PRICE JUMPS AND VOLATILITY IN U.S. AGRICULTURAL FUTURES MARKETS

BY

ANABELLE COULEAU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Agricultural and Applied Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Philip Garcia, Chair and Director of Research
Professor Teresa Serra, Chair and Director of Research
Professor Scott Irwin
Assistant Professor Andrés Trujillo-Barrera, University of Idaho
ABSTRACT

Agricultural commodity futures markets have changed with the arrival of electronic trading. Electronic trading platforms have facilitated the emergence of automated systems in these markets which are now experiencing a race among traders to gain speed in implementing transactions. This new trading environment raises questions about the increasing role of high-speed traders and their effect in agricultural futures markets. In the first two essays in this dissertation, I examine how recent structural changes associated with increased speed of trading and release of public information experienced by these markets affect prices and volatility dynamics. In the third essay, I investigate whether more flexible research approaches should be employed to provide market participants and policy markets more accurate volatility forecasts within the context of the new more heterogeneous trading environment.

The first essay 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. The U.S. live cattle futures prices have experienced high levels of intraday price variance, starting in 2015, which have raised concerns about the possible impact of microstructure noise from high frequency trading on market instability. 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.

The second essay demonstrates that jumps in corn futures prices have increased with electronic trading and the shift to real-time announcement of USDA reports. Using intraday prices from 2008 to 2015, we employ a nonparametric test to detect jumps and variance analysis to estimate jump or execution risk. Real-time trading of major USDA reports has substantially increased the frequency and clustering of price jumps, and results in higher market liquidity costs. In contrast, while the presence of jumps on non-announcement days has doubled recently, their magnitude has declined as have transactions costs during their occurrence. The largest jump risk or execution risk is experienced by high frequency traders due to heightened microstructure noise during price jumps.

The third essay investigates the ability of artificial neural network (ANN) to forecast realized volatility in the corn futures market. Forecasting volatility is complicated by heterogeneous expectations from a diversity of traders and by nonlinearities such as seasonality or public information shocks (USDA announcements). Recent applications of artificial neural networks in econometrics suggest this model is particularly suited in capturing unknown nonlinearities forms. Using corn futures prices observed between 2009 and 2017, this paper compares the volatility forecasting performance of nonlinear autoregressive ANN models against other alternative linear specifications, such as the heterogeneous autoregressive (HAR) model which account for heterogeneity in volatility expectations. Our findings indicate that the nonlinear ANN model works better at all horizons (1-day, 1-week, and 1-month) than the standard HAR model even when the HAR model is augmented for seasonality or public information shocks, pointing to the importance of accounting for unknown forms of nonlinearities through more flexible approaches.
ACKNOWLEDGEMENTS

I would like to thank my advisors, Philip Garcia and Teresa Serra, for their guidance, discussions, and advice during my time as a Ph.D. student. Their advice always made me want to learn more. Thanks for the time and energy you devoted to making me grow as a researcher. The thesis also benefited from comments and suggestions made by Scott Irwin and Andrés Trujillo-Barrera. Thank you for willing to be part of my dissertation committee.

I would like to thank all the faculty members I had the chance to interact with. In particular, I thank David Bullock and Nick Paulson for their helpful guidance while I was their research and teaching assistant, respectively. I am also grateful to the members of the Department of Agricultural and Consumer Economics staff, in particular, Pam and Melissa who make our working environment efficient and enjoyable.

Completing this work would have been all the more difficult without the support from my friends at the University of Illinois and back home, in France. Thanks for always cheering me up. Finally, I would like to thank Leo, my husband, who always believed in me and whose love and encouragement are with me in whatever I pursue. Thank you for your continued support during this journey.
# TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION .............................................................................................................. 1

CHAPTER 2: MICROSTRUCTURE NOISE AND REALIZED VARIANCE IN THE LIVE CATTLE FUTURES MARKET ......................................................................................................................... 4

CHAPTER 3: ARE CORN FUTURES PRICES GETTING “JUMPY”? .................................................. 41

CHAPTER 4: FORECASTING THE CORN FUTURES REALIZED VOLATILITY ................................ 87

CHAPTER 5: CONCLUSIONS ................................................................................................................. 141
CHAPTER 1
INTRODUCTION

Over the last decade, automation in trading has led to considerable changes in U.S. agricultural commodity futures markets. Computerized algorithms reached 32% of the trading volume in the livestock futures market in 2014 and increased to 59.5% by October 2018, while they increased from 39% to 55.2% of the grain and oilseed futures trading volume (Haynes and Roberts 2015; 2017; 2019). Such rapid increase not only questions the role of heightened speed in agricultural futures markets but also how conventional and new traders interact in this new trading environment and the implications for market behavior. The increased availability of futures markets intraday data has facilitated new research that aims to respond to these research questions. This thesis provides three essays that shed light on two main issues. The first and the second essays focus on how the new trading environment has affected the functioning of agricultural futures markets. The third essay investigates whether more flexible research methods should be adopted to provide decision makers with more accurate information.

The first essay investigates the role of high frequency trading (HFT) on the efficiency of the live cattle futures market. More specifically, I aim at responding to what extent high frequency traders can move the futures market price away from its fundamental value by means of their trading strategies. This research was motivated by the sharp decline in the live cattle futures market price and increased intraday price volatility in 2015 and 2016, which increased beef producers’ difficulties in managing risk through futures contracts. The National Cattlemen’s Beef Association attributed the heightened variance to high frequency traders who, through high-speed trading activities, may have increased intraday execution risk. This essay investigates whether bursts of HFT intensify price fluctuations, add noise to the market, and thus confound the price discovery
process. To do so, we identify both the magnitude and duration of the bias caused by market microstructure noise in measuring efficient price variance in the live cattle futures market.

The second essay studies the impacts of a policy change under the new trading environment. More specifically, we investigate the effects of a change in the release of USDA public information in the era of electronic trading. Traditionally, reports were released after trading hours, but in a recent policy change, USDA now releases them when markets are open, allowing for real-time trading on new information. Commercial traders have complained that the new policy leaves them no time to digest information released, and creates an unfair playing field that favors nonconventional high frequency trading firms that can quote, revise, and execute orders in milliseconds. While previous research (Adjemian and Irwin 2018) has already documented the presence of large volatility spikes that dissipate within a few minutes during USDA report releases, it has not assessed the relevance of price jumps within these spikes. Price volatility has two components: a jump-free volatility component that is easy to predict and hedge, and a jump component that cannot be predicted and is difficult to hedge. To the extent that jumps are present during the USDA announcement days, price discovery is less predictable, and hedging is more problematic. Additionally, jumps lead to a substantial increase in execution risk that is relevant to all traders willing to take or liquidate positions. This second essay pays close attention to the impacts of USDA announcements on the jump component of price volatility in the corn futures market.

The third essay investigates the ability of artificial neural networks (ANN) to forecast price volatility under the new trading environment. The ability to forecast agricultural futures price volatility is critical to inform market participants of their future risk exposure, and to guide their production, hedging, and inventory decisions. Changes in market behavior in the era of electronic
trading call for a reassessment of the issue. Forecasting volatility in the presence of long-memory
 can be complicated by heterogeneous volatility expectations from a diversity of traders. Depending
 on their information needs, market participants are likely to be interested in futures price volatility
 at different horizons. The nonlinearities such as seasonality or public information shocks (USDA
 public announcements) also influence volatility persistence, complicating volatility forecasts in
 the corn futures market. Recent applications of artificial neural networks suggest that this model
 is particularly suited to capture unknown form of nonlinearities. This third essay explores the
 benefits of using nonlinear realized volatility models estimated through ANN using intraday prices
 and compares them to linear model specifications.

1.1. References

Papers Series

White Papers Series

White Papers Series
CHAPTER 2
MICROSTRUCTURE NOISE AND REALIZED VARIANCE IN THE LIVE CATTLE FUTURES MARKET

2.1. Introduction

The U.S. live cattle futures market experienced particularly high intraday variance in 2015 and 2016. Beef producers placed the primary responsibility on high frequency trading (HFT) activities. HFT became possible with the emergence of electronic trading in agricultural futures markets in 2006. By 2011, livestock futures trading on the electronic platform reached 80% of total trading volume (Irwin and Sanders 2012) and about 95% in 2015 (Gousgounis and Onur 2016). 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 debate on the effects of HFT on market efficiency is active (Brogaard, Hendershott, and Riordan 2014; Conrad, Wahal, and Xiang 2015; Hasbrouck 2015; Wang 2014). 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).

---

1 See the letter that the National Cattlemen’s Beef Association (NCBA) addressed to Terrence A. Duffy, CEO of CME on January 13, 2016 (http://www.beefusa.org/CMDocs/BeefUSA/Media/NCBAlettertoCMEreHFT.pdf).
Hence, while noise may be present throughout the day as events in the market unfold, it can only be identified using HF data.

In 2015 and 2016, intraday live cattle futures prices experienced heightened variance (Figure 2.1). 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. 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, NCBA met with 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. HF 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 HF 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 (Hasbrouck 2015; Wang 2014; Hagströmer and Nordén 2013; O’Hara 2015). Nevertheless, it is possible to shed light on the situation by identifying market microstructure noise using HF data and the extent to which this noise distorts efficient price variance measures. Specifically, in the context of 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 the 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 HF 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 one cent 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.
2.2. 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 mean zero and is time dependent which can be captured by its stationary autocovariance, \( \pi(s) = E(\ln(v_t), \ln(v_{t+s})) \), where \( v_t \) 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 (2.1),

\[
\ln(v_t) = \ln(p_t) - \ln(p_t^e),
\]

where \( p_t \) is the observed price at time \( t \), and \( p_t^e \) 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 HF quoting which generates noise in quote prices and increases frictional variance (Hasbrouck 2015; Wang 2014).

The latent or efficient log-price process is assumed to follow a Brownian semi-martingale (Hamilton 1994, p. 477) and can be represented as (2.2),

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

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,

\[
IV \equiv \int_0^T \sigma(t)^2 \, dt.
\]

\( IV \) 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,
\([t_{i-1,m}, t_{i,m}]\) with \(t_{0,m} = 0 < \cdots < t_{m,m} = T\). Using (2.1), intraday observed realized returns for each interval can be written as,

\[
    r_{i,m} = r_{i,m}^* + e_{i,m},
\]

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

\[
    RV^{(m)} = \sum_{i=1}^{m} \left( \ln(p_{t_{i,m}}) - \ln(p_{t_{i-1,m}}) \right)^2 = \sum_{i=1}^{m} r_{i,m}^2.
\]

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 (5), and provides a biased and generally an inconsistent estimate of the \(IV\) (Bandi and Russell 2003).

\[
    RV^{(m)} = \sum_{i=1}^{m} (r_{i,m}^*)^2 + 2 \sum_{i=1}^{m} e_{t_i} r_{t_i}^* + \sum_{i=1}^{m} (e_{t_i})^2.
\]

The three components on the right-hand side of (2.5) 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 \(IV\) (Kalnina and Linton 2008; Hansen and Lunde 2006).

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.
2.2.1. **Nonparametric Identification of Noise**

This section presents a nonparametric approach to measure $NB$ that is based on the comparison between $RV$ and $IV$ using HF 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,

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

(2.6)

where $\hat{\gamma}_j \equiv \sum_{i=1}^{N} r_{i+j}^2$, $\hat{\gamma}_0 \equiv r_t^2$, and $k \geq 2$ is the order of autocorrelation. A progressive increase in $k$ will change (2.6) as the amount of noise filtered increases. Hansen and Lunde (2006) choose the value of $k$ that renders (2.6) 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 (2.6) for each day using intraday tick data.

To draw conclusions from this approach, we then visually compare, using variance signature plots, an average of $IV$, $RV_{ACNW}^{1tick}$, to an average of the daily observed price $RV$ (denoted by $RV_t^m$ and measured using calendar-time sampling to approximate the duration of noise). $RV_t^m$ is estimated for progressively longer intraday time intervals $m$. The difference between the $RV_t^m$ and the $RV_{ACNW}^{1tick}$ represents the $NB$ for interval $m$. By increasing the length of $m$, we can observe the length of time needed for $NB$ to disappear.
2.2.2. 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,

$$ p_{t_i} = \begin{pmatrix} \text{transaction price at time } t_i \\ \text{corresponding ask price at } t_i \\ \text{corresponding bid price at } t_i \end{pmatrix}, \quad (2.7) $$

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

$$ \Delta p_{t_i} = \alpha \beta' p_{t_{i-1}} + \sum_{j=1}^{l-1} \Gamma_{1j} \Delta p_{t_{i-j}} + \sum_{j=0}^{l-1} \Gamma_{2j} vol_{t_{i-j}} + \mu + \epsilon_{t_i}, \quad (2.8) $$

where $\mu = \alpha \rho$ 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 and $\alpha$ is a 3x2 matrix (Hansen and Lunde 2006). Following Hasbrouck (1991), our VECM specification includes trade volumes ($vol_{t_{i-j}}$) 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 (2.8). The error, $\epsilon_t$ ($i = 0, 1, \ldots, l$), is assumed to be an uncorrelated error vector, $\beta$ is a 3x2 matrix, $l$ is the number of lags, and $\Gamma_1$ and $\Gamma_2$ 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 $(\beta_1^t \mathbf{p}_t)$ represents the long-run link between the transaction price and the quotes, and the second $(\beta_2^t \mathbf{p}_t)$ is used to represent the long-run pattern of the bid-ask spread. As a result, $\beta$ can be expressed as

$$\beta = (\beta_1, \beta_2) = \begin{pmatrix} 1 & 0 \\ -1/2 & 1 \\ -1/2 & -1 \end{pmatrix}. \quad (2.9)$$

To identify $\alpha$ and $\beta$, the following normalization vectors are imposed, $\beta_\perp = (1 1 1)'$ and $\alpha_\perp = (1 1 1)'=1$, where $\beta_\perp$ and $\alpha_\perp$ are 3 x 1 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

$$\hat{p}_t^* = (\hat{\alpha}_\perp \Gamma_1 \beta_\perp)^{-1} \sum_{j=1}^{l} \hat{\alpha}_j^t \hat{\epsilon}_t^j. \quad (2.10)$$

The corresponding efficient price intraday return is given by,

---

2 We test for the restrictions imposed on the cointegrating vectors and find them to hold on nearly 90% of the days in the sample.

3 There are alternative definitions of the common stochastic trend. Gonzalo and Granger (1995) derive it through a linear combination of observed prices. The alternative by Hasbrouck (1995; 2002) and Hansen and Lunde (2006) requires the common trend to be a martingale. This martingale property makes it the most appropriate definition of the efficient price.
\[ r^*_t \equiv \frac{\hat{a}_1 \hat{e}_t}{(\hat{a}_1, \hat{r}, \hat{\beta})}. \]  \hfill (2.11)

From (2.10), the \( RV \) of \( \hat{p}_t \) is,

\[ RV_p, = \sum_{l=1}^{I} (\hat{r}^*_t)^2 \]  \hfill (2.12)

which represents the semiparametric IV estimator. Under this semiparametric approach, derivation of a noise series is straightforward \( \ln(\hat{\nu}_t) = \ln(p_t) - \ln(\hat{p}_t) \), with \( \hat{e}_t = \ln(\hat{\nu}_{t,m}) - \ln(\hat{\nu}_{t-1,m}) \). From here, we can decompose \( RV^{(m)} \) into its three components thus identifying the two components of NB: \( 2 \sum_{l=1}^{I} e_{t_i} r^*_t \) and \( \sum_{l=1}^{I} (e_{t_i})^2 \).

2.3. Data and Empirical Findings

The research focuses on the period from January 2011 through December 2016 which is characterized by a predominance of electronic platform trading in live cattle futures markets. As noted, livestock futures trading volumes on the electronic platform represented about 80% in 2011 (Irwin and Sanders 2012) and reached 100% when the CME live cattle future pit closed in July 2015. Intraday price variance increased substantially in 2015 and 2016 (Figure 2.1). The Wilcoxon test for the intraday price RV before and after August 2015 rejects the null that the average RV did not change.

2.3.1. Data and Data Pre-processing

The data consist of transaction prices, quotes and trade volumes from CME Group’s BBO (Best-Bid-Offer) and market depth datasets for live cattle futures contracts traded on the electronic platform. We concentrate on the day-trading session which contains the bulk of market activity.

The CME live cattle futures contract is 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 rollover 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: (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 limit-price moves (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).\(^4\)

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), 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

\(^4\) 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 and were also deleted from the analyzed sample.
and December), nine LPM days in 2016 (three up and six down). Careful examination of the LPM revealed 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 fill 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 1/6 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). CTS is used in the derivation of observed price $RV$ ($RV^{(m)} = \sum_{i=1}^{m} r_{i,m}^2$). 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).

2.3.2. 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, $n$,

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

(2.13)

$RV$ 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 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 2.2 presents the variance signature plots. The horizontal line in each panel is the $IV$ for that period measured as $\overline{RV}_{ACNW_{30}}^{1 \text{tick}}$. 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 to four-minute sampling frequency, $RV$ estimates appear unbiased as they converge to the estimate of the $IV$.\footnote{Convergence is defined when the difference between the $RV$ of prices and the estimated $IV$ is less than 1%.

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 $IV$s. However, in both years, RVs converge to the $IV$s more quickly, reaching the $IV$ in only three minutes.

Table 2.1 presents selected details of the signature plots. For each period, the table provides estimated $IV$ (which corresponds to the horizontal line in Figure 2.2.), 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 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 2.1 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 Wilcoxon signed rank nonparametric test.

2.3.3. 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 (2.8) 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 (1000 iterations for each day) is used in the estimation process.

Table 2.2 presents average estimates from the cointegration model for the three periods identified earlier. Daily \( IV \) is derived from equation (2.12), as the square returns of the efficient price (Table 2.2, column 5). These semiparametric measures are close to the \( IV \) from the variance
signature plots, providing a robustness check. Table 2.2 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 $(\tilde{\alpha}^t_t, \tilde{\alpha}^b_b, \tilde{\alpha}^a_a) = (1/2, 1/4, 1/4)$, where $t$, $b$ and $a$ superindices represent the transaction, bid and ask prices, respectively, suggesting that transaction prices are more informative of the efficient price than quotes. This pattern becomes more accentuated in 2015, when a substantial increase in the transaction price’s coefficient takes place (Table 2.2, 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, i.e., RV is systematically above IV. However, these plots do not identify the sign of the correlation between the efficient price and noise. Using equation (2.5), 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 2.2). 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.

