left to right are: 2016-02-01 when live cattle futures contracts were included in the CME MEP (red vertical line); 2016-02-29 when CME modified the day trading hours in the live cattle futures market (red vertical line); 2016-03-16 when the number of cattle slaughtered strongly increased (green vertical line); and 2016-08-05 when the CME modified the live cattle futures contract specifications (red vertical line). Finally, the last vertical line (blue) is 2016-10-05 when the USDA incorporated information from the first online auction implemented by the fed cattle exchange.
In addition, we also include several other discrete variables. Dummy variables are included to control for limit-price move days \((LPM_t)\) equals one on the limit price move days, and the large increase in slaughter numbers in mid-March reflecting cyclical supply changes (Hurt 2016; \(D4\), equals one beginning on March 16, 2016 and after, zero otherwise). We also include a dummy variable to reflect efforts by the cattle industry to improve cash market transparency and reduce unnecessary variance in futures contracts by providing additional transaction information. Beginning on October 5, 2016, the USDA agricultural marketing service began including online auction transactions from Superior Livestock Auction's fed cattle exchange \((D5_t)\) equals 1 on October 5 and afterwards, and zero otherwise). Finally, we allow for seasonal effects (Karali and Power 2013) through three dummies capturing spring, summer, and autumn \((SPR_t, SUM_t, AUT_t\) respectively).

Because of endogeneity, autocorrelation, and heteroscedasticity concerns, we estimate equation (15) using GMM procedures. The equation is estimated using fitted \(IV_t\) (based on an autoregressive model of order 5-AR5) and fitted \(VOL_t\) (based on an AR3). Table 3 presents the estimated findings. General tests indicate that the instruments are valid (Wald F-statistic) and well selected (Kleibergen and Paap 2006) for under-identification and instrument redundancy at the 1% level. GMM residuals are absent of autocorrelation and heteroscedasticity (Pagan and Hall 1983; Cumby and Huizinga 1992) at the 5% level.

Table 3. GMM Results of Noise Bias in the Live Cattle Futures Market, 2015–2016

<table>
<thead>
<tr>
<th>Dependent variable: (NB_t)</th>
<th>Estimate GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.73e-06***</td>
</tr>
<tr>
<td>(NB_{t-1})</td>
<td>0.14</td>
</tr>
<tr>
<td>(lagged noise variance)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>(IV_t)</td>
<td>1.24***</td>
</tr>
<tr>
<td>(integrated variance)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>(VOL_t)</td>
<td>-7.64e-11*</td>
</tr>
<tr>
<td>(daily trading volume)</td>
<td>(3.95e-11)</td>
</tr>
<tr>
<td>(LPM_t)</td>
<td>-1.00e-06***</td>
</tr>
<tr>
<td>(limit price moves)</td>
<td>(2.62e-07)</td>
</tr>
<tr>
<td>(D1_t)</td>
<td>1.75e-07</td>
</tr>
<tr>
<td>(MEP)</td>
<td>(1.10e-07)</td>
</tr>
<tr>
<td>(D2_t)</td>
<td>2.71e-07***</td>
</tr>
<tr>
<td>(trading hours change)</td>
<td>(9.01e-08)</td>
</tr>
<tr>
<td>(D3_t)</td>
<td>8.53e-07 ***</td>
</tr>
<tr>
<td>(futures contract</td>
<td>(2.44e-07)</td>
</tr>
<tr>
<td>specifications changes)</td>
<td></td>
</tr>
<tr>
<td>(D4_t)</td>
<td>6.87e-07 ***</td>
</tr>
<tr>
<td>(supply shock)</td>
<td>(2.29e-07)</td>
</tr>
<tr>
<td>(D5_t)</td>
<td>-1.74e-07</td>
</tr>
<tr>
<td>(online auction)</td>
<td>(2.81e-07)</td>
</tr>
<tr>
<td>(SPR_t)</td>
<td>4.44e-07**</td>
</tr>
<tr>
<td>(SUM_t)</td>
<td>(1.79e-07)</td>
</tr>
<tr>
<td>(AUT_t)</td>
<td>1.95e-08</td>
</tr>
<tr>
<td></td>
<td>(1.11e-07)</td>
</tr>
<tr>
<td></td>
<td>4.92e-07***</td>
</tr>
<tr>
<td></td>
<td>(1.60e-07)</td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses. Significance levels of 1%, 5%, and 10% are represented by ***, **, and *, respectively. We use two instrumental variables that are \(a\) fitted values using an AR(5) for \(IV_t\), and \(b\) fitted values using an AR(3) for \(VOL_t\). Noise bias \(NB_{t-1}\) is the daily lagged noise bias of transaction prices at one-second sampling frequency estimated from the nonparametric approach, \(IV_t\) is the daily integrated variance from the nonparametric approach, \(NV_t\) and \(IV_t\) were divided by the corresponding transaction price at the end of trading day, \(LPM_t\), and \(D_i\) \((i = 1, \ldots, 5\) denote the dummy variables representing limit-price moves and days with policy or supply shocks, respectively. Further, \(SPR_t\), \(SUM_t\), and \(AUT_t\) denote spring, summer, and autumn, respectively. These are described in the text in further detail.

Traders do not usually hold overnight positions (Baron, Brogaard, and Kirilenko 2012), hence their intraday impact is not transmitted to the next day.

To date, CME changes to increase stability in the contract appear to not have been fruitful. Both the change to trading hours to realign liquidity with user demands, and changes to the specifications of the contract have resulted in an increase in \(NB\). Of these two, the changes to the contract resulted in a more than three times \((8.43e-07 compared to 2.71e-07)\) larger effect on \(NB\). These increases in \(NB\) may be a reflection of added uncertainty introduced into the market with a change in the nature and perhaps even
existence of the contract. Informatively, the MEP variable had little effect on NB. This lack of importance points to the limited (if any) presence of high frequency traders in the live cattle market.