\(^6\) Live cattle futures prices declined from 170 to nearly 95 cents per lb in this period.
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 2.12) from the observed transaction log-prices (equation 2.1). 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,

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

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 \(EMs\) is developed (Figure 2.3). \(EM\) is small, on average, one cent per pound during 2011-2016. In 2011-2014, the median \(EM\) is 0.79% of transaction prices, and 0.67%, 0.86% in 2015 and 2016 respectively.\(^7\)

2.3.4. 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 September 2015-December 2016, which starts just before the dramatic drop in prices and increased variance.

\(^7\) When LPM days are included, there is no change at 2 decimal places.
Figure 2.4 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 USDA agricultural marketing service introduced online auction is also included (blue) (AMS-USDA 2016). Visual examination substantiates the statements that variance has increased. Recall from our earlier analysis, 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:

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

(2.15)

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. $NB_t$ 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 to also 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

---

8 Note these measures represent variance for a particular day in contrast to the variances in the signature plots which reflect averages for specific intervals over a year. As a result, they appear to be more volatile.
is transmitted across days. $VOL_t$ is the daily volume traded in the nearby contract. Volume, which is an indication of the information arrival, also can 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 2.4. 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 the addition of live cattle futures contracts to the Message Efficiency Program (MEP) on February 1, 2016. MEP is designed to reduce potentially harmful high frequency activity in the market. HF 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. $D_1_t$ 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. $D_2_t$ 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

---

9 More details on MEP and further discussion of its limited success can be found on their website (CME 2017).
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. \( D_3 \) equals one on August 5, 2016 and afterwards, and zero otherwise.

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); the large increase in slaughter numbers in mid-March reflecting cyclical supply changes (Hurt 2016) (\( D_4 \) 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, USDA agricultural marketing service began including online auction transactions from Superior Livestock Auction’s fed cattle exchange (\( D_5 \) 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 (2.15) 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 AR3). Table 2.3 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 (Cumby and Huizinga 1992; Pagan and Hall 1983) at the 5% level.

Based on statistical significance and size of the coefficients, the variables differ greatly in importance. A primary factor is the IV which confirms our earlier assessment that the NB and longer-term IV levels are positively correlated. It also suggests that relevant market price adjustments due to market participants incorporating information about fundamentals can lead to higher market noise as traders alter their positions to new market information. Informatively, once this is accounted for, increases in market liquidity (VOL) appear to marginally reduce market noise. It appears that added trading volume not only reduces bid-ask spreads (Frank and Garcia 2011), but also can reduce harmful NB. As expected, NB is not transmitted across trading days as the coefficient of NB_{t-1} does not differ from zero. This finding is consistent with the observation that HF traders do not usually hold overnight positions (Baron et al. 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 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 HF traders in the live cattle market.

Two factors which 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. Ex post this is rather clear, 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 which are somewhat in tandem with normal seasonal price patterns (Hurt 2015), providing another indication that noise accompanies the large swings in market prices.  

2.4. 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 (HF) traders who can operate on these platforms for the high intraday price variance observed in 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

---

10 The interested reader can find the program code files of this article in a supplementary appendix online.
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 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 HF traders do not hold overnight positions in the live cattle market. 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.
2.5. Figures and Tables

Figure 2.1. Realized variance on CME live cattle futures nearby transaction prices ($RV_i$), 2011-2016

Notes: $RV_i$ is divided by the corresponding last transaction price of the day session to standardize it for the level of prices that occurred in the period. Calendar time sampling (CTS) is used in the derivation of observed price realized variance, i.e. a normal day trading session contains i.e. 14,100 seconds.
Figure 2.2. Variance signature plots for the live cattle nearby futures contracts, 2011-2016

Notes: 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.
Figure 2.3. Daily median economic measure (EM) in live cattle nearby futures contracts, 2011-2016

Note: The observations are the daily median noise as a percent of transaction price.
Figure 2.4. Daily realized variance ($RV_t$), integrated variance ($IV_t$), and noise bias ($NB_t$), September, 2015-December, 2016

Notes: $RV_t$ 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
Figure 2.4. (Continued)

2016-10-05 when the USDA incorporated information from the first online auction implemented by the fed cattle exchange.
Table 2.1. Estimated $IV$, $RV$, and $NB$ at One-Second Frequency for the Live Cattle Nearby Futures Contracts, 2011-2016

<table>
<thead>
<tr>
<th>Year</th>
<th>$\hat{IV}$</th>
<th>$RV_t^{(1\text{ sec})}$</th>
<th>$RV_t^{(1\text{ sec})} - \hat{IV}$</th>
<th>$RV_t^{(1\text{ sec})} - \hat{IV}$</th>
<th>$\hat{IV}$</th>
<th>$RV_t^{(1\text{ sec})}$</th>
<th>$RV_t^{(1\text{ sec})} - \hat{IV}$</th>
<th>$RV_t^{(1\text{ sec})} - \hat{IV}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.45E-05</td>
<td>1.05E-04</td>
<td>1.58E-04</td>
<td>1.58E-04</td>
<td>2.99E-04</td>
<td>3.69E-04</td>
<td>6.24E-04</td>
<td>2.35E-04</td>
</tr>
<tr>
<td></td>
<td>(6.30E-05,6.66E-05)</td>
<td>(1.00E-04,1.05E-04)</td>
<td>(1.56E-04,1.60E-04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The $RV$, $IV$, and $NB$ are the realized and integrated variances, and noise bias of transaction prices. The confidence intervals for $IV$ are constructed using

$$CI(\bar{\sigma}^2) \equiv \exp\left(\log(\bar{\sigma}^2) \pm \delta \left(c_{1-\frac{\alpha}{2}}\right)\right)$$

where $c_{1-\frac{\alpha}{2}}$ are the 5 and 95 quantiles of the standard normal distribution, Hansen and Lunde (2006). Days with limit-price moves longer than 30 minutes were excluded in the left panel.
Table 2.2. Cointegration Results for the Nearby Contracts by Period, 2011-2016

<table>
<thead>
<tr>
<th>Year</th>
<th>$\alpha_t^t$</th>
<th>$\alpha_t^b$</th>
<th>$\alpha_t^a$</th>
<th>RV$_{p^t}$</th>
<th>$\sum e_t r_t^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011-2014</td>
<td>0.53</td>
<td>0.24</td>
<td>0.23</td>
<td>8.21E-05</td>
<td>-4.77E-05</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>0.63</td>
<td>0.19</td>
<td>0.18</td>
<td>1.32E-04</td>
<td>-6.12E-05</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.07)</td>
<td>(0.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>0.56</td>
<td>0.22</td>
<td>0.22</td>
<td>1.93E-04</td>
<td>-1.04E-04</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table presents the daily sample average from 2011 to 2016 of $\alpha_t$, RV of $\tilde{p}_t^t$ -- the measure of IV using the kernel estimator from Table 2.1, and the covariance between noise and returns. Super indices $t$, $b$ and $a$ on the alphas represent transaction price, and bid and ask quotes, respectively. The cointegration specification is estimated each day. The parameters presented are averaged over days. The numbers in parentheses for the orthogonal alphas are standard deviations while the confidence intervals are reported for $IV_t$. The last column measures the covariance between noise returns and efficient price returns. The limit-price move days were excluded.
Table 2.3. 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></td>
<td>(5.52e-07)</td>
</tr>
<tr>
<td>$NB_{t-1}$</td>
<td>0.14</td>
</tr>
<tr>
<td><em>(lagged noise variance)</em></td>
<td>(0.09)</td>
</tr>
<tr>
<td>$IV_t$</td>
<td>1.24***</td>
</tr>
<tr>
<td><em>(integrated variance)</em></td>
<td>(0.26)</td>
</tr>
<tr>
<td>$VOL_t$</td>
<td>-7.64e-11*</td>
</tr>
<tr>
<td><em>(Daily Trading Volume)</em></td>
<td>(3.95e-11)</td>
</tr>
<tr>
<td>$LPM_t$</td>
<td>-1.00e-06***</td>
</tr>
<tr>
<td><em>(limit price moves)</em></td>
<td>(2.62e-07)</td>
</tr>
<tr>
<td>$D1_t$</td>
<td>1.75e-07</td>
</tr>
<tr>
<td><em>(MEP)</em></td>
<td>(1.10e-07)</td>
</tr>
<tr>
<td>$D2_t$</td>
<td>2.71e-07***</td>
</tr>
<tr>
<td><em>(trading hours change)</em></td>
<td>(9.01e-08)</td>
</tr>
<tr>
<td>$D3_t$</td>
<td>8.53e-07***</td>
</tr>
<tr>
<td><em>(futures contract specifications changes)</em></td>
<td>(2.44e-07)</td>
</tr>
<tr>
<td>$D4_t$</td>
<td>6.87e-07***</td>
</tr>
<tr>
<td><em>(supply shock)</em></td>
<td>(2.29e-07)</td>
</tr>
<tr>
<td>$D5_t$</td>
<td>-1.74e-07</td>
</tr>
<tr>
<td><em>(online auction)</em></td>
<td>(2.81e-07)</td>
</tr>
<tr>
<td>$SPR_t$</td>
<td>4.44e-07**</td>
</tr>
<tr>
<td></td>
<td>(1.79e-07)</td>
</tr>
<tr>
<td>$SUM_t$</td>
<td>1.95e-08</td>
</tr>
<tr>
<td></td>
<td>(1.11e-07)</td>
</tr>
<tr>
<td>$AUT_t$</td>
<td>4.92e-07***</td>
</tr>
<tr>
<td></td>
<td>(1.60e-07)</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. Significance level at 1%, 5%, and 10% is represented by ***, **, and *, respectively. We use two instrumental variables that are (1) fitted values using an AR(5) for $IV_t$, and (2) 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 $D_i$ ($i = 1, \ldots, 5$) are the dummy variables representing limit-price moves and days with policy or supply shocks, respectively. $SPR_t$, $SUM_t$, $AUT_t$ denote spring, summer and autumn, respectively. These are described in the text in further detail.
2.6. References


CHAPTER 3
ARE CORN FUTURES PRICES GETTING “JUMPY”? 

3.1. Introduction

In efficient markets price volatility arises as participants, by buying and selling, incorporate their new information and expectations on fundamentals into the market (Hwang and Satchell 2000; Fama 1970). In agricultural commodity markets, researchers have established that United States Department of Agriculture (USDA) reports are an important source of fundamental information, as unanticipated information in the reports has been shown to affect price volatility in a substantial manner. Traditionally, reports were released after trading hours, but in a recent policy change USDA now releases them when markets are open, allowing for real-time trading on new information. Commercial traders have complained that the new policy leaves no time to digest information released, and creates an unfair playing field that favors nonconventional high-frequency trading firms that can quote, revise, and execute orders in milliseconds. With lower monitoring costs, these traders are able to continuously assess market conditions and take positions based on the most current information available. These concerns are consistent with Budish, Cramton, and Shim’s (2015) assessment that high-speed market activity combined with the release of public information, can lead to sharp market price movements as high frequency traders’ race to take advantage in multiple ways of an environment created by the new information.

Adjemian and Irwin (2018) show that real-time trading on USDA crop announcements leads to large volatility spikes in grain futures prices that dissipate within a few minutes. But price volatility has two components, a jump-free quadratic variation of a continuous price path (bi-power variation) and a quadratic variation of an instantaneous and discrete price move (jump variation). Previous research has identified the importance of disentangling these components, since bipower
variation is easier to predict and to hedge, in contrast to the jump component that cannot be predicted and is difficult to hedge (Aït-Sahalia 2004; Bollerslev, Law and Tauchen 2008; Todorov and Tauchen 2011). Andersen et al. (2007) have separated jumps from the continuous sample path variation of returns and show that almost all the predictability in return volatility comes from the non-jump component. While the magnitude of jump variation is difficult to predict, the timing of jumps at least in financial and currency markets appears somewhat more systematic. Early studies by Goodhart et al. (1993) and Almeida, Goodhart and Payne (1998) report price jumps following news releases. More recently, Andersen et al. (2003) confirm the link between news releases and jumps in exchange rates and show that announcement surprises (i.e., the deviation between market expectations and macroeconomic announcements) influence the jumps. These findings suggest that price discontinuities—jumps—may be the source of the volatility spikes identified by Adjemian and Irwin (2018). To the extent that jumps are present during the USDA announcement days, price discovery is less predictable and hedging is more problematic. Additionally, jumps lead to a substantial increase in execution risk that is relevant to all traders willing to take or liquidate positions. These traders span from non-commercials operating at high or low speeds to commercial traders exercising their price judgement to hedge and participate in the pricing process, and even to producers who simply are liquidating their market positions.

Without much empirical support, recent jumps in agricultural markets have been attributed to public information shocks (Dreibus and Sparshott 2014), and to the reduced trading latency brought by the adoption of new trading technology by exchanges and traders (Stebbins 2013; Miller and Shorter 2016). High-frequency trading (HFT) entered agricultural commodity futures markets in 2006 after the Chicago Board of Trade (CBOT) launched its electronic trading platform that increased access and the speed of operations (Irwin and Sanders 2012). Increased speed not
only brought new types of traders to the market, but also changed the way conventional traders operate as they adapt to the new low latency environment (O’Hara 2015). After the adoption of the new platform, the Commodity Futures Trading Commission (CFTC) identified a number of large price movements or “flash” events in the corn futures market between 2010 and 2015, which raised concerns about the price risk faced by market participants (CFTC 2015). Some traditional commodity investors have already announced company closures due to their inability to react quickly and effectively to increased price risk (Onstad 2018; Meyer 2018).

Disentangling jumps from volatility is key for managing the risk of positions taken on USDA announcement days. Price jumps are not felt equally by all traders. Market participants who watch markets at periodic but infrequent intervals could perceive price changes in between observation points as price jumps, even when the underlying process is jump free. This may occur when prices in between observation points are continuous, but highly volatile (volatility bursts). As a result, identification of genuine discrete and instantaneous price movements requires fine resolution (intraday) price data. Fine resolution data allow differentiating between a jump and a volatility burst and thus avoid spurious jump identification (Christensen, Oomen and Podolskij 2014). While allowing us to disentangle jumps from volatility bursts, high frequency intraday price data are characterized by microstructure noise that induces autocorrelation and discreteness in returns (Hansen and Lunde 2006). Huang and Tauchen (2005) show that noise can bias jump-test results towards finding more jumps. This requires careful selection of research methods.

The primary purpose of the research is to study the impacts of USDA announcements on the jump component of price volatility in the corn futures market. We also identify price jumps on non-announcement days. This is done for comparative purposes and to identify any systematic patterns in timing and magnitude of these jumps in the market, which should provide additional
insights into the execution risk that market participants face. We use nearby corn futures transaction prices, tick data time-stamped to the nearest second and observed from January 2008 to December 2015. We rely on the methods by Lee and Mykland (2008; 2012) and Christensen et al. (2014) who propose nonparametric approaches to detect intraday jumps, estimate jump risk and identify microstructure noise.

The main findings show that the transition to the real-time release policy has increased the number of jumps on announcement days, and increased jump clustering which has increased the cumulative jump size. Increased trading on announcements has also increased liquidity costs. Non-announcement jumps also have increased, but their size and liquidity costs are much smaller. We also show that traders operating at high frequency face more jump risk than traders operating at smaller frequency, which is due to the noise occurring during jumps.

The article makes several contributions. First, we identify the timing when jumps occur, their economic magnitude and the impact of the announcement surprise on the magnitude of the price jumps. We also characterize market conditions in the presence of jumps using the bid-ask spread, the trading volume around jump times and identifying any major crop report releases taking place in jump days. Our results add to the findings of Adjemian and Irwin (2018) by showing that prices jumps are the process by which markets react during announcements, which helps explain the recent volatility spikes identified. Second, we provide empirical evidence of the magnitude of intraday jump risk faced by traders operating at different temporal frequencies in agricultural futures markets. We call this the “sampling frequency of jumps”. Understanding the sampling frequency of jumps is important since automated trading in agricultural commodity markets has polarized trading latency, allowing nonconventional market participants to take and cancel positions at ultra-high speed, while hedgers continue to operate at a much slower pace. Third, since
we work with high-frequency price observations affected by microstructure noise, we disentangle the portion of daily price variance due to jumps in the efficient price from that due to market microstructure noise blurring efficient price jumps. We thus shed light on jump risk composition and allow for a better understanding of volatility and its dynamics.

3.2. Literature review

A few studies discuss agricultural commodity price jump risk in the context of futures and options prices modelling, and reach the conclusion that proper modeling of jumps can reduce forecasting errors. Hilliard and Reis (1999) show that large price changes cause return non-normality in commodity markets and propose a jump-diffusion model to better capture futures prices behavior. Koekebakker and Lien (2004) estimate jumps size and intensity for wheat options prices, assuming futures prices follow a jump-diffusion process. They develop a futures option pricing model and find that accounting for jumps reduces forecasting errors. Schmitz et al. (2014) model large price movements in U.S. corn, soybean and wheat spot prices using a Poisson jump-diffusion process with stochastic volatility. They find jump parameters to be significant and forecasting errors lower than errors from a stochastic model without jumps. While these studies point toward the relevance of accounting for jumps in modelling daily agricultural prices, they don’t examine their intraday presence and behavior.

Recent studies have also examined jumps in commodity futures markets using daily prices. Dempster, Medova and Tang (2018) use daily copper and oil futures data and differentiate between short term and long-term jumps. They find that futures contracts with shorter maturity usually exhibit much larger price jumps than those with longer maturity, which they argue implies the existence of jumps in the convenience yields. Nguyen and Prokopczuk (2019) focus on the jumps and co-jumps in 29 commodity markets and use daily data. The authors employ a nonparametric
test by Barndorff-Nielsen and Shephard (2006) which provides a measure of the jump variation component over a month. They find that commodity markets have less jump variance than the stock market. The test is repeated for each month and applied using 21 observations (number of days in a month). Diewald, Prokopczuk and Wese Simen (2015) study whether jumps in commodity markets are equally distributed in time. They identify extreme daily returns in the heating oil, natural gas, corn and soybeans futures markets by an \textit{ad hoc} rule that consists in taking the top and bottom $2.5\%$ returns. The authors assess the time-varying probability of a jump from January 2, 1991 to December 30, 2011. Their modeling approach is parametric, and their findings suggest that the equal jump intensity assumption over time is not supported. Instead, jumps are characterized by seasonal behavior (peak during cold months for energy markets and summer months for grain markets).

With increased trading speed, jumps occur and fade quickly. The literature studying the presence of jumps using intraday data often focuses on financial markets and is primarily concerned about the relative contribution of jumps to total price variance (Huang and Tauchen 2005; Andersen, Bollerslev, and Diebold 2007; Andersen, Benzoni, and Lund 2002; Tauchen and Zhou 2011). Along these lines, Wu et al. (2015) examine jumps in agricultural futures markets using a model-free approach and 5-minute sampled returns. They identify jumps in corn futures transactions prices by taking the difference between the annualized standard deviation of realized variance and the bipower variation. Similar to Wu et al. (2015), most literature assessing intraday jumps, uses a 5-minute sampling frequency to eliminate microstructure noise, which can confound jump identification. Christensen et al. (2014) suggest that jump occurrence is quite small ($1\%$ of the realized annualized variance) at the millisecond environment. They explain that the use of noise-filtered millisecond data reduces the likelihood of confounding volatility bursts with real
jumps. They also show that price-jumps have a higher impact at lower sampling frequency (e.g. 5-minute or 15-minute) than at ultra-high frequency.

The factors influencing intraday jumps and the market characteristics during jumps have been investigated in financial markets. Several studies explore the impact of news on intraday price jumps (Boudt and Petitjean 2014; Bjursell, Gentle and Wang 2015; Chan and Gray 2017; Jiang, Lo and Verdelhan 2011). Boudt and Petitjean (2014) distinguish between jumps related to firm news and to macro-announcements, and explore how jumps are linked to market liquidity measures such as bid-ask spreads. Christensen et al. (2014) argue that liquidity measures appear to have more significant jumps than prices during extreme market events (e.g. flash crashes or earthquakes) at the millisecond lens. Both studies find that market liquidity measures worsen following a jump. Brogaard et al. (2018) investigate whether extreme price movements are caused by high frequency traders. By using two main methodological approaches, one that is indifferent and the other that controls for time-varying volatility, they conclude that high frequency traders do not cause extreme market price movements, but instead act as liquidity suppliers during extreme price events.

3.3. Jump identification methods

With the arrival of high frequency trading and data, a variety of nonparametric tests have been developed to detect the jumps component in price variation (Barndorff-Nielsen and Shephard 2006; Aït-Sahalia and Jacod 2009; Lee and Mykland 2008). Dumitru and Urga (2012) use Monte Carlo simulation to compare alternative nonparametric jump testing procedures and conclude that the approach by Lee and Mykland (2008) is the most effective, which nonetheless might be oversized under extremely volatile processes. Lee and Mykland's (2008) test allows to time-stamp jump occurrence, but is not robust to the presence of microstructure noise and thus must be applied
to non-overlapping noise-filtered data (see appendix 3.8.1.1.). Microstructure noise is induced by trade frictions such as tick size, the bid-ask spread, or discretely sampled data, which create discreteness in the recorded price data that may be confounded with changes in the underlying efficient price.

Few studies have explicitly considered the effect of market microstructure noise on jump identification. Huang and Tauchen (2005) and Andersen et al. (2007) examine the effect of noise on jump detection assuming identically and independently distributed (i.i.d) noise, while Lee and Mykland (2012) and Christensen et al. (2014) adopt a more realistic assumption of non-i.i.d. noise, which is consistent with Hansen and Lunde (2006) assumptions and findings. Lee and Mykland's (2012) procedure assesses the intensity of jumps in intraday time intervals, relative to the total daily price variation in the presence of noise. However, their procedure does not time-stamp jumps. Christensen et al. (2014) estimate the jump variation ($JV$) component of total price variation by relying on noise-filtered price realized volatility ($RV$) and bipower variation ($BV$) estimators and identify the jump location using Lee and Mykland's (2008) test on 5-minute sampled returns and tick-sampled returns filtered for microstructure noise. We follow their approach.

We first identify and time-stamp intraday jumps using Lee and Mykland's (2008) jump identification test applied on noise-filtered tick price data (Christensen et al. 2014). We compare jump characteristics occurring during major USDA crop reports release days (Table 3.1) with jumps on non-announcement days and we characterize market liquidity around jump time. Second, for the days with jumps, we estimate the contribution of jump risk to total price risk by computing the ratio of $JV$ relative to $RV$. Intraday jumps may not affect all traders in the same fashion. Market participants who watch markets at periodic but infrequent intervals could perceive price changes in between observation points as price jumps, even when the underlying process is continuous. In
contrast, high frequency traders operating at high resolution are likely to face less efficient price risk. To measure these differences, we draw the JV signature plots at different sampling frequencies (e.g. 1-tick, 5 minutes, or 10 minutes). These signature plots are also used to show the relevance of market microstructure noise during jump occurrence. In the following sections, we offer a detailed description of the Lee and Mykland (2008) jump test and the process to derive the jump variation component and related signature plots.

3.3.1. Jump detection and location

Intuitively, the Lee and Mykland (2008) test identifies intraday price jumps by observing the sequence of intraday returns within the day. The test compares the value of returns relative to a volatility measure calculated over the immediately preceding returns that is impervious to jumps. More specifically, returns are divided by the integrated volatility (JV) estimated using the jump-robust BV in a time window whose size is controlled by parameter W (the number of previous returns considered to compute the preceding volatility).\(^\text{11}\) To establish statistical significance, Lee and Mykland (2008) study the distribution of the maximums of this test under the null hypothesis of no jumps, which allows us to define the critical values for the test rejection and thus identify the jumps.

More formally, consider the log efficient transaction price, which is free from microstructure noise and follows a martingale. It is represented by \(P^*(t)\) and modeled as

\[
dP^*(t) = \mu(t)dt + \sigma(t)dZ(t) + Y(t)dJ(t),
\]

where \(t \in [0, T]\) indexes time, \(Z(t)\) is an \(F_t\)-adapted standard Brownian motion, with \(F_t\) being a right-continuous information filtration for market participants. \(\mu(t)\) is a drift and \(\sigma(t)\) is a

\(^{11}\) When working with tick data, an approximate link between \(W\) and trading time can be established, by computing the average time needed for a transaction to occur.
stochastic volatility process, both \( F_t \)-adapted processes with an underlying Ito process with continuous sample paths. \( Y(t) \) is the predictable jump size, with a mean \( \mu_y(t) \) and a standard deviation \( \sigma_y(t) \). \( J(t) \) is assumed to follow a non-homogenous Poisson distribution (i.e., jumps arrival time is independently distributed but, for instance, they can arrive more frequently at a certain time of the day).

Market microstructure noise contaminates prices observed at high frequency and thus we do not observe \( P^*(t) \). To clean observed prices of noise, we use the method by Lee and Mykland (2012) described in appendix 3.8.1.1., which filters for serial correlation using resampling techniques, and filters for any remaining noise by pre-averaging the resampled prices over non-overlapping windows. After filtering, we obtain an estimate of the efficient market price (\( \hat{P} \)) on which we apply the jump detection test.

We take a fixed time horizon (the day trading session) \( T \) with \( N \) observations and observe the sequence of intraday returns. To test if the return \( \hat{r}_{ti} = \hat{P}_{ti} - \hat{P}_{ti-1} \) contains a jump, we compare its magnitude against \( \hat{\sigma}(t_i) \), the realized \( BV \) during the previous \( W \) returns. The jump detection test statistic is denoted by \( L(i) \) and defined as:

\[
L(i) = \frac{\hat{r}_{ti} - \hat{r}_{ti-1}}{\hat{\sigma}(t_i)}. \tag{3.2}
\]

The \( BV \) estimator of \( \hat{\sigma}(t_i) \) in the window \( W \) is computed as follows,

\[
\hat{\sigma}(t_i) = \sqrt{\frac{1}{W-2} \sum_{j=i-W+2}^{i-1} \left| \hat{r}_{t_j} \right| \left| \hat{r}_{t_{j-1}} \right|}. \tag{3.3}
\]

The optimal window size \( W \) is chosen to ensure robustness of \( \hat{\sigma}(t_i) \) to jumps. Lehecka, Wang and Garcia (2014) show that USDA announcement effects, when announcements are released outside trading hours, are usually absorbed by the market in ten minutes. Adjemian and Irwin (2018) identify volatility spikes right after the release of the reports during trading hours, but spikes
dissipate in ten to fifteen minutes. We define $W$ so that it covers on average one hour and a half.\footnote{This corresponds to $W = 176$. In appendix 3.8.2, we provide the intraday distribution of jumps for $W = 185$ (which is approximately equivalent to using the data over the preceding 100 minutes), with the main results holding for the different window sizes. The robustness check conducted is limited by data availability within a day trading session.} This should lead $BV$ measures that are robust to jumps, as jumps are likely to occur in the first minutes after the release.

The test (3.2) is performed for each intraday return in a day. The performance of the jump test is measured by its ability to identify actual jumps and avoid type I statistical errors (i.e., reject the null when there is no jump). By performing the test repeatedly within each day, the number of jumps spuriously detected converges to the test significance level (Bajgrowicz et al. 2016). For example, if a jump test is conducted with a 5% significance level over 200 intraday returns, on average, 10 jumps will be erroneously identified. Some studies have adopted an ad hoc response to the spurious jump detection problem by using very conservative significance levels (0.1%), e.g. Bollerslev, Law and Tauchen (2008) and Giot, Laurent and Petitjean (2010). Lee and Mykland (2008) however, address the problem through the critical values of the maximum of the test statistic, which increase as the number of intraday tests performed increase (Dumitru and Urga 2012).\footnote{Bajgrowicz et al. (2016) have characterized Lee and Mykland approach (2008) as excessively conservative when applied to high frequency data.} Under the null of no jumps, the test statistic $\mathcal{L}(i)$ takes a small value and follows approximately a normal distribution. Lee and Mykland (2008) identify the null hypothesis’ rejection region by studying the asymptotic distribution of the maximums of the test statistic under the null of no jumps during the interval $(t_{i-1}, t_i]$. For this purpose, they show that $\mathcal{L}(i)$ sample
maximums converge to a Gumbel variable. The null of no jump in $\hat{r}_{t_i}$ will be rejected if $|L(i)| > G^{-1}(1 - \alpha)S_n + C_n$, where $G^{-1}(1 - \alpha)$ is the $(1 - \alpha)$ quantile function of the standard Gumbel distribution and $C_n$ and $S_n$ are defined as,

$$C_n = \frac{(2 \log N)^{1/2}}{c} - \frac{\log \pi + \log (\log N)}{2c (2 \log N)^{1/2}},$$

$$S_n = \frac{1}{c (2 \log N)^{1/2}},$$

where $c = \sqrt{\frac{2}{\pi}}$ and $N$ denotes the number of intraday tests or number of intraday observations.

With the probability $\alpha$ of type I error, we reject the null hypothesis of no jump if $|L(i)| > \beta^* S_n + C_n$ with $\beta^*$ defined such that $\exp(-\exp(\beta^*)) = 0.99$ for 1% significance level, implying $\beta^* \approx 4.6001$. When the Lee and Mykland (2008) test identifies a jump, we stamp it at $t_i$, corresponding to the timing of the last price observed in $W$. We define the jump size in cents/bushel as the difference between the noise-filtered price (not in logarithm form) between $t_{i-1}$ and $t_i$, i.e. $\exp(\hat{p}(t_i)) - \exp(\hat{p}(t_{i-1}))$.

### 3.3.2. Jump variation component

Price returns are characterized by two stochastic components (equation 3.1), a jump-free stochastic process and a jump stochastic process. While Lee and Mykland's (2008) test identifies the number of statistically significant jumps in a day, their timing and magnitude, it does not reflect their importance relative to the overall stochastic process, nor does it reflect who faces jump risk. Here, we offer a measure of the relative importance of jumps by decomposing the RV of returns$^{14}$ into the bipower and the jump variation components and express the latter as a share of the former.

---

$^{14}$ Which is an efficient estimator of the returns’ quadratic variation ($QV$) in the absence of microstructure noise (Andersen et al. 2001).
Jumps will not be equally felt by all traders and we study how traders checking the market at different time intervals (from every tick to every ten minutes) will perceive the jumps.

More specifically, for the jump days, we estimate the daily proportion of $JV$ relative to the noise-filtered price $RV$ estimate of Christensen et al. (2014):

$$RV^c = \frac{N}{N-K+2} \frac{1}{K \psi_K} \sum_{i=0}^{N-K+1} \left( \hat{r}_{t_{i+1}} \right)^2 - \frac{\hat{\omega}^2}{\theta^2 \psi_K}$$

(3.4)

where $N$ is the total number of intraday observations, $\hat{r}_{t_i}$ is the noise-filtered return, $\psi_K = \frac{(1+2K^{-2})}{12}$, with $K = \theta \sqrt{N} + o(N^{-1/4})$ and $\hat{\omega}$ is estimated using $\hat{\omega}_{AC} = \frac{1}{N-1} \sum_{i=2}^{N} \left| \hat{r}_i \right| \left| \hat{r}_{i-1} \right|$ (Oomen 2006). Noise under the Christensen et al. (2014) approach is filtered using overlapping pre-averaging techniques described in appendix 3.8.1.2. The $RV^c$ is the sum of the $JV^c$ and a jump-robust estimator of the integrated variance. The latter is approximated by the $BV^c$ (Barndorff-Nielsen and Shephard 2004) which we also define on filtered prices as follows,

$$BV^c = \frac{N}{N-2K+2} \frac{1}{K \psi_K} \frac{\pi}{2} \sum_{i=0}^{N-K+1} \left| \hat{r}_{t_i} \right| \left| \hat{r}_{t_{i+1},K} \right| - \frac{\hat{\omega}^2}{\theta^2 \psi_K}.$$  

(3.5)

A consistent estimator of the $JV$ component in presence of noise is thus given by,

$$JV^c = RV^c - BV^c \xrightarrow{P} \sum_{i=1}^{N_i} f_i^2.$$  

(3.6)

The magnitude of $JV^c$ expressed as a portion of total $RV^c$ is given by equation (3.7),

$$JV_{\text{share}}^c = \frac{RV^c - BV^c}{RV^c}.$$  

(3.7)

Annualized $JV^c$ (expressed as a proportion of $RV^c$) signature plots can be developed to identify the importance of $JV^c$ at different sampling frequencies and thus provide a measure of the jump risk faced by traders taking positions at different speeds.

Through these signature plots, we can also compare $JV^c$ to $JV$, the latter being based on observed, non-filtered prices, defined as $JV = RV - BV$, being $RV = \sum_{i=1}^{N} (r_{t_i})^2$ and $BV = \sum_{i=1}^{N} (r_{t_i})^2$.
\[ \frac{N}{N-1} \frac{\pi}{2} \sum_{i=2}^{N} |r_{t_i}| \left| r_{t_{i-1}} \right|. \] The difference between \( JV^c \) and \( JV \) shows the \( JV \) portion that can be attributed to noise. As the sampling frequency declines, noise dissipates and \( JV \) and \( J V^c \) converge. The duration of noise variation is approximated by the sampling frequency for which the two measures converge.

3.4. Empirical design and results

We use tick data for corn transaction prices from CME Group’s BBO (Best-Bid-Offer), time-stamped to the nearest second and traded on the electronic platform. The sequence of transactions occurring within a second is preserved using a sequence number. The sample period is from January 14, 2008 to December 4, 2015, resulting in 1983 trading days. The corn futures contracts are traded with five delivery months: March, May, July, September and December. We use the nearby series, defined as the nearest contract delivery month with the highest trading volume. The nearby is the most liquid contract that attracts most of the trading activity, and where high frequency traders are more likely to operate (Brogaard, Hendershott and Riordan 2014).\(^{15}\) We center our attention on the day trading hours,\(^{16}\) which present most of the trading activity. Days with limit-price moves (LPM) in which the prices stay locked most of the day trading session and

\(^{15}\) We also examine presence of price jumps in the first deferred contract in period 3. We find that the total number of jumps is 152 compared to 298 in the nearby, which was expected since most trading activity (pricing and hedging) occurs in the nearby contract, particularly for large algorithmic traders which might influence the presence of jumps (Budish, Cramton and Shim 2015; Brogaard et al. 2014). Investigation of the differences when markets were in backwardation showed little evidence that this was influencing this behavior.

\(^{16}\) Trading hours considered are: before May 21, 2012, from 9:30 to 13:15; May 21-December 31, 2012: 7:00 to 14:00; January 2 - April 5, 2013, from 9:30 to 14:00; since April 8, 2013, from 8:30 to 13:15, and since July 6, 2015, from 8:30 to 13:20. We use a wider time window for the period from May 21, 2012 to December 31, 2012, to observe market behavior around the report release at 7:30:00 CT.
do not offer enough data to apply the jump test are inadequate for the empirical analysis and they are thus excluded.\textsuperscript{17}

Figure 3.1 depicts, in the top panel, the daily nearby corn contract closing transaction prices, and in the bottom panel, the annualized realized volatility of noise-filtered (Lee and Mykland 2012) transaction prices for the sample period. We report the descriptive statistics of the transaction prices in table 3.2. Episodes of high intraday volatility are observed during 2008-2010 and after 2013. Since 2013, the corn futures price volatility is characterized by salient daily spikes corresponding, in most cases, to the monthly or quarterly USDA reports (Table 3.1). The change in volatility dynamics since 2013 suggests that the release of USDA reports when markets are open has changed the intraday price behavior appreciably.

We identify jumps in transactions prices and present our results for USDA announcement and non-announcement days separately. The announcement days correspond to monthly WASDE and quarterly Grain Stock (GS) report days, which have been identified by previous research (Adjemian and Irwin 2018) to have a major impact on corn futures prices (see Table 3.1 for details).\textsuperscript{18} Since January 2013, both reports are released at 11:00:00 am CT. Following Adjemian and Irwin (2018), we refer to this period as the “real-time” era, as opposed to the halt era before 2012, and call it period 3. From June to December 2012, the reports were also released in real time,

\textsuperscript{17} Only 5 LPM days (3 corresponding to announcement days) are excluded (all in the period from January 2008 to May 2012).

\textsuperscript{18} The monthly Cattle on Feed and the quarterly Hog and Pig reports are released at 2pm CT, after the day trading session closes. We test for the presence of jumps in price at the opening of the trading session on the day following these reports’ release and we do not identify any systematic jump process (only one jump on June 29\textsuperscript{th}, 2015 at 8:30:36 am is identified). As a result, we do not to include livestock reports in the definition of announcement days. Notice that this jump is however included in the group of non-announcement days.
but earlier at 7:30:00 am CT when trading volume is usually lower. We call this period (period 2) and allow for different jump behavior since market liquidity conditions may vary. From January 2008 to May 2012, the report release time was 07:30:00 when markets were closed, which we call period 1. The average daily trade volume over our sample period is 134,391 contracts during announcement days and 81,690 during non-announcement days.