Two factors that influenced the NB appreciably were the LPM and cattle slaughter supply variables. This confirms that the presence of LPM reduces realized variance and its components. The magnitude of change is the largest among the discrete variables. The increase in cattle slaughter starting in mid-March 2016 seems to have signaled the beginning of the end of the cattle cycle. This is rather clear ex post, but during the process the signals in the market may have been uncertain and the behavior of traders unclear. This tumultuous period may have led to errors in interpreting the market and large mistakes (and thus noise) in the positions taken.

Finally, cattle producers’ efforts to provide added information to the cash market have decreased the NB but not significantly in a statistical sense. Changing seasonal patterns suggest a peak in the early spring and autumn that is somewhat in tandem with normal seasonal price patterns (Hurt 2015), providing another indication that noise accompanies the large swings in market prices.10

Concluding Remarks

Agricultural commodity futures markets experienced an important change with the emergence of electronic trading in 2006, which enabled the use of computerized algorithms for decision making, order entry, and cancelation. U.S. beef producers have blamed high frequency traders who can operate on these platforms for the high intraday price variance observed in the live cattle futures market in 2015 and 2016. Using high frequency data and nonparametric and semiparametric methods, this article generates the realized variance (RV) of observed prices in the U.S. live cattle futures market from 2011 to 2016, and provides estimates of the noise and integrated variance components. We also examine the effects of recent changes by the CME and cattle producers to reduce realized variance in hopes of reducing market execution risk.

Realized variance can be decomposed into its two components. Noise bias (NB) is influenced by the frictions in the market including the bid-ask bounce, tick size, infrequent trading, and high frequency trading activities. Integrated variance (IV) reflects the stochastic arrival of new information, and provides a measure of how market participants, through their buying and selling activities, incorporate information and expectations about fundamentals in the market. Examination of the estimated intraday variances over time and for different temporal intervals provides insights into their relationships and their sources.

Over time, our findings point to the notion that level of variance is heavily influenced by fundamental changes in the market. For instance, the magnitude of noise bias (NB) was the lowest in 2011–2014 (a relatively stable period), and then increased gradually in 2015 and in 2016. Integrated variance (IV) also followed a similar pattern, increasing in recent years. The pattern is consistent with events in the live cattle market, particularly in 2015 and 2016. In 2015, high supplies in cold storage coupled with very heavy feedlot sales and high Australian imports of beef products pushed prices down and increased variance (Mathews and Haley 2015). Changes in the cattle cycle through increasing slaughter also appears to have come into play in 2015 and 2016 (Hurt 2016), driving down prices but increasing variance. Pronounced changes result in added information to the market as participants modify their positions. These adjustments in a market venue that permits quick response inevitably leads to heightened variance and added noise, which may trigger large intraday price movements. This interpretation of the importance of fundamental factors in affecting realized variance is supported by the strong IV, supply shock (reflecting the increase in cattle slaughtering starting in mid-March 2016), and seasonal effects in the GMM estimation.

Assessment of the estimated variances for different temporal intervals permits a more detailed view of the relationships. The magnitude of NB depends heavily on sampling frequency and decreases quickly as the interval is expanded. Transaction price RV has the highest embedded NB (at the one-second frequency, it is approximatively four times the IV), followed by the bid, ask, and mid-quote variances. Informatively, when using the

10 Program coding files can be found in the supplementary appendix online.
mid-quote, which reduces frictions caused by the normal bid-ask bounce, the RV is less than twice as large as the IV. Regardless, all RVs converge to the IV in a span of three to four minutes, indicating that noise dissipates quickly. When compared to other markets in the literature, the length of the noise bias is relatively short (Andersen et al. 2000; Kalnina and Linton 2008).

To date, changes by the CME and cattle producers to increase price stability have not been fruitful. Intuitively, changing trading hours and contract specifications to align more closely in time and form with market needs should add liquidity and lead to less noise. But changing market conditions and uncertainty about the impacts of these changes on trading and hedging may have overwhelmed the expected effects. In contrast, cattle producers’ efforts to provide added information to the cash market through the introduction of online auction transactions have decreased the NB, but not significantly. Perhaps, over a longer period, these changes will have their desired and expected effects. An aspect of market environment that did reduce NB was the presence of the CME’s limit-price structure. But the reduction in NB on a given day can come with an added cost as market participants are unable to close their positions. Regardless, research should be initiated to consider more carefully the effect of price-limits and perhaps how they might be more effectively structured to reflect market conditions. This research can be directly motivated by the magnitude and changing trading execution costs that participants face.

Overall, the analysis finds little to support the notion that high frequency traders were responsible for added intraday variance in the live cattle futures market in recent years. There appears to be little direct carryover of NB from day to day, which is consistent with the observation that high frequency traders do not hold overnight positions in the live cattle market. The absence of a HFT effect is also supported by the limited success of the CME’s messaging efficiency program specifically designed to reduce this type of NB. Noise bias did increase in 2015 and 2016 relative to other years. However, IV also increased in 2015 and 2016, which is consistent with the sharp decline in the general level of prices due to fundamental factors in the market. In a pricing context, the presence of noise bias from the arrival of information dissipates quickly and the absolute value of the difference between the efficient price and the observed price is small—less than 1% of price. However, in a hedging context, execution risk remains. High execution risk can limit hedging activities and in the longer-term adversely affect the sustainability of the contract.

Supplementary Material

Supplementary material are available at American Journal of Agricultural Economics online.

Endnotes


References