3.4.1. Jump detection

Here, we present the results of the nonparametric test by Lee and Mykland (2008) used to identify jumps, their timing of occurrence and size. We filter observed prices for noise using the techniques explained in appendix 3.8.1.1. We conduct the test in (3.2) for each intraday efficient price return by comparing it to BV calculated over the preceding $W = 176$ noise-filtered returns which is approximately equivalent to using the data over the preceding 90 minutes.¹⁹ When moving to the next intraday return, the window $W$ is rolled one observation to the right.

Lee and Mykland (2008) intraday jump test identifies jumps when observed returns are large relative to the preceding volatility. The jump identification process is equivalent to a market participant walking through the trading session and considering that normal returns are the ones experienced during the preceding 90 minutes. Whenever a new return is high in absolute values, relative to the baseline of 90 minutes, the market participant will consider the return to be a jump. To avoid losing observations at the beginning of the trading day, the first rolling window ($W$) to

¹⁹ We assess the robustness of Lee and Mykland's (2008) test results by increasing the window size to $W = 185$. The results, presented in appendix 3.8.2, are similar to the ones using $W = 176$. 

56
conduct the test starts at 03:00:00 am. The low trading overnight and/or the morning halt before May 21st, 2012 requires going back to 03:00:00 am in order to have enough observations.

Table 3.3 presents a summary of the jump test results. We find 269 days with at least one jump, representing 14% of the total trading days, and a total of 446 jumps indicating the presence of multiple-jump days. Jump behavior differs by period and by type of day. As a percentage of the number of days in each period, jumps are more prevalent in period 3 (20.92% of the days) followed by period 2 (11.54%) and 1 (8.89%). In terms of the number of jumps on announcement days in each period, a clear increasing pattern can be seen. Only 11.48% of the announcement days in period 1, 55.56% of the announcement days in period 2, and 88.37% of the announcement days in period 3 experienced at least one jump. The average number of jumps per announcement day also increased from 0.18 in period 1, to 0.89 in period 2, and 3.02 in period 3. This last statistic reflects the jump clustering on announcement days that has occurred in the real-time era. During period 3, 79% of the announcement jump days had at least two jumps clustered within 2 minutes after the release of the report. informatively, while the cumulative absolute jump size is the largest (9.19 cents/bushel) in period 3,20 not all clustered jumps move in the same direction. In fact, many of the clustered jumps in period 3 for announcement and non-announcement days move in opposite directions, perhaps pointing to the uncertainty around the announcement time and trader activity.

---

20 We verify the daily cumulative absolute jump size (sum of jumps size in absolute value when jumps are clustered) on announcement days when WASDE and Grain Stock reports are released simultaneously. The average jump size on these days (in bold in Table 3.1) is twice the size on other WASDE-only days. To further understand this difference, we also compute the average size of jumps on Grain Stock reports without simultaneous release of WASDE report (June, March, and September). We find that the cumulative jumps size is 20% higher than the jump size on WASDE only days. Cumulative jump magnitude is thus higher on days with Grain Stocks report release compared to other announcement days.
Somewhat unexpectedly, jumps on non-announcement days are not trivial. There are 297 jumps occurring on 219 non-announcement days in the entire sample. Period 3 experiences the largest number of non-announcement jump days and jumps. Jump days as percent of non-announcement days in each period increased through time from 8.74% in period 1 to 16.74% in period 3. Notice the cumulative absolute jump size is much smaller on non-announcement days than announcement days (e.g., in period 3, 1.26 cents/bushel on non-announcement compared to 9.19 cents/bushel on announcement days). Jump size on non-announcement days has declined over time, suggesting that technological changes affecting agricultural commodity markets may have increased liquidity provision and reduced jump size outside announcement sessions.\(^\text{21}\) Despite their smaller size, similar to the announcement pattern, jumps on non-announcement days also cluster, particularly in period 3 where 23% of jump days have at least two jumps clustered and where nearly half of the jumps in clusters move in the opposite direction.

The findings suggest that jumps are more prevalent since the USDA has started releasing reports when the market is open. Jumps on announcement days have increased in size, but jumps on non-announcement days have declined in size. In the most recent period, jumps tend to cluster on both announcement and (non-announcement) events with 73% (48%) consequent jumps moving in opposite directions. Our percentage of jump days is above Bjursell et al. (2015) who, using 5-minute sampling returns, find energy price jumps to affect between 4% and 7% of the total trading days. They further find only a low jump rate (9%) associated to inventory announcements. Our percentage of jump days is however lower than Lahaye, Laurent and Neely (2011) who find 25% of the trading days with at least one jump in the foreign exchange market. They also find that

\(^{21}\) Larger jumps in non-announcement days in period 1 may have also been related to the financial crisis turmoil in 2008-2009 which would lead to a similar temporal pattern. We thank a reviewer for identifying this point.
jumps in foreign exchange markets, financial index futures and 30-year U.S. treasury bonds futures markets tend to cluster around public announcements time. Maheu and McCurdy (2004) also find that price jumps cluster when new information is incorporated into the market, reflecting the structure of the information arrival process.

To assess the location of the jumps within the day, we provide Figure 3.2 which is a histogram of jumps for intraday time intervals in the three periods. In period 1, jumps occurred slightly more often (38%) in the first interval of the day trading session, with a relatively even distribution throughout the rest of the trading day. The incorporation of information, both private and public, at the market opening is likely to create more frequent price jumps. In period 2, jumps occurred most often in the second and first intervals of the day, with 28 and 24% of the jumps, respectively, with the first interval coinciding with a lowly traded period when the USDA reports were released. Finally, in period 3, more than half of the intraday jumps (53%) are detected from 10:30:00 am to 11:29:59 am, which coincides with the USDA report release time discussed and the clustered jumps identified previously. Changes in USDA report release times have shifted the timing and the structure of the absorption of fundamental information in the corn futures market, increasing the proportion of price jumps occurring around the release time and making price jumps timing more predictable.

Figure 3.3 shows the average jump absolute size (in cents/bushel) for intraday intervals in the three periods. In periods 1 and 2, the largest jumps are located in the first interval of the trading day, reflecting USDA report release times and the liquidity in the interval. In period 2, low liquidity during the interval from 7:30 to 9:30 coupled with the release of USDA reports in the same interval explains the large jump size at the open. In period 3, the largest jumps are registered in the middle of the day trading session, corresponding to the USDA report release time. Notice that the average
jump size when the USDA release occurs in period 3 is roughly 56% larger than the first interval when the USDA release occurred when the market was closed. Consistent with earlier discussion, the non-announcement jumps in period 3 are small and below non-announcement jumps in the other periods.

3.4.2. Jumps and market liquidity

In figure 3.4, we examine market conditions during jumps by presenting the behavior of the bid-ask spread and trading volume in the minutes preceding and following jumps on announcement and non-announcement days. Specifically, we examine the maximum bid-ask spread and total volume within 1-minute bins for a 10-minute window before and after the jump. The panels are sorted by period (each pair of columns represents a period) and type of day (the first row (red plots) represents USDA announcement days and the second row (black plots) presents the non-announcement days). Liquidity measures are compared to the same measures on no-jump days in that period.22 Filled bullet points correspond to the cases when the Wilcoxon test’s null hypothesis of no significant difference in mean between two series is rejected at 1% significance level. In general, spreads and volumes have an n-shape around jumps, with the peak occurring right after the jump. Relative to no-jump days, spreads and volumes become significantly higher a few minutes before jump occurrence and usually do not return back to normal levels within the ten-minute interval considered. For jumps on non-announcement days, spreads reach about 0.6 and 0.5 cents/bushel in periods 1 and 2, respectively, but do not exceed 0.4 cents/bushel in period 3. The volume around jumps on non-announcement days has remained stable with a maximum of about 2,000 contracts/minute. For jumps on announcement days, suppression of the morning

---

22 Note that the counterfactual liquidity variables are measured using the maximum price change on days when no significant jump is detected.
trading halt has widened spreads around jumps, from a maximum of about 0.7 cents/bushel in period 1, 1.1 cents/bushel in period 2 and 1.2 in period 3. Bid-ask spreads for jumps on announcement days are substantially larger than spreads for jumps on non-announcement days.

Volume has also increased from a maximum of 2,000-2,500 contracts/minute in periods 1 and 2, to nearly 9,000 contracts in period 3.

In sum, price jumps are usually accompanied by an increase in trading volume that is higher for announcement than non-announcement days. While volume on non-announcement jump days has not changed substantially over the period studied, volume on announcement jump days has sharply increased since 2013. For the spread, while most jumps cause spreads to widen slightly (around 2 ticks on non-announcement days), real-time trading of USDA reports and sharp increases in trading resulted in spreads changing more than 5 ticks in period 3. Real-time trading of USDA reports appears to have resulted in a substantial increase in transactions costs which is consistent with Christensen et al (2014) and Boudt and Petitjean (2014).

### 3.4.3. Jump magnitude and inventory surprises

Here, we investigate the relationship between price jump magnitude and the surprises in USDA corn inventory releases using a regression framework. To measure the relationship between the size of jumps and the magnitude of the surprise, we define the absolute value of the surprise on the corn ending stocks from WASDE reports as $|S_t| = |\log(A_t) - \log(E_t)|$. $A_t$ is the value of corn ending stocks in the monthly WASDE report and $E_t$ is the expected value. Previous literature usually approximates $E_t$ by private analyst forecasts published by news services prior to the public announcements. We use the ending stocks from Bloomberg to represent private analysts’ forecasts (available starting in September 2009). The magnitude of jumps is $\sum |J_S_t|$, and computed as the sum of the absolute value of log-price jumps size within the day, which allows for the possibility
of more than one daily jump. A similar framework has been used by a number of researchers to identify the degree to which surprises in USDA information affects price changes. We expect the surprise effect to be linked to the magnitude of price jumps, with larger surprises leading to larger jumps.

The regression is specified as:

\[ \sum |S_t| = \alpha_0 + \alpha_1 |S_t| + \alpha_2 ANNxPOL_t + \alpha_3 SPR_t + \alpha_4 SUM_t + \alpha_5 AUT_t + \alpha_6 D_{time-to-rollover} + e_{1,t} \]

(3.8)

where \( e_{1,t} \) is error term, and \( ANNxPOL_t \) is an interaction between announcement days (one on announcement days, and zero otherwise) and report release policy (one since May 21st, 2012, and zero before) to allow for a differential effect with real-time announcements. \( D_{time-to-rollover} \) is dummy variable equal to one for 40 days before rollover and zero otherwise to allow for a possible Samuelson effect that jumps are larger as maturity approaches and trading increases.\(^{23}\) Seasonal dummies are included for Spring (\( SPR_t \)), Summer (\( SUM_t \)), and Autumn (\( AUT_t \)) to allow for the possibility that well-understood corn crop volatility patterns may be influencing price jumps (Egelkraut, Garcia and Sherrick 2007).

The estimation results are presented in table 3.4. The magnitude of the adjusted R squared is reasonably high, and the residuals show no signs of autocorrelation and heteroscedasticity at the 5% level of significance. The seasonal dummies and time to rollover do not have a statistically significant effect on price jumps magnitude. This is due to Lee and Mykland (2008) test robustness to heightened volatility during these periods, which is captured by dividing returns by the prior 90

\(^{23}\) Investigation of the number of jumps prior to expiration pointed to the possibility that larger jumps occurred 40 days immediately prior to the roll to the next contract.
minutes $BV$ (see denominator in equation 3.2). This makes it harder for large absolute value price returns during high volatility periods to be considered as jumps.

As expected, the magnitude of the surprises is positively and statistically significantly related to the cumulative absolute jump size, which indicates that the larger the surprise, the larger is the jump magnitude. The coefficient for $ANNxPOL_t$ further indicates that the real-time trading of USDA reports is more likely to increase jump size, which is essentially due to the jump clustering in the real-time era. Overall, the notion that jump behavior is influenced by surprises is consistent with the studies that have explored the impact of various forms of news on intraday price jumps in financial markets (Boudt and Petitjean, 2014; Bjursell, Gentle and Wang, 2015; Chan and Gray, 2017; Jiang, Lo and Verdelhan, 2011).

3.4.4. Jump variation component

Following Christensen et al. (2014), the estimator of the daily $JV$ component of the total price daily $RV$ is defined as the difference between $RV$ and $BV$ and calculated using prices for days with at least one statistically significant jump identified. We provide relative measures of the $JV$ by expressing it as a share of the $RV$. The shares are calculated for the observed and noise-filtered prices. Specifically, $AJV^{share}$ is the daily annualized $JV$ as a proportion of daily annualized $RV$ calculated using observed transaction prices. $AJV^{c,share}$ is calculated using the noise-filtered transaction prices (see appendix 3.8.1.2.). Jump shares measured on the noise-filtered prices correspond closely to the prices used earlier in the analysis, and represent the jump portion of efficient price volatility. Jumps shares measured on actual transaction prices represent the execution risk since market participants do not trade on efficient prices when establishing or closing a position.
The jump shares are presented at different sampling frequencies for each period in Table 3.5. Figure 3.5 provides a visual representation of the two measures for period 3. Recall the measures reflect the portion of the price volatility attributed to jumps at specific sampling frequencies. The difference between the two shares in any period is a reflection of the market-friction noise that arises with the jump.

Several patterns are readily apparent. First, in almost all cases, the price volatility attributed to jumps as measured by both shares increases from period 1 to period 3. This finding is consistent with earlier results that highlighted the increasing prevalence of jumps and their magnitudes. The results are also compatible with increased jump risk due to real-time trading of announcements, which is in line with previous literature (Janzen and Adjemian 2017; Adjemian and Irwin 2018; Bunek and Janzen 2015). Second, the largest differences between the jump shares measures occur at the shortest sampling frequency (1-tick) where noise plays its largest role in the pricing process.

Third, the difference between the two measures declines and finally disappears at longer sampling frequencies as the importance of noise dissipates. The equalization of the jump shares occurs more rapidly on announcement (e.g., 10 seconds in period 3) rather than non-announcement days (e.g., 2 minutes in period 3) as more volume and liquidity are drawn to the market in response to the reports. Hence, price discovery seems to occur relatively faster on announcement days than on non-announcement days, which may reflect increased presence of high-frequency liquidity providers during these days. This interpretation is in line with Brogaard, Hendershott, and Riordan (2014) who find that during stressful periods, high frequency traders tend to supply liquidity and trade in opposite direction to price discovery errors.

Finally, to put the jump volatility in the corn market in a larger context, consider the noise-filtered jump share at 1-tick horizon in period 3, which can arguably be considered as reflective of
major agricultural markets in the new electronic-trading era. Combining information from table 3.5, we calculate a period 3 weighted average jump share using the proportion of days as weights. Based on 38 USDA announcement days (5.2% of trading days in period 3) with jump shares of 7.51%, 116 non-USDA days (15.8%) with a jump share of 6.10%, and the remaining 79% of the days with no jump risk, we find a 1.36% jump share for all trading days in the period. This statistic is very close to the 1.3% reported by Christensen et al. (2014) for DJIA constituent markets, but above 0.4% and – 1.3% for foreign exchanges and equity index markets. This finding also contrasts with Nguyen and Prokopczuk (2019) that finds that commodity markets have less jump variance than the stock market. These differences may reflect differences in time periods of analyses, procedures, and data used. But the notion that prices are spiky in agricultural commodity markets which are affected by stochastic factors is well established.

What are the implications of jump shares for market participants that trade at different time frequencies? Using the jump shares calculated with actual transaction prices, it is clear that the largest jump risk or execution risk is experienced by high frequency traders. These traders seek high returns and bear a higher risk. Market participants that trade at lower frequencies also face execution risk due to jumps, but it can be smaller. For instance, market participants trading at the 10-minute frequency face sometimes as little as 20% of the jump risk encountered by high frequency traders.

3.5. Conclusions

USDA report release policy during normal market trading hours has raised a series of concerns. Commercial traders have complained that the new USDA report release policy, which leaves no time to digest the information released, coupled with the recent technological changes in futures markets, essentially favor high-speed traders. In this article, we study the impacts of USDA
announcements on the jump component of price volatility in the corn futures market. Jumps are unpredictable, instantaneous and discrete price moves that have relevant implications for managing risk of positions taken during announcement days. We also identify price jumps in non-announcement days for comparative purposes and to identify any systematic patterns in timing and magnitude of these jumps in the market, which should provide additional insights into the execution risk that market participants face. Using high frequency data from 2008 to 2015 and nonparametric methods, we identify price jumps, their magnitude and timing. We also examine market conditions around price jumps, the factors underlying jump size, the magnitude of intraday jump risk faced by traders operating at different temporal frequencies and the microstructure noise affecting price jumps.

We find that 14% of the days in the sample contain jumps. Recent years have seen an increased presence of jumps on USDA report release days that appear to be driven by changes in the release policy to allow for real-time trading. Intraday timing of jumps and their magnitude, and regression results identifying a large differential increase in the cumulative jump size during the real-time trading era are also highly consistent with changes in USDA release policy. Jumps on announcement days have increased in number and cumulative size, but jumps on non-announcement days have increased in number and declined in size. Recent jumps on announcement days have given rise to higher liquidity costs, but liquidity costs around jumps on non-announcement days have not increased measurably. These differences jump size and market liquidity suggest that recent technological changes affecting trading in agricultural commodity markets may have increased liquidity provision to non-announcement jumps. Informatively, in recent periods, jumps cluster and frequently move in opposite directions. Clustering has been observed in other markets, but clustered jumps in opposite directions have not been reported. Since
the recent overall uncertainty in the corn market has not been extremely high, this finding may point to a race by high frequency traders to establish more profitable positions. Overall, it appears that higher frequency traders have been active in liquidity provision and in searching for profitable opportunities based on the speed of their trading operations.

The new report policy has increased execution risk, particularly around announcement times, which can limit hedging activities and affect the sustainability of commercial traders. While the efficient price jump risk is lower for high frequency traders than for conventional market participants, they face more execution risk due to heightened microstructure noise during jumps. These traders seek high returns and bear a higher risk. While the presence of jumps per non-announcement days has more than doubled in recent times, their magnitude has decreased by almost half and they are accompanied by lower transaction costs in the most recent period. Hence, outside USDA announcements, execution risk is larger, but the economic implications are smaller than previously experienced.

Overall, our results complement the findings by Adjemian and Irwin (2018) that real time trading of USDA announcements leads to volatility spikes, by showing that price jumps are the process by which markets react during announcements. However, our analysis also demonstrates the real, and sometimes large, difference that exists between efficient price volatility and realized price volatility. Finally, it is important to note that the corn market is relatively more susceptible to jump events in efficient price volatility when compared to other volatile non-commodity markets (Christensen et al. 2014). Well-understood stochastic factors and perhaps lower liquidity than in other financial and foreign exchange markets may help explain the higher levels of jump volatility and non-trivial non-announcement jumps identified here.
3.6. Figures and Tables

Figure 3.1. Daily nearby corn futures transaction prices (closing price) in logarithm form (top panel) and annualized realized volatility of noise-filtered transaction prices (bottom panel), January 15, 2008 – December 4, 2015

Notes: Noise-filtered transaction prices are obtained using Lee and Mykland's (2012) approach (see appendix 3.8.1.1.).
Figure 3.2. Distribution of intraday jumps across intraday time intervals in each period, January 15, 2008 – December 4, 2015.
Figure 3.3. Average jump sizes (absolute value) in cents/bushel per intraday time intervals in each period, January 15, 2008 – December 4, 2015.
Figure 3.4. Minute-by-minute maximum bid-ask spread (BAS) and trading volume using a window of 10-minute before and after the jump (dashed grey line) on announcement and non-announcement days (rows) by period (columns).

Notes: The first two panels on each row represent period 1, the following two panels represent period 2, and the last two panels represent period 3. Tests are performed to assess differences compared to non-jump day measures. The filled dark bullet points refer to the null of the two-way Wilcoxon test of no significant difference in mean between the two series being rejected at 1% significance level.
Figure 3.5. Annualized jump volatility shares (%) in period 3.

Notes: “+” curve corresponds to $AJV_{s\,h\,a\,r\,e} = \frac{ARV - ABV}{ARV}$ at different sampling frequencies, where $AJV_{s\,h\,a\,r\,e}$ is the proportion of daily jump risk including noise. $ARV$ is the annualized daily realized volatility ($\sqrt{252} \times RV$) and $ABV$ is the annualized daily bipower variation ($\sqrt{252} \times BV$). “o” curve corresponds to the $AJV_{c\,s\,h\,a\,r\,e} = \frac{ARV^c - ABV^c}{ARV^c}$ at different sampling frequencies, where $AJV_{c\,s\,h\,a\,r\,e}$ is the proportion of daily jump risk filtered for noise. $ARV^c$ is the noise-filtered daily annualized realized volatility ($\sqrt{252} \times RV^c$) and $ABV^c$ is the noise-filtered daily annualized bipower variation ($\sqrt{252} \times BV^c$).

<table>
<thead>
<tr>
<th>Year</th>
<th>World Agricultural Supply and Demand Estimates (WASDE) report</th>
<th>Grain Stock reports</th>
</tr>
</thead>
</table>

Notes: The WASDE report is monthly, while the Grain Stock report is quarterly and are both released at 11:00:00 am CT. In bold are days when both reports are released at the same time. The Crop Production reports are released at the same time as WASDE reports, and thus we capture their aggregate effect. The March Grain Stock report is released at the same time as the annual Prospective Plantings report, while the June Grain Stock report is released at the same time as the annual Acreage report.


Table 3.2. Descriptive statistics of the transaction prices of the nearby series (cents/bushel)

<table>
<thead>
<tr>
<th></th>
<th>Period 1</th>
<th>Period 2</th>
<th>Period 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>304.75</td>
<td>512.25</td>
<td>318.25</td>
</tr>
<tr>
<td>Max</td>
<td>799.25</td>
<td>849.00</td>
<td>746.25</td>
</tr>
<tr>
<td>Mean</td>
<td>563.64</td>
<td>725.55</td>
<td>453.94</td>
</tr>
<tr>
<td>Standard</td>
<td>2.74</td>
<td>3.25</td>
<td>1.77</td>
</tr>
<tr>
<td>deviation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of days</td>
<td>1091</td>
<td>156</td>
<td>736</td>
</tr>
</tbody>
</table>

Notes: The total number of days are 1983 from January 2008 to December 2015. The min, max, and mean correspond to minimum, maximum and average of all intraday prices within each period, and the standard deviation corresponds to the median standard deviation of the intraday prices across days in each period. The mean of intraday prices is higher in period 2, which corresponds to a period of high prices compared to periods 1 and 3.

Table 3.3. Number of jumps and jump days per period and type of day

<table>
<thead>
<tr>
<th></th>
<th>Period 1</th>
<th>Period 2</th>
<th>Period 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of jumps</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Announcement</td>
<td>11</td>
<td>8</td>
<td>130</td>
</tr>
<tr>
<td>Non-announcement</td>
<td>113</td>
<td>16</td>
<td>168</td>
</tr>
<tr>
<td>Total</td>
<td>124</td>
<td>24</td>
<td>298</td>
</tr>
<tr>
<td>Number of jump days (percent over total number of days)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Announcement</td>
<td>7 (11.48%)</td>
<td>5 (55.56%)</td>
<td>38 (88.37%)</td>
</tr>
<tr>
<td>Non-announcement</td>
<td>90</td>
<td>13</td>
<td>116</td>
</tr>
<tr>
<td>Total</td>
<td>97 (8.89%)</td>
<td>18 (11.54%)</td>
<td>154 (20.92%)</td>
</tr>
<tr>
<td>Number of clustered jumps</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Announcement – clustered jumps in same direction</td>
<td>0</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Announcement – clustered jumps in opposite direction</td>
<td>1</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Non-announcement – clustered jumps same direction</td>
<td>8</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Non-announcement – clustered jumps opposite direction</td>
<td>4</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Cumulative jump sizes in absolute value (cents/bushel)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Announcement</td>
<td>7.21</td>
<td>5.08</td>
<td>9.19</td>
</tr>
<tr>
<td>Non-Announcement</td>
<td>2.08</td>
<td>1.52</td>
<td>1.26</td>
</tr>
</tbody>
</table>

Notes: the size of positive price jumps on announcement days in period 1 is high and driven by the limit price move on 2010-10-08 of about 31 cents/bushel.
### Table 3.4. Jump size regression

|                      | $\sum | J S_t |$ Jump size |
|----------------------|--------------------------------------------------|
| $| S_t |$                |                                                                  |
| Surprise in absolute values | 0.139***                                           |
| $ANNXPOL_t$           | Dummy announcement day (real-time trading)         |
|                      | 0.009***                                           |
| $SPR_t$               | Dummy spring                                       |
|                      | 0.002                                              |
| $SUM_t$               | Dummy summer                                       |
|                      | 0.003                                              |
| $AUT_t$               | Dummy autumn                                       |
|                      | 0.002                                              |
| $D_{time-to-rollover}$| Dummy time-to-rollover                             |
|                      | 0.0003                                             |
| Constant              | 0.002                                              |

Observations 228

Adjusted $R^2$ 0.46

Notes: Significance levels: *p<0.1; **p<0.05; ***p<0.01. Standard errors are reported in parenthesis. $\sum | J S_t |$ corresponds to the summation of the absolute value of log-prices jumps size. The surprises in absolute value correspond to $|\log(A_t) - \log(E_t)|$ where $A_t$ is the observed U.S. corn ending stock in the monthly WASDE report and $E_t$, representing the expected stocks value, corresponds to the surveyed value of U.S. corn ending stocks conducted by Bloomberg (available starting on September 2009). $ANNXPOL_t$, representing real time trading of the reports, is an interaction term between the dummy announcement (one on announcement days and zero otherwise) and the dummy policy (one since May 21st, 2012 and zero before). To account for seasonality effect, we include seasonal dummies for Spring ($SPR_t$), Summer ($SUM_t$) and Autumn ($AUT_t$). The dummy time-to-rollover ($D_{time-to-rollover}$) is equal to one 40 days before rollover and zero otherwise. The Breusch–Godfrey and ARCH Engle tests fail to reject the null hypothesis of no serial correlation and ARCH effects in the residuals at the 5% significance level, respectively.
Table 3.5. Jump variation component as a proportion of overall volatility.

<table>
<thead>
<tr>
<th></th>
<th>Period 1</th>
<th></th>
<th>Period 2</th>
<th></th>
<th>Period 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$AJV_{\text{share}}$</td>
<td>$AJV_{\text{c.share}}$</td>
<td>$AJV_{\text{share}}$</td>
<td>$AJV_{\text{c.share}}$</td>
<td>$AJV_{\text{share}}$</td>
<td>$AJV_{\text{c.share}}$</td>
</tr>
<tr>
<td>1-tick</td>
<td>0.2167</td>
<td>0.0328*</td>
<td>0.2970</td>
<td>0.0190*</td>
<td>0.3639</td>
<td>0.0751*</td>
</tr>
<tr>
<td>1s</td>
<td>0.1406</td>
<td>0.0290*</td>
<td>0.1348</td>
<td>0.0401*</td>
<td>0.1501</td>
<td>0.0727*</td>
</tr>
<tr>
<td>5s</td>
<td>0.0957</td>
<td>0.0537*</td>
<td>0.1090</td>
<td>0.0346*</td>
<td>0.1286</td>
<td>0.0998*</td>
</tr>
<tr>
<td>10s</td>
<td>0.0972</td>
<td>0.0548*</td>
<td>0.0928</td>
<td>0.0411*</td>
<td>0.1265</td>
<td>0.1034</td>
</tr>
<tr>
<td>15s</td>
<td>0.0825</td>
<td>0.0555</td>
<td>0.0747</td>
<td>0.0879</td>
<td>0.1165</td>
<td>0.1027</td>
</tr>
<tr>
<td>30s</td>
<td>0.0941</td>
<td>0.0446</td>
<td>0.0649</td>
<td>0.0501</td>
<td>0.1081</td>
<td>0.0949</td>
</tr>
<tr>
<td>1m</td>
<td>0.0493</td>
<td>0.0500</td>
<td>0.0312</td>
<td>0.0445</td>
<td>0.1269</td>
<td>0.0970</td>
</tr>
<tr>
<td>2m</td>
<td>0.0404</td>
<td>0.0499</td>
<td>0.0563</td>
<td>0.0650</td>
<td>0.1235</td>
<td>0.0970</td>
</tr>
<tr>
<td>3m</td>
<td>0.0559</td>
<td>-</td>
<td>0.0654</td>
<td>-</td>
<td>0.1173</td>
<td>-</td>
</tr>
<tr>
<td>4m</td>
<td>0.0631</td>
<td>-</td>
<td>0.0875</td>
<td>-</td>
<td>0.1397</td>
<td>-</td>
</tr>
<tr>
<td>5m</td>
<td>0.0508</td>
<td>-</td>
<td>0.0872</td>
<td>-</td>
<td>0.1237</td>
<td>-</td>
</tr>
<tr>
<td>10m</td>
<td>0.0636</td>
<td>-</td>
<td>0.0638</td>
<td>-</td>
<td>0.1325</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Period 1</th>
<th></th>
<th>Period 2</th>
<th></th>
<th>Period 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$AJV_{\text{share}}$</td>
<td>$AJV_{\text{c.share}}$</td>
<td>$AJV_{\text{share}}$</td>
<td>$AJV_{\text{c.share}}$</td>
<td>$AJV_{\text{share}}$</td>
<td>$AJV_{\text{c.share}}$</td>
</tr>
<tr>
<td>1-tick</td>
<td>0.1829</td>
<td>0.0183*</td>
<td>0.3399</td>
<td>0.0240*</td>
<td>0.3223</td>
<td>0.0610*</td>
</tr>
<tr>
<td>1s</td>
<td>0.1424</td>
<td>0.0418*</td>
<td>0.1419</td>
<td>0.0407*</td>
<td>0.1453</td>
<td>0.0775*</td>
</tr>
<tr>
<td>5s</td>
<td>0.1205</td>
<td>0.0544*</td>
<td>0.1201</td>
<td>0.0523*</td>
<td>0.1365</td>
<td>0.0773*</td>
</tr>
<tr>
<td>10s</td>
<td>0.1005</td>
<td>0.0614*</td>
<td>0.1006</td>
<td>0.0625*</td>
<td>0.1285</td>
<td>0.0947*</td>
</tr>
<tr>
<td>15s</td>
<td>0.0886</td>
<td>0.0585*</td>
<td>0.0883</td>
<td>0.0586*</td>
<td>0.1183</td>
<td>0.0874*</td>
</tr>
<tr>
<td>30s</td>
<td>0.0681</td>
<td>0.0521*</td>
<td>0.0691</td>
<td>0.0499*</td>
<td>0.1079</td>
<td>0.0769*</td>
</tr>
<tr>
<td>1m</td>
<td>0.0616</td>
<td>0.0512</td>
<td>0.0630</td>
<td>0.0533</td>
<td>0.0871</td>
<td>0.0645*</td>
</tr>
<tr>
<td>2m</td>
<td>0.0495</td>
<td>0.0512</td>
<td>0.0520</td>
<td>0.0463</td>
<td>0.0671</td>
<td>0.0644</td>
</tr>
<tr>
<td>3m</td>
<td>0.0423</td>
<td>-</td>
<td>0.0465</td>
<td>-</td>
<td>0.0650</td>
<td>-</td>
</tr>
<tr>
<td>4m</td>
<td>0.0438</td>
<td>-</td>
<td>0.0490</td>
<td>-</td>
<td>0.0627</td>
<td>-</td>
</tr>
<tr>
<td>5m</td>
<td>0.0411</td>
<td>-</td>
<td>0.0481</td>
<td>-</td>
<td>0.0587</td>
<td>-</td>
</tr>
<tr>
<td>10m</td>
<td>0.0446</td>
<td>-</td>
<td>0.0530</td>
<td>-</td>
<td>0.0563</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes: This table presents detailed results of jump shares for announcement and non-announcement days in the three periods. $AJV_{\text{share}} = \frac{ARV - ABV}{ARV}$ is the proportion of daily jump risk including noise on overall realized variance, where $ARV$ is the daily annualized realized volatility ($\sqrt{252 \times RV}$) and $ABV$ is the daily annualized bipower variation ($\sqrt{252 \times BV}$). $AJV_{\text{c.share}} = \frac{ARV^c - ABV^c}{ARV^c}$ is the proportion of daily jump risk filtered for noise, where $ARV^c$ is the noise-filtered daily annualized realized volatility ($\sqrt{252 \times RV^c}$) and $ABV^c$ is the noise-filtered daily annualized bipower variation ($\sqrt{252 \times BV^c}$). We exclude the same limit price moves days as in the jump identification test analysis for which the $ARV^c$ is zero making the computation of the share impossible. The * indicates when the Wilcoxon test null hypothesis is rejected at the 1% significance level, suggesting that the two curves are different.
3.7. References


CFTC. 2015. “Remarks of Chairman Timothy Massad before the Conference on the Evolving
Structure of the U.S. Treasury Market.” Available at:
http://www.cftc.gov/PressRoom/SpeechesTestimony/opamassad-30


Meyer, G. 2018. “Last commodities hedge funds go off beaten track.” Financial Times. Available at: https://www.ft.com/content/fbe5a554-36b3-11e8-8eee-e06bde01c544


*Reuters.* Available at: https://www.reuters.com/article/us-commodities-funds-computers/data-overload-commodity-hedge-funds-close-as-computers-dominate-idUSKBN1FW0E4


3.8. Supplementary Information

3.8.1. Noise filtering approaches

As explained in Jacod et al. (2009) the literature has provided different methods to filter for microstructure noise, including resampling and pre-averaging techniques. Microstructure noise is only present at very high frequency data and has been shown to dissipate at five-minute or lower sampling frequencies. Resampling techniques are based on these findings and choose a resampling frequency accordingly. Pre-averaging techniques, in contrast, average prices over a window whose length increases with the number of intraday observations. While pre-averaging was initially applied on non-overlapping windows (Podolskij and Vetter 2009), procedures have been recently developed to allow for a moving-average window (Jacod et al. 2009) that limits the number of observations that are lost. It is important to select a proper filtration, as not all filtration methods are compatible for all statistical analyses. We apply two main techniques in our article, the Lee and Mykland (2008) jump test and the Christensen et al. (2014) method to identify the bipower and jump variation components of realized variance. Below we describe the microstructure noise cleaning methods that we use, depending on the technique applied.

3.8.1.1. Lee and Mykland’s (2012) pre-averaging approach

Lee and Mykland’s (2008) test assumes returns to have stationary and independent increments. Non-overlapping pre-averaging techniques yield asymptotically independent returns (Gonçalves, Hounyo and Meddahi 2014), which is not the case with overlapping methods that create moving average structures. Hence, we employ the non-overlapping noise-filtering approach by Lee and Mykland (2012), a technique also applied in recent research for jumps detection (e.g. Brogaard et al. 2018). This method combines resampling and pre-averaging methods. First, autocorrelation in returns is removed by subsampling every \( k \) observations, where \( k-1 \) is the autocorrelation order.
Second, subsampled prices are averaged within non-overlapping windows of size $M$ (see Lee and Mykland 2012). In our study, daily autocorrelation functions of the transactions price identify an average serial correlation of four on average. As a result, we subsample prices every five ticks and we then smooth the subsampled prices over the non-overlapping intervals of $M = C \lfloor N/k \rfloor^{1/2}$ observations each, where $N$ is the number of intraday observations and $C$ is a parameter whose optimal value needs to be identified. We choose the $C$ parameter so that the annualized realized volatility ($ARV$) of the filtered price is the closest to the $ARV$ at 5-minute sampling frequency. The latter is supposed to reflect the efficient price realized volatility, as 5-minute resampled prices are commonly accepted noise-free prices (Hansen and Lunde, 2006; Wu et al., 2015). The distance between the two measures is minimized at $M = 3.1$ (see figure 3.6. below). Note that the resampling and non-overlapping pre-averaging results yield to noise-filtered prices observed, on average, every 30 to 60 seconds.

Figure 3.6. Comparison of annualized realized volatilities for prices subsampled every 5-minutes and for the noise-filtered tick transaction prices

Note: The $ARV^c$ is an average over the different days of the annualized daily realized volatility (squared root of $\sum_{i=1}^N (r_t)^2$) of noise-filtered prices. The benchmark $ARV$ 5-min is constructed using 5-minute sampled transaction prices.
3.8.1.2. Christensen et al.’s (2014) pre-averaging approach

To identify the bipower and jump variation components of realized variance, Christensen et al. (2014) pre-average observed prices in an overlapping local neighborhood of \( K \) observations.

\[
\hat{r}_{t_{i,K}} = \frac{1}{K} \left( \sum_{j=K/2}^{K-1} \hat{P}_{t_{i+j}} - \sum_{j=0}^{K/2-1} \hat{P}_{t_{i+j}} \right)
\]

where \( K = \theta \sqrt{N} + o(N^{-1/4}) \). Their asymptotics are based on overlapping pre-averaging techniques and we thus follow their filtering approach. Note the similarities between \( M = C \left[ N/k \right]^{1/2} \) and \( K \); the main difference between the two parameters is \( k \) which corresponds to the observations lost due to autocorrelation resampling. The \( \theta \) parameter is chosen so that the annualized realized volatility (ARV) of the filtered price is the closest to the ARV at 5-minute sampling frequency. We find that \( \theta = 0.4 \) is the optimal value to obtain a noised filtered price series (see figure 3.7.).

![Figure 3.7. Robustness check on \( \theta \) and \( M \) regarding the annualized realized volatility.](image)

Notes: The ARV\(^c\) is an average over the different days of the annualized squared root of

\[
\frac{N}{N-K+2} \frac{1}{K} \sum_{i=0}^{N-K+1} \left( \hat{r}_{t_{i,K}} \right)^2 - \frac{\tilde{\omega}^2}{\vartheta^2 \psi K} \hat{r}_{t_{i,K}} \quad \text{is the returns based on pre-averaged prices using} \quad K = \theta \sqrt{N} \quad \text{where} \quad N \quad \text{is the number of intraday prices. The benchmark ARV 5-min is constructed using 5-minute sampled transaction prices.}
3.8.2. Robustness analysis of the jump test results with a longer window size: $W = 185$.

Figure 3.8. Distribution of intraday jumps out across intraday time intervals in each period with $W = 185$, January 15, 2008 – December 4, 2015.

With $W = 185$ (which is approximately equivalent to using the data over the preceding 100 minutes), a total 259 days have at least one jump, and a total of 439 jumps are detected. On announcement days, the average number of jumps per day is 0.20 in period 1, 0.67 in period 2, and 3.12 in period 3; while on non-announcement days, the average number of jumps per day is 0.10, 0.12, and 0.24 in periods 1, 2, and 3, respectively.
CHAPTER 4
FORECASTING THE CORN FUTURES REALIZED VOLATILITY

4.1. Introduction

The ability to forecast agricultural commodity futures price volatility is critical to inform market participants of their future risk exposure, and to guide their production, hedging, and inventory decisions. Proper volatility modeling and forecasting should allow for the long-memory usually identified in volatility patterns (Jin and Frechette 2004; Baillie et al. 2006; Karali and Power 2013; Choi, Yu and Zivot 2010; Wang 2014) that leads to persistent price fluctuations. Traders in futures contracts often adjust their hedging strategies to this persistence by hedging across multiple contracts or trading in both distant and nearby contracts (Fett and Haynes 2017). Volatility modelling and forecasting should also allow for the diversity of agricultural futures market participants with different expectations at various time horizons. Fett and Haynes (2017) study the composition of the corn futures contracts from the Chicago Mercantile Exchange (CME) on June 2016 on a volume basis. They find a high participation of individuals and corporations who usually hedge business risks on a medium or long-term horizon. Principal trading firms are also present and act as short-term intermediaries commonly in intraday trading. Additionally, futures prices volatility modeling should allow for processes such as seasonality and changing market conditions that can lead to nonlinear fluctuations in volatility.

In the era of artificial intelligence and machine learning, the use of algorithms that can learn from and can forecast the data has gained momentum. Influenced by biological neural networks, artificial neural networks (ANN) learning algorithms are nonlinear statistical tools used to identify patterns in data and make predictions. The popularity of ANN has surged with growing evidence of their accuracy in approximating nonlinear functions and providing better forecasts relative to
alternate methods (Franses and van Dijk 2000). Since the literature suggests that different forms of nonlinearities are present in agricultural futures price volatility, ANN models emerge as a potentially useful forecasting tool.

Given the diversity of corn futures market participants identified by previous literature (Fett and Haynes 2017), the objective of this research is to forecast corn futures price volatility through a nonlinear heterogeneous autoregressive (HAR) model estimated through ANN using intraday prices and compare it to the standard HAR model. The linear HAR framework (Corsi 2009), based on realized volatility, builds on the heterogeneous market hypothesis and assumes that the interaction between traders operating at different time horizons leads to long-memory volatility. The HAR model captures the heterogeneous nature of traders, information arrival into the market and the long-run memory process through a multicomponent volatility model. The intuition behind HAR is straightforward, it models daily realized volatility forecasts as a function of past daily, weekly and monthly components. While the HAR developed by Corsi (2009) has proven to efficiently capture in- and out-of-sample long-term dependencies in the realized volatilities, other time-series properties such as structural breaks or time-varying persistence in volatility may interfere with HAR model forecasting abilities (Maheu and McCurdy 2002; Audrino and Knaus 2016). Kuan and White (1994) discuss the potential of ANN to model multiple forms of nonlinearity and its applicability to time series econometrics.

To our knowledge, no previous study has used ANN to model and predict agricultural price volatility. Since the literature suggests that different forms of nonlinearities are present in agricultural futures price volatility, ANN models may be particularly useful for our research purposes. In this article, we model and predict, both nearby and more distant horizons, daily
agricultural commodity futures price volatility. We use intraday data and daily exogenous variables to estimate both a linear HAR and nonlinear HAR approach through ANN.

The contributions of this article are several. First, this work is pioneer in using ANN methods to model agricultural futures prices volatility. Corn futures volatility is characterized by strong nonlinearities that are related to: U.S. Department of Agriculture (USDA) report release days (Adjemian and Irwin 2018); strong seasonality (Karali and Power 2013; Egelkraut, Garcia and Sherrick 2007); time-to-delivery effects (Goodwin and Schnepf 2000); and changes in macroeconomic conditions (Karali and Power 2013). Combined, these effects create nonlinearities of unknown forms that require highly flexible and adaptive methods such as ANN models. Second, by using a HAR framework, we allow for the heterogeneous effects of different market participants on the volatility which reflects shorter- and longer-term trading activity. Third, we also conduct multi-step horizons forecasting. While forecasting corn volatility using option prices, previous research has traditionally used a 2-month horizon due to options expirations (e.g. Wu et al. 2015), other market participants may rely on medium- or short-term forecasting to adjust their risk strategies. For instance, traders relying on USDA weekly export sales reports24 (such as shippers) or weekly corn options traders25 will be more interested in a shorter-term forecasting of the corn futures prices volatility. Futures markets volatility forecasts are also crucial in the pricing of the options, a higher volatility will result in a higher priced options (Kroner, Kneafsey and Claessens 1995). On the other hand, high-speed liquidity providers might rely on short-run volatility forecasts to decide how to adjust their strategies for the next day. As a result, our findings should be useful to a wider range of market participants.


25 https://www.cmegroup.com/trading/agricultural/weekly-options-on-grain-futures.html
Our findings indicate that the nonlinear ANN model works better at all horizons (1-day, 1-week, and 1-month) than the standard HAR model even when the HAR model is augmented for seasonality or public information shocks. Overall, this finding points to the importance of accounting for complex nonlinearities through general models when forecasting realized volatility in the corn futures market. These results are robust for various window sizes of the out-of-sample forecasting rolling approach. Additionally, we analyze the time-varying dynamics of the lags structure in the ANN model. The results suggest that the flexible lags structure is preferred to the fixed lags structure (1, 5, 22) in the ANN specification supporting the findings by Audrino, Huang and Okhrin (2018) that fixed lag structure of the HAR model is not always accurate.

4.2. Relevant Literature

The literature assessing long-memory in agricultural futures market volatility is relatively recent and there is still no consensus on the causes of long-memory in volatility, or whether it could be spurious. Crato and Ray (2000) show that commodity markets (including agricultural) volatilities have higher long-run dependencies than currency futures market volatilities. Jin and Frechette (2004) further confirm evidence of long-term memory in price volatility of 14 agricultural commodities through fractionally integrated models, FIGARCH. The authors hypothesize the sources of long-memory come from the agricultural sector’s structural characteristics such as the low supply elasticity due to annual or multi-annual production cycles and the time required to adjust stock levels. Although they emphasize the characteristics of the agricultural markets as the main sources of long-memory, they also consider trader heterogeneity as a possible driver of long-memory due to asynchronous switching in the traders’ demand curves when new information arrives. However, they do not probe into this hypothesis in more detail. Other articles providing
evidence of long-memory in agricultural commodity price volatility are Chen, Daigler and Parhizgari (2011) and Baillie et al. (2006).

Other authors suggest that persistence in volatility might not be relevant or might be spurious and result from model misspecification. Karali and Power (2013) find only weak evidence of volatility persistence when using a spline-GARCH to separate the high- and low-frequency components of the realized volatilities for 11 commodity futures markets. Wang (2014) examines the sources of long-memory in grain futures markets using daily settlement prices from 1989 to 2011. Using a FIGARCH model, he observes large part of the long-memory process fades after volatility models account for seasonality and structural breaks. Choi, Yu and Zivot (2010) use 30-minute returns to compute the daily log-RV of the Deutschmark/Dollar, Yen/Dollar, and Deutschmark/Yen exchange rates. They find that removing the structural breaks in the mean greatly reduces persistence in the realized volatility.

The discussions on the sources of long-memory and whether it improves volatility forecasting at various horizons re-emerged around the now popular HAR model (Corsi 2009). The model allows for traders’ heterogeneity and relies on realized volatility measures. Corsi (2009) uses both a simulation approach and an empirical case based on tick-by-tick data for U.S. dollars/Swiss Franc exchange rate, S&P500 futures, and 30-year US treasury bond futures. He fits the model to the data with standard ordinary least squares (OLS) with a Newey-West covariance correction for serial correlation. He finds the HAR captures well the slow decay of the daily realized volatility in a similar manner to a fractionally integrated model specification. The main findings indicate that the HAR model generates better out-of-sample forecasts than short-term memory models. Since Corsi (2009), several extensions of the HAR have flourished in the volatility forecasting literature. Notable progress has been made using more flexible approaches that allow for a time-varying
predictor of the RV measure. Tian, Yang, and Chen (2017) are among the first to forecast realized volatility in agricultural markets. They study six agricultural commodities futures contracts traded in Chinese exchanges. They employ a time-varying HAR by modeling a gamma autoregressive process to capture the nonlinear dynamics of positive-valued time series and find it improves the forecasting ability of the HAR model. While they account for complex nonlinearities in the HAR model, they do so by imposing parameters’ prior distributions, which can constrain the model in selecting the time-varying set of covariates. Recently, Baillie et al. (2019) highlight the importance of accounting for time-varying parameters when modeling long-memory process, however they do not assess the model forecasting performance. In parallel, machine learning techniques have also recently appeared in the sphere of long-memory volatility. Audrino and Knaus (2016) use a least absolute shrinkage and selection operator (LASSO) model that reveals structural breaks in the long-memory process. However, the authors find no specific improvements in forecasting at a one-day horizon when using their LASSO model compared to the standard HAR model and provide no evidence of the model’s performance at longer horizons.

Only a small set of studies has explored agricultural commodities markets using ANN and they mostly focus on forecasting price levels (Kohzadi et al. 1996; Hamm and Brorsen 1997). While the ANN model can approximate general forms of nonlinearities, the literature on the use of ANN models to predict volatility still remains scarce, offers mixed results regarding their usefulness, and has much scope for contributions. Previous studies have mostly focused on currency (Diebold and Nason 1990), financial markets (Hamid and Iqbal 2004; Hillebrand and Medeiros 2010; McAleer and Medeiros 2011; Fernandes et al. 2014; Donaldson and Kamstra 1997) or energy markets (Baruník and Krehlík 2016).
Very few studies employ ANN models to account for nonlinearities in volatility (Donaldson and Kamstra 1997; Hamid and Iqbal 2004; McAleer & Medeiros 2011; Fernandes et al. 2014; and Baruník and Krehlík 2016). Donaldson and Kamstra (1997) use a semi-nonparametric ANN-GARCH model to capture the nonlinear relationship between past return innovations and future volatility of stock indexes. They examine the out-of-sample performance of the ANN-GARCH model in a rolling window approach and find, for one-step ahead forecasts, it explains variance effects not captured by alternative models (simple GARCH, EGARCH, GJR). Hamid & Iqbal (2004) forecast volatility of the S&P 500 daily prices using neural network models and compare their performance in forecasting accuracy with implied volatility forecasts. Their findings suggest that forecasts from neural networks outperform implied volatility forecasts. McAleer & Medeiros (2011) employ an ANN model of the HAR and show that it captures pretty well the nonlinear behavior of realized volatility of the S&P500 and FTSE100 futures indices when Bayesian Regularization (BR) is employed and outperforms the simple HAR. Fernandes et al. (2014) compare HAR and asymmetric HAR to ANN-HAR-X to forecast the VIX index. They find that it is very hard to beat the standard HAR model which suggests there are few nonlinearities in the volatility index. These studies highlight that the data frequency and the type of volatility measures can affect the results. Others have compared the linear HAR and HAR-ANN models using a wide range of RV volatility measures and compare their forecasting performance to conditional volatility measure (GARCH) for energy markets (Baruník and Krehlík 2016). Using the model confidence set approach (Hansen, Lunde and Nason 2011) with average Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) loss functions, the ANN models outperform the HAR, ARFIMA, and GARCH at longer horizon while the forecasting performances are closer between

26 Bayesian regularization uses a network training function based on a Levenberg-Marquardt algorithm.
the models at shorter horizons depending on the energy market considered. In sum, while some studies find modest or no improvement from using the ANN model to forecast volatility, they mostly look at short-term horizon (one-day ahead). More recent studies find that ANN can be useful for forecasting RV at longer horizons, but much work remains to be done to assess the performance of ANN model for volatility forecasting especially at different horizons. This is particularly important for agricultural markets that are marked by strong seasonality, severe weather shocks or information shocks. Such shocks may affect volatility differently in the short-run versus long-run, which could result in trading activities at different horizons.

4.3. Methods

To forecast volatility in the presence of heterogeneous market participants, we adopt Corsi’s (2009) HAR specification that models daily realized volatility against past daily, weekly and monthly volatility. Since the causes of long-memory are not well established, we compare the HAR against a purely statistical model not linked to any specific theory that also captures long-memory, the ARFIMA model. To evaluate the role of known nonlinearities on RV, we assess how seasonal terms with Fourier forms and dummies for days with USDA public announcements improve forecast accuracy of the linear HAR model.

Although the HAR specification captures long-term dependencies, it does not allow for nonlinearities such as structural breaks, which we model using the artificial neural network model. First, we specify an ANN nonlinear structure with the same three components of the HAR (daily, weekly, and monthly), noted HAR-ANN. Second, we adopt an ANN specification with 22 daily realized volatility lagged input variables. The latter represents a more flexible lag structure than the HAR model recently proposed to capture the time-varying dynamics of volatility (Barunik and

27 More details about these studies can be found the appendix 4.8.1.
Krehlik 2016; Audrino and Knaus 2016; McAleer and Medeiros 2011; Yang et al. 2017). We refer to the ANN with unrestricted lag structure as the ANN model. To further assess the performance of the ANN models, we specify the ANN-X model with unrestricted lag structure and exogenous variables. Finally, we evaluate the forecast accuracy of the competing models using the pairwise Modified Diebold and Mariano test and the Model Confidence Set approach.

4.3.1. Linear HAR Model

With increasing availability of intraday high frequency data, refined measures of integrated volatility (IV) have been developed. Andersen et al. (2003) show that realized volatility (RV) can approximate integrated volatility (IV) better than other parametric conventional models such as GARCH. Realized volatility can be expressed as

$$RV_t = \sqrt{\sum_{j=1}^{N_j} [p_{j,t} - p_{j-1,t}]^2} = \sqrt{\sum_{j=1}^{N_j} r_{j,t}^2},$$

where $p_{j,t}$ is the $j$th intraday log-price observation on day $t$, the $j$th intraday log-return is $r_{j,t}$, and $N$ is the number of returns within the day. In practice, however, intraday returns are plagued with microstructure noise that creates return autocorrelation (Bandi and Russell 2006; Hansen and Lunde 2006) and renders RV a biased estimator of IV. To address this problem, we use a consistent daily IV estimator that corrects for autocorrelations in returns caused by microstructure noise (Jacod et al. 2009; Christensen et al 2014),

$$\tilde{IV}_t^d = \frac{N}{N-2K-2} \frac{1}{K \varphi_K} \sum_{l=0}^{N-2K-1} |r_{l,K}^*|^2 - \frac{\hat{\omega}}{\theta^2 \varphi_K}$$ (4.1)

where returns are calculated on smoothed log-price series that consist of the original log-prices pre-averaged in a local neighborhood of $K$ observations: $r_{l,K}^* = \frac{1}{K} \left( \sum_{j=K/2}^{K-1} p_j - \sum_{j=0}^{K/2-1} p_j \right)$, with $K = \theta \sqrt{N} + o(N^{-1/4})$, $\varphi = 1 + 2K^{-2}/12$ and $\theta = 0.4$ a parameter whose value is chosen following Couleau et al. (2018). The last term on the right hand side in (4.1) is a bias correction term that serves the purpose of removing residual noise $\hat{\omega} = -\frac{1}{N-1} \sum_{l=2}^{N} |r_{l}^*||r_{l-1}^*|$ (Oomen 2006).
We define the weekly and monthly IV as equations (4.2) and (4.3), respectively,

\[
\hat{IV}_{t-1}^w = \frac{1}{5} \sum_{i=1}^{5} \hat{IV}_{t-i}^d 
\]

where \(\hat{IV}_{t-1}^w\) is the average lagged weekly realized volatility\(^{28}\) and

\[
\hat{IV}_{t-1}^m = \frac{1}{22} \sum_{i=1}^{22} \hat{IV}_{t-i}^d 
\]

where \(\hat{IV}_{t-1}^m\) is the average lagged monthly realized volatility. The simplest logarithmic version\(^{29}\) of the HAR model is defined as

\[
\log(\hat{IV}_t^d) = \alpha_0 + \alpha_1 \log(\hat{IV}_{t-1}^d) + \alpha_2 \log(\hat{IV}_{t-1}^w) + \alpha_3 \log(\hat{IV}_{t-1}^m) + u_t \quad (4.4)
\]

where \(u_t \sim N(0, \sigma_u^2)\). The HAR model is easily estimated by ordinary least squares. Its simplicity and tractability make it a popular model to forecast volatility.

The HAR model uses past volatility information to forecast future volatility. Including exogenous variables may improve HAR forecasting accuracy, especially at long-term horizons (Fernandes et al. 2014). Hillebrand and Medeiros (2010), McAleer and Medeiros (2011) and Fernandes et al. (2014) have all considered different exogenous variables in their HAR specifications such as past cumulative returns, macroeconomic variables, or dummies for the day of the week and for macroeconomic announcement days. Following this research, we use a generalized version of the HAR that includes a vector of market-specific, seasonal and other exogenous variables. We denote this model as HAR-X and express it as

\[
\log(\hat{IV}_t^d) = \alpha_0 + \alpha_1 s_t + \alpha_2 \log(\hat{IV}_{t-1}^d) + \alpha_2 \log(\hat{IV}_{t-1}^w) + \alpha_4 \log(\hat{IV}_{t-1}^m) + y x_t + u_t \quad (4.5)
\]

\(^{28}\) We use the term realized volatility to refer to the realized integrated volatility (free of microstructure noise) in the rest of the paper as it is usually done in the literature (Andersen et al. 2003).

\(^{29}\) We use the logarithm form as the \(\hat{IV}\) is not normally distributed. In addition, it does require imposing the non-negativity constraint on the estimation (Corsi 2009; Duong and Swanson 2015).
where \( s_t \) is the seasonal Fourier function, \( x_t \) is a vector that contains dummy variables for announcement days and the lagged daily trading volume and \( y \) is a vector of parameters. Note that jumps in volatility are mainly captured by announcement days (Couleau, Serra and Garcia 2018).

4.3.2. Artificial Neural Network Model

The Artificial Neural Network (ANN) model application to nonlinear time series in empirical analysis has been discussed in Franses and van Dijk (2000). We provide an intuitive explanation of an ANN through the single layer feedforward neural network schematically represented in Figure 4.1. The two neurons of the hidden (H) layer receive a signal from two inputs (I) and become active only after input activity passes a certain threshold, which captures the nonlinear features of the ANN model. The hidden layer activation function is denoted by \( G(I) \) and its role is to turn “on or off” the signal from the inputs. \( G(I) \) is often specified as a sigmoid (logistic) function \( G(I) = \frac{1}{1+\exp(-I)} \), whose shape allows a nonlinear smooth transition from the “on” to the “off” position. Neurons of the hidden layer process the information from the input and send it to the output (O). The output function \( F(I, G) = 01 \) processes the hidden units’ information into output and can either predict a categorical variable (e.g., bankruptcy, yes or no) or a numerical variable (e.g., annual income). We use the latter type (also known as regression type ANN) as we want to forecast price return volatility levels and specify \( F(I, G) \) as in equation (4.6) below. Additional bias units (denoted by B1 and B2 in Figure 4.1) are added to the hidden and output layers to capture the intercepts in equation (4.6).
Figure 4.1. Schematic artificial feedforward neural network.

We specify $F(.)$ as an auto-regressive neural network with exogenous variables, ANN-X (Barušík and Krehlík 2016) as follows

$$y_t = F(\phi_0, G(Y_{t-1}, X_t)) =$$

$$\phi_0 + \sum_{m=1}^{M} \lambda_m \left( \gamma_{0,m}y_{t-1} + \gamma_{1,m}y_{t-2} + \gamma_{22,m}y_{t-22} + \delta_{1,m}x_{1,t} + \cdots + \delta_{l,m}x_{l,t} \right) + u_t,$$  \hspace{1cm} (4.6)

Where $y_{t-i} = \log\left(\overline{IV}_{t-i}^d\right), i = 0, \ldots, 22$. Note the original HAR developed by Corsi (2009), which centers on lags 1, 5, and 22, is a restricted case of the ANN model (equation 4.6) with unrestricted AR terms (lags from 1 to 22). Here, we propose to work with the unrestricted, more flexible ANN model. However, results from the HAR-ANN model are presented for comparison, i.e. the ANN model is specified with lags aggregated at the daily, weekly and monthly frequency as it is the case in the linear HAR. Variables $x_{1,t}$ to $x_{l,t}$ are exogenous variables where $l$ is the number of exogenous variables. The bias terms, $\gamma_{0,m}$ for $m = 1$ to $M$, are the number of hidden units and $\phi_0$ and correspond to B1 and B2 in Figure 4.1, respectively. They capture the intercepts as in a regression analysis. Other model parameters are $\lambda_m, \gamma_{1,m}$ to $\gamma_{p,m}$, and $\delta_{1,m}$ to $\delta_{l,m}$. Finally, $u_t$ is
the error term with zero-mean and finite variance and $G(.)$ is specified as a logistic function\(^{30}\) which results in a $F(Y_{t-1}, X_t)$ expression as follows:

$$F(Y_{t-1}, X_t) = \phi_0 + \sum_{m=1}^{M} \frac{\lambda_m}{1 + e^{-y_m Y_{t-1} - \delta_m X_t}} + u_t,$$

where $Y_{t-1} = (y_{t-1}, ..., y_{t-22})$, $X_t = (x_1, ..., x_l)$, $Y_m = (y_{0,m}, y_{1,m}, ..., y_{22,m})$, and $\delta_m = (\delta_{1,m}, ..., \delta_{l,m})$. Note this ANN corresponds to a single hidden layer and $M$ hidden units. Parameter $M$ is often selected by cross-validation, which is based on minimizing the Mean Squared Error (MSE) with $M$ varying between 1 and the total number of inputs (see e.g. Liang et al. 2006). The ANN model has parameters called weights $\gamma_m$ and $\delta_m$, which are weighting the information from input variables depending on their importance. To estimate the model parameters, we use the sum of squared errors as our measure of fit (error function),

$$R(\Gamma) = \frac{1}{2} \sum_{t=1}^{T} [y_t - F(Y_{t-1}, X_t)]^2$$

where $\Gamma = (\phi_0, \gamma_m, \delta_m)$ and $T$ is the sample size. ANN models “learn” from the data to estimate the model parameters. The learning process is reflected on the adjustment of weights through a learning algorithm. The backpropagation algorithm is commonly used (Werbos 1974; Rumelhart, Hinton and Williams 1986). Intuitively, this procedure consists of repeatedly adjusting the weights in the network to minimize the difference between the actual output vector and the desired output vector. Practically, the standard backpropagation algorithm requires knowledge on how the weights and the biases change the objective function $R(\Gamma)$, which requires the gradient $\nabla R(\Gamma)$.

\(^{30}\) Another function could be used, for instance, the hyperbolic tangent function is also often used as an activation function in feedforward backpropagation neural networks, but mostly when the output consists of a classification problem. Hence, we use the logistic function here.
of function $R(\Gamma)$. The signs and sizes of the derivatives guide the search for the optimal weights and biases. Given the estimates of $\hat{\Gamma}$, the fitted values and the residuals are computed. These residuals are fed back into the network iteratively, and the weights and bias parameters are adjusted until the objective function is minimized. A slightly different version of the backpropagation algorithm, the resilient backpropagation algorithm, based on the gradient descent, was developed by Riedmiller and Braun (1993) to increase computational efficiency. The resilient backpropagation algorithm relies only on the sign of the gradient instead of its magnitude (as is the case in the standard backpropagation algorithm which is often imprecise). We use the resilient backpropagation algorithm in this research.

The input variables of the HAR-ANN model are the lags of the log-IV for the three daily, weekly, and monthly components as defined in the linear HAR model. For the ANN model the input variables are the lags from 1 to 22 of the daily log-IV, which allows assessing the advantage of using a flexible lag structure in volatility forecasting while nonlinearities are taken into account.

For out-of-sample forecasting assessment, we split the sample into a training set and a test set. We choose the latter to represent at least 30% of the total size.\textsuperscript{31} The training sample size is used to train the ANN and to evaluate the forecasting accuracy of the model. Before we start estimating the ANN model, a number of initial conditions need to be determined, which are rarely discussed in econometric applications of ANN models. The choice of the model parametrization, evolving algorithms, and types of neural network can vary from one study to another which makes it difficult to compare the diverse specifications and contributes to the “black-box” reputation of these models. In the following subsections, we provide these details, specifically discussing data

\textsuperscript{31} Zhang, Patuwo and Hu (1998) review the neural network literature but find there is no rule for splitting the sample into the training and test samples.
normalization, initial weights, learning rates, number of epochs, and threshold of the partial derivative.

4.3.3. Data normalization

The behavior of neural networks is sensitive to the scale of the variables. To improve the estimation of the neural network, we follow previous research and scale the variables to be between [0,1] using

\[ Y_t^{scaled} = \frac{Y_t - \min(Y_t)}{\max(Y_t) - \min(Y_t)} \quad \text{and} \quad X_t^{scaled} = \frac{X_t - \min(X_t)}{\max(X_t) - \min(X_t)}. \]

The advantage of this normalization is that it makes the distribution of these variables (input and output) approximately uniform which is important to facilitate the learning approach since the logistic function output domain lies between 0 and 1. Once the ANN model is trained, the forecasted values are unscaled by taking, \( Y_t^{scaled} \times (\max(Y_t) - \min(Y_t)) + \min(Y_t). \)

Initial weights

Initial weights are randomly selected from a uniform distribution (-1,1) (Rumelhart et al. 1986; Riedmiller and Braun 1993). The choice of the initial weights set is critical to find the global minimum if local minima exist. Kaastra and Boyd (1996) suggest that five to ten random sets of starting weights can improve the chances of reaching a global minimum. Donaldson and Kamstra (1997) draw 5 different randomly produced sets of the weights and choose the estimated ANN that best fits the in-sample model, while Hamm and Brorsen (2000) estimate their ANN with 10 randomly chosen values of the starting parameters and select the ANN model with the starting values with the lowest sum of squared errors. This approach has the advantage of making the study replicable (Hamm and Brorsen 1997). In this study, we use 20 randomly selected starting values of the weights and bias (equation 4.6) parameters and select the starting parameter set that leads to the smallest MSE.
Learning rate

The learning rate is used to scale the first-derivative and has an important role on the time needed to reach convergence. A higher leaning rate will increase the training time. We allow the learning rate factor to vary between 0.5 and 1.2 as recommended by Riedmiller and Braun (1993). Every time the partial derivative of the error function of the corresponding weight change its sign, the update-value (of the weight) is decreased by a factor (‘decrease’ factor) to reach a minimum. If the derivative keeps its sign, the update-value is slightly increased to accelerate convergence.\textsuperscript{32} Indeed, for the decrease factor, they suggest that on average, taking half of the update-value is a good guess to reach the global minimum. Regarding the increase factor, the value of 1.2 provides good results with the fastest speed, while slightly varying its value does not change significantly the convergence time.

The number of iterations (epochs)

The number of iterations (or epochs) corresponds to each time the model parameters (weights and biases) are adjusted. The learning time of the model is often reported as the number of iterations the model needed to minimize the objective function. When estimating the ANN model on the full sample, 7,573 iterations are required to reach convergence, while in the out-of-sample forecasting with rolling window size of 800 observations, the number of iterations varies between 1,191 to 392,607 with a mean at 9,945.\textsuperscript{33}

\textsuperscript{32} For more details on the learning algorithm, the reader can refer to Riedmiller and Braun (1993).

\textsuperscript{33} Riedmiller and Braun (1993) suggest that a number of iterations lower to 15,000 is acceptable with the resilient backpropagation algorithm. However, compared to their analysis, we use a higher number of input variables which increases the computational costs of the ANN training.
Threshold of the partial derivative

The threshold value corresponds to the change in MSE during an iteration that stops the optimization process and takes the last estimated bias vector and weight matrix as the optimal. We define the threshold as 0.005 (0.5%) as a compromise between computation time versus model accuracy.

4.3.4. Linear ARFIMA model

The long-memory autoregressive fractionally integrated moving average process was first developed by (Granger and Joyeux 1980; Hosking 1981). Oomen (2001) and Andersen et al. (2003) proposed to model the realized volatility using the univariate ARFIMA. We will only use the ARFIMA model as a benchmark in the forecast evaluation analysis as a way to assess the robustness of the HAR model to capture long-memory (Corsi 2009). If we assume that the log($\hat{I}^d_t$) follows an ARFIMA(p,d,q) process, we have the following equation,

$$\alpha(L)(1 - L)^d \left( \log(\hat{I}^d_t) - \mu \right) = \beta(L) \epsilon_t$$

Where $\alpha(L)$ is a polynomial of order p, $\beta(L)$ is a polynomial of order q, and $\epsilon_t$ is a i.i.d. residual term. The ARFIMA model is stationary under the assumption that the roots of the $\alpha(L)$ and $\beta(L)$ lie outside the unit circle. The difference parameter d is positive and strictly lower than $1/2$. The AR and MA orders are chosen using the Akaike Information Criteria (AIC). This fractional parameter is estimated using the Geweke and Porter-Hudak (1983).

4.4. Empirical application

4.4.1. Data description

We employ corn futures transactions prices time-stamped to the nearest second from the Chicago Mercantile Exchange (CME) Group’s Best Bid and Offer (BBO) database from January 2, 2009
to April 31, 2017. We use day trading session hours,\textsuperscript{34} log transactions price returns to compute our estimates of integrated volatility based on equation (4.1).\textsuperscript{35} The corn futures contracts consist in five delivery months: March, May, July, September and December. We roll from the nearby to the next deferred contract when the trading volume of the deferred contract is higher than the nearby contract. This usually occurs three calendar weeks before the date of contract expiration.

Our intraday data cleaning procedure follows Barndorff-Nielsen et al. (2009). In addition, we exclude five limit price move days during which volatility is artificially reduced and, as a result, the log-IV cannot be computed for those days. After data pre-processing, our sample period length results in 2,129 trading days.

\textit{4.4.2. Corn futures volatility modeling}

Figure 4.2 depicts the temporal evolution of daily log-prices. Figure 4.3 shows the estimated daily log of the integrated volatility computed using equation (4.1). Note that we use the log($\widehat{IV}_t^d$) instead of the $\widehat{IV}_t^d$ to have an estimate of the volatility more symmetrically distributed and with lower kurtosis, a common approach in the literature (Fernandes et al. 2014) (Table 4.1). The daily log($\widehat{IV}_t^d$) presents a strong seasonal pattern with higher volatility in the summer. From January 2013, the integrated volatility series shows more frequent jumps related to USDA reports released during normal trading hours (Adjemian and Irwin 2018; Couleau et al. 2018). In Figure 4.4, we

\textsuperscript{34} These vary during our sample period as follows, before May 21, 2012, from 9:30 to 13:15; May 21-December 31, 2012: from 7:30 to 14:00; January 2, - April 5, 2013, from 9:30 to 14:00; since April 8, 2013, from 8:30 to 13:15, and since July 6, 2015, from 8:30 to 13:20.

\textsuperscript{35} Note that we estimate the $\widehat{IV}_t^d$ using equation (4.1) which requires us to define $\theta$ in order to determine the autocorrelation correction parameter $K$. As mentioned, we use $\theta = 0.4$ as in Couleau et al. (2018).
plot the autocorrelation function (ACF) of the $\log(\hat{V}_t^d)$ which suggest that the estimated series is highly persistent.

To evaluate the potential nonlinearity in the $\log(\hat{V}_t^d)$ series, we compute the linearity in mean test developed by Teräsvirta, Lin and Granger (1993) using five lags. This test not only captures structural changes, but also smoother nonlinearities in a time series; in other words, it tests for nonlinearities of various forms (Lee, White and Granger 1993). This test relies on a Taylor series expansion of a sigmoid function as the activation function to arrive to a suitable test statistic. The p-value is reported in Table 4.1 and indicates the null hypothesis of linearity in mean is rejected at the 5% significance level. We also conduct unit root tests. The ADF and PP unit root test p-values indicate the daily percentage returns and $\log(\hat{V}_t^d)$ are stationary over our sample period at the 5% significance level.

Table 4.2 reports the results of the estimation of the log-linear HAR models (Equations 4.4 and 4.5). The standard errors suggest that all coefficient estimates representing $\log(\hat{V}_t)$ at different time horizons are highly significant. Corsi (2009) points out that the estimation of the weekly and monthly volatilities contains more information on the volatility process as they are aggregated over longer horizons. This would explain the higher weights found for the weekly and monthly variables in Table 4.2. Ljung-box tests for serial correlation of order 10 or 20 for standardized residuals (squared and not squared) fail to reject the null of no serial correlation. In the fourth and fifth columns in Table 4.2, dummies for USDA announcement days are included. The dummy

---

36 We included the lagged log volume in the HAR and the HAR with dummy for announcement days and found it to be non-significant at 5% significance level.
variable for announcements is positive and significant. However, the parameter associated with this dummy decreases slightly when seasonal terms are included.

Adding seasonal terms increases slightly the lagged monthly log IV parameters, while the lagged daily and weekly coefficients decline, thus reducing the relevance of the short-run in favor of mid- and long-run effects. This result contrasts with Wang (2014) who find a reduction in the long-memory parameter after including seasonal components in their FIGARCH model. Differences in results might be related to how the FIGARCH model captures short-run dynamics compared to the HAR model. Nonetheless, adding the dummy variable for announcement days has the opposite effect; it decreases the value of the estimate of the monthly component (compare columns 2 and 4) while it increases those of the daily and weekly components. This finding is consistent with the literature that finds a lower long-memory parameter after accounting for structural breaks (Choi et al. 2010; Granger and Hyung 2004; Wang 2014).³⁷

We now move to the neural network specifications. First, we estimate the ANN model with a flexible lags structure (22 input nodes), as suggested by Hillebrand and Medeiros (2010). For comparison, we also specify a HAR-ANN model where we keep HAR lag structure with the three frequency components (daily, weekly, and monthly). Specifying the ANN requires to choose the number of hidden layers and the units in each layer. We consider between 1 and 8 layers and between 1 and 8 units in each layer and then select the alternative yielding the minimum RMSE. We find that two hidden layers with 2 units in the first layer and 4 hidden units in the second layer

³⁷ We tried lags 44 and 66 in the HAR model specifications and the parameters were not significant. This result indicates that there is no predictive information in aggregated volatility beyond one month. Audrino and Knaus (2016) also find that lags beyond 22 days are rarely significant for nine stocks of the S&P500 index. As a result, we set the maximum lags to 22 in further model evaluations.
perform best (lowest RMSE, i.e. 0.169). Such a specification indicates the presence of complex nonlinearities in the series that cannot be captured by a linear model. The estimated parameters of the ANN models are not presented as they do not have a direct interpretation and cannot be compared to the OLS parameters. As a result, we focus on the forecasted values.

The next subsection evaluates in-sample fit and out-of-sample forecasting performance of the realized volatility-based models (HARs and ANNs models).

4.4.3. In-sample fit and out-of-sample forecasting

To evaluate the forecasting ability of the different models, we first assess in-sample fit accuracy and then the out-of-sample forecasting performance of the models. The in-sample approach compares observed values (used in the estimation) to fitted values, while the out-of-sample predictive approach compares observed values (that were not used in the estimation) to model predicted values. Because the in-sample period has to be representative of the population to be compared to the out-of-sample analysis, we follow the approach by Fernandes et al. (2014).

First, we estimate the model on the full sample and provide the fitted values. We report the forecast statistics which are easy to understand and related to the needs of the decision-makers, i.e., root mean square error (RMSE), the mean absolute error (MAE), and the root mean square percentage error (RMSPE). In contrast to the MAE, measures based on the squared error, such as the RMSE, strongly penalize large errors which is important for practitioners as they might have costly consequences.

---

38 However, the weights parameters of the ANN model estimated in-sample are presented for completeness in appendix 4.8.2.

39 There are different ways to define the in-sample and out-of-sample sets (see the summary of the literature on ANN and volatility forecasting in appendix 4.8.1.).
**In-sample evaluation**

In-sample results are presented in Table 4.3. The smallest RMSE, MAE and RMSPE are indicated in bold. The ANN model has the smallest RMSE, MAE, and RMSPE, indicating ANN outperforms the linear HAR models for all measures of forecast accuracies. Consistent with Corsi (2009) and according to in-sample RMSE and MAE, the HAR model performs slightly better than the ARFIMA model, but the forecast accuracy measures are very close. Among the linear HAR models, the S-HAR-X has the best performance suggesting that adding nonlinear features improves the linear HAR model fit.

Amongst the ANN models, for comparison purposes, we evaluate the in-sample performance of the ANN-X, the HAR-ANN and the S-ANN models. Surprisingly, the ANN-X has the worst in-sample performance suggesting the USDA announcement dummies worsen the nonlinear ANN model specification in fitting the data. The HAR-ANN model performs worse than the ANN model. We attribute the improved performance of the ANN relative to the HAR-ANN model to the higher lag structure flexibility of the ANN specification. We explore the relevance of allowing for time-varying lags through the flexible lag structure of the ANN model. Given the difficulties in interpreting ANN parameters (Paliwal and Kumar 2009), we follow Olden, Joy and Death (2004) and measure “variable importance” as the product of the raw input-hidden and output-hidden connection weights between each input and output neuron and add the product across all hidden neurons. In figure 4.5 (bottom panel), we present a ‘heatmap’ of this measure. The x-axis represents the dates from 2012-03-15 to 2017-05-31, and the y-axis represents the ANN input variables (lagged log($\hat{\text{IV}}_t^d$) for lags from 1 to 22). The red indicates higher variable importance while the light yellow indicates lower variable importance. From figure 4.5, we appreciate a time-varying lag structure, with lags 1-2, 5-6, 9-10, 13-16, 19-22 becoming more relevant after 2014,
while lag relevance is more homogeneous before 2014. The fact that the lags selected by the ANN model and their intensity vary over time, helps understanding the better performance of the more flexible ANN over the HAR-ANN model. This finding corroborates the conclusions by Audrino, Huang and Okhrin (2018) that flexible lag structure works better in an unstable environment.

The S-ANN worse performance relative to ANN is attributed to the fact that ANN already takes into account seasonality. The degree to which ANN models are able to capture seasonality has been the subject of debate. While some studies find that ANN models perform better on deseasonalized data (Nelson et al. 1999), others such as Sharda and Patil (1992) find that the seasonality of time series does not affect the performance of the ANN models as these models are able to capture it implicitly. Additionally, Franses and Draisma (1997) investigate how well ANN can capture the seasonal pattern in macroeconomic time series and find these models are useful for seasonal pattern recognition. More recently, neural networks are used with success in atmospheric science to forecast periodic patterns of air quality (Kolehmainen, Martikainen and Ruuskanen 2001; Grivas and Chaloulakou 2006).

**Out-of-sample forecasts evaluation**

Next, we turn to the out-of-sample forecast ability of the models considered. The forecast error corresponds to \( e_{j+h} = \log(I\hat{V}_{t+h}) - \log(I\hat{V}_{t+h|t}) \) where \( \log(I\hat{V}_{t+h|t}) \) denotes the out-of-sample forecast of \( \log(I\hat{V}_{t+h}) \) based on \( F_t \), the information set available at time \( t \). We produce the out-of-sample forecasts following the widely used fixed rolling window approach (Fernandes et al. 2014; Baruník and Krehlík 2016; Liu, Pantelous and von Mettenheim 2018). We use two windows with sizes 800 (38% of the total sample), and 1,200 (56% of the total sample) daily observations to explore how results vary by window size. The starting periods associated to each window are 2009-01-02 to 2012-03-14 (for \( W = 800 \)), and 2009-01-02 to 2013-10-14 (\( W = 1,200 \)). For each
sample, the models are estimated, and the \( h \)-step ahead out-of-sample forecasts are produced. One-step ahead forecast is equivalent to the fitted value of the estimated model over the full or partitioned sample, while multi-step ahead forecasts are computed based on the iterated forecasting technique (by re-inserting the estimated output variable back into the model iteratively). Then, the window is rolled one observation ahead and the process is repeated again until the end of the sample. For the out-of-sample forecasts, we compute the three forecasting evaluation measures as:

\[
RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^{N} e_{j+h}^2},
\]

\[
MAE = \frac{1}{N} \sum_{j=1}^{N} |e_{j+h}|, \text{ and}
\]

\[
RMSPE = \sqrt{\frac{1}{N} \sum_{j=1}^{N} \frac{e_{j+h}^2}{\log(\hat{V}_{t+h|x})^2}}.
\]

We present the out-of-sample results in Table 4.4, with two general patterns emerging. First, with a few exceptions, the three forecast accuracy measures tend to decrease as the window size increases. This trend indicates that increasing the number of observations tends to improve the models’ estimation or “training”. For the ANN-X, however, smaller windows might permit a more flexible response and hence more accurate forecast in the presence of announcement dummies. In the same vein, in the absence of dramatic structural breaks, as the number of observations increases, the forecast performance is more precise. Second, forecast accuracy declines at longer horizons as reflected in higher RMSE, MAE, and RMSPE values.

Among the HAR specifications, the S-HAR model is often preferred, independent of the criteria used on medium and long-horizon, while the HAR-X or S-HAR-X are preferred on short-horizon, depending on the window size. This result confirms that accounting for announcement days through dummies decreases the performance of the models on long-horizon. The ANN model outperforms the HAR models at all horizons and all window sizes. A result in line with McAleer
and Medeiros (2011) who find that flexible ANN and HAR-ANN models approximate well the behavior of the realized volatility compared to their linear counterpart. Baruník and Krehlík (2016) also find a superior forecasting performance of the ANN and the HAR-ANN over the HAR model. The improved performance of the ANN model is often attributed to the learning from repeated patterns, which is the case for seasonality or USDA public announcements. Among the ANN models, the ANN-X performs worse than the ANN except in terms of percentage error at $h = 5$. This result suggests that adding dummies to capture events limits the flexibility of the nonlinear models’ forecast accuracy.

In summary, accounting for nonlinearities is important and can largely improve out-of-sample forecasts throughout the ANN model, but using dummies could lead to large forecast errors in volatility. This finding indicates that nonlinearities might be time-varying and can take various forms that cannot be properly captured through standard dummy variables that represent an abrupt and discontinuous change in volatility.

4.4.4. Testing models’ forecasting performance

To formally compare the forecasting performance of the different models, we first use the Modified Diebold-Mariano (MDM) test (Harvey, Leybourne and Newbold 1997) which conducts a pairwise comparison of the loss function derived from each model. MDM is computed as

$$ MDM = \frac{H - 1}{\sqrt{H} \sum_{t=1}^{H} (d_t - \bar{d})^2} $$

---

40 The MDM test accounts for the fact that DM test could be oversized for forecasting horizons higher than 1, that is $h \geq 2$. Harvey, Leybourne and Newbold (1997) propose the MDM test in order to alleviate this problem.
where \( d_t = g(e_{t,1}) - g(e_{t,2}) \), being \( g(e_{t,i}), i = 1, 2 \) the loss function for each of the models, which we specify as the MSE\(^{41}\), \( H \) is the number of out-of-sample forecasts, and \( \bar{d} \) is the average of the difference. The test statistic is computed for 1-, 5-, and 22-step ahead forecasts. The null hypothesis is defined as \( E(d_t) = 0 \), that is the two models have equivalent forecasting performance.

We present the results of the MDM test in Table 4.5 for two window sizes (\( W = 800 \) and \( W = 1,200 \)). The bold values refer to p-values inferior to 0.05, that is the null hypothesis that the two models have equivalent forecast performance is rejected at 5% significance level. The results indicate that among the HAR models, it is difficult to disentangle difference in predictive accuracy between the ARFIMA, HAR, and S-HAR depending on the horizon. However, the HAR-X and S-HAR-X appear to have better predictive accuracy than those models. Among the ANN models, the ANN has better predictive accuracy than the ANN-X model at all horizons. These results are reinforced when increasing the window size at \( W = 1,200 \). The MDM test results are somewhat robust to various window sizes. In sum, accounting for the seasonal component in the HAR model does not improve forecasting performance but accounting for information shocks related to public announcement days improves forecasting performance of the linear model in the short-horizon. Most importantly, taking into account the nonlinearities of unknown forms through the ANN model provides superior forecasting performance than the linear models with exogenous shocks in all horizons.

\(^{41}\) This loss function is often preferred as it accounts for large errors compared to the MAE. Additionally, RMSE and MSE are reaching the same conclusion since both uses mean square error. The squared root is a monotonic function, so the ranking of the model’s results is similar with the RMSE or the MSE.
Second, we employ the model confidence set (MCS) by Hansen et al. (2011). Hansen et al. (2011) go a step further to testing pairwise models and propose a selection procedure. Intuitively, the MCS procedure consists in selecting the set of ‘best’ models ($M^*$) given a collection of candidate forecast models ($M_0$), where ‘best’ is defined by an evaluation forecast measure selected by the modeler. The procedure sequentially identifies $M^* \subset M_0$ given a confidence interval $\alpha$. In practice, an equivalence test is applied to the collection of models in $M_0$. If the null hypothesis is rejected, there is evidence that the objects in $M_0$ are not equally “good” and an elimination rule is used to remove from $M_0$ the model with poor sample performance. This procedure is repeated until the null is accepted and the MCS is defined by a set of “surviving” models. We define a set of models $M_0 = \{1, \ldots, m_0\}$ and sequentially\(^{42}\) test the null hypothesis of the equal prediction accuracy, i.e.

$$H_0: E(d_{ij,t}) = 0 \ \forall i, j \in M_0$$

where $d_{ij,t} = g(e_{i,t}) - g(e_{j,t})$ is the loss differential between models $i$ and $j$ in the collection of models where $g(.)$ is the MSE function. Following Hansen et al. (2011), we use two types of statistics for testing the null hypothesis, the range statistics, $T_R$, and the semi-quadratic statistics $T_{SQ}$, both relying on the following t-statistic:

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\text{var}(\bar{d}_{ij})}}$$

for $i, j \in M_0$, with $\bar{d}_{ij} = \frac{1}{H} \sum_{t=1}^{H} d_{ij,t}$. The t-statistic, $t_{ij}$, provides scaled information on the average difference in the point forecast quality of models $i$ and $j$. Additionally, $\text{var}(\bar{d}_{ij})$ is an estimate of $\text{var}(d_{ij})$, obtained by using the block bootstrap of Gonçalves and White (2005) as proposed by Hansen et al. (2011).

\(^{42}\)Hansen et al. (2011) use the approach of the stepdown Holm adjusted p-values to account for multiple testing bias (see discussion in their paper top of page 474).
This approach is usually used when the number of models in the set is high. The range statistics, $T_R$ and the semi-quadratic statistics, $T_{SQ}$ are given by,

$$T_R = \max_{i,j \in M_0} |t_{ij}| = \max_{i,j \in M_0} \frac{|d_{ij}|}{\sqrt{\text{var}(d_{ij})}}$$

$$T_{SQ} = \max_{i,j \in M_0} t_{ij}^2 = \sum_{i,j \in M_0} \frac{d_{ij}^2}{\text{var}(d_{ij})}.$$

Finally, the MCS procedure assigns p-value, $\hat{p}_i$, to each model $i$ in the initial set $M_0$. The resulting optimal set of model(s) is denoted by $\hat{M}_{1-\alpha}$ if and only if $\hat{p}_i \geq \alpha$.

We present MCS p-values per statistic in Table 4.6. The ANN model is the unique model selected for $h = 1$, $h = 5$ and $h = 22$ and for $W = 800$ and $W = 1,200$, confirming the results from the MDM test. Hansen, Lunde, and Nason (2011) argue that the size of the set of models selected by the MCS informs us on the accuracy of the data involved in the forecasts. Indeed, less informative data will result in difficulty of the MCS to distinguish the ‘best’ model while more informative data will lead to the ‘best’ model being selected. The selection of a unique model suggests that the log-IV is an informative measure of volatility.
4.5. Concluding remarks

Efficient volatility forecasts can improve social welfare by adjusting economic decisions (Kenyon, Jones and Mcguirk 1993). However, forecasting volatility in the presence of long-memory can be complicated by heterogeneous volatility expectations from a diversity of traders. Depending on their information needs, market participants are likely to be interested in futures price volatility at different horizons. Nonlinearities such as seasonality or public information shocks (USDA public announcements) also influence volatility persistence, complicating volatility forecasts in the corn futures market. Recent applications of artificial neural networks suggest that this model is particularly suited to capture unknown forms of nonlinearities.

To assess how nonlinearities affect volatility persistence forecasting, we forecast corn futures price volatility through a nonlinear heterogeneous autoregressive (HAR) model approximated through an artificial neural network (ANN) using intraday prices and compare it to linear model specifications of the realized volatility.

Using intraday transaction prices from 2009 to 2017, we document a long-memory process in the realized volatility through the linear HAR model. Accounting for the seasonal component in the HAR model does not improve forecasting performance but allowing for information shocks related to public announcement days does in the short-horizon. Our main finding is that taking into account nonlinearities of unknown forms through the ANN model provides superior forecasting performance than any of the linear models in all forecasting horizons. Out-of-sample forecast analysis and forecasting evaluation tests are consistent with the selection of the ANN model as the one having the best forecasting performance. This is especially true at long forecasting horizons. This result is also consistent with findings by McAleer and Medeiros (2011) and Baruník and Krehlík (2016) who find a superior forecasting performance of the flexible ANN over the HAR
model. To further understand the good performance of the ANN model, we analyze the time-varying dynamics of the lags structure by comparing the ANN with the HAR-ANN. A flexible lags structure is preferred to the fixed lags structure (1, 5, 22). This finding supports the conclusions by Audrino, Huang and Okhrin (2018) that fixed lag structure is not always accurate in an unstable environment.

Future research might explore the benefits of forecast combinations (linear with nonlinear models) for situations where nonlinearities are less obvious (such as in live cattle futures market for instance). Bates and Granger (1969) were the first to show that forecast combinations can lead to improved forecast accuracy. Since then a number of different combination techniques have been proposed but a forecast combination puzzle have emerged as well. The forecast combination puzzle consists in that combining forecast with equal weights performs much better than combining forecast with optimal weights constructed to have superior forecasting accuracy using the MSE measure (Smith and Wallis 2009). The MCS approach can also be used to combine forecasts from models selected in the optimal set if more than one model is selected (Samuels and Sekkel 2017). Additionally, the agricultural economic literature provides evidence that implied volatility embedded in corn option prices can lead to better forecasts of realized price volatility than those based on historical volatility information (Egelkraut et al. 2007; Giot 2003). However, Hamid and Iqbal (2004) and Martens and Zein (2004) show that realized volatility forecasts can provide new information compared to the implied volatility estimates. Revisiting this research question by comparing implied volatility with volatility models based on realized volatility in agricultural futures markets in the presence of long-memory and nonlinearities would be an extension of this research. Finally, the ANN model forecasting performance could be investigated in other economic situations such as predicting the directional changes in volatility.
4.6. Figures and tables

Figure 4.2. Nearby corn futures transaction log-prices from January 2, 2009 to May 31, 2017
Figure 4.3. The log-realized volatility on noise-filtered transaction prices, $\log(IV_t^d)$, from January 2, 2009 to May 31, 2017

Notes: The variable $\log(IV_t^d)$ is estimated using the log of the IV in equation (4.1).

Figure 4.4. Autocorrelation Function (ACF) of the $\log(IV_t^d)$

Notes: The variable $\log(IV_t^d)$ is the log of the IV in equation (4.1) and $\theta = 0.4$. The ACF maximum lag is 800. The confidence intervals at 5% significance level are represented in dashed blue horizontal lines.
Figure 4.5. Daily log($\hat{\bar{W}}_t^d$) from 2012-03-15 to 2017-05-31 (top panel) and heatmap of the importance of the weights’ parameters in the ANN for the out-of-sample analysis (with $W = 800$) (bottom panel)
Figure 4.6. Number of iterations for the training of the ANN model in the out-of-sample analysis (with $W = 800$)

Note: The vertical dash grey line corresponds to April 5th, 2016 after which the number of iterations increases drastically.
Table 4.1. Descriptive statistics of the \( \log(\hat{IV}_t^d) \), 01/02/2009-05/31/2017

<table>
<thead>
<tr>
<th></th>
<th>( \log(\hat{IV}_t^d) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>-5.1503</td>
</tr>
<tr>
<td>Max</td>
<td>-2.8333</td>
</tr>
<tr>
<td>Mean</td>
<td>-4.3977</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.3207</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.580</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.617</td>
</tr>
<tr>
<td>Jarque-Bera test p-value</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>ADF p-value</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>PP p-value</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Terasvirta et al. (1993) test p-value</td>
<td>7.60e-12</td>
</tr>
</tbody>
</table>

Notes: Squared interday returns are constructed using \( 100 \ast \log(\frac{P_t}{P_{t-1}})^2 \), and \( \log(\hat{IV}_t^d) \) corresponds to the log of the IV in equation 4.1. Five limit-price move days with limit moves spanning almost all day were excluded as their variance was close to zero. Lags of the augmented Dickey-Fuller (ADF) test are selected using Bayesian Information Criterion (BIC). Phillips and Perron (PP) test uses the Newey-West standard errors to allow for autocorrelation. The number of lags to estimate the asymptotic variance is set at 12(T/100)^1/4. Teräsvirta et al. (1993)’s test p-value is obtained for five lags. The p-value is still lower than 0.01 for further lags (10 and 20).
Table 4.2. Coefficient estimates of the log-linear HAR models’ estimation.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.423***</td>
<td>-0.519***</td>
<td>-0.488***</td>
<td>-0.579***</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.12)</td>
<td>(0.09)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>Lag(IV_{t-1d})</td>
<td>0.205***</td>
<td>0.201***</td>
<td>0.221***</td>
<td>0.218***</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Lag(IV_{t-5d})</td>
<td>0.321***</td>
<td>0.290***</td>
<td>0.349***</td>
<td>0.323***</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Lag(IV_{t-22d})</td>
<td>0.378**</td>
<td>0.390***</td>
<td>0.325***</td>
<td>0.334***</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>$D_t^\text{usda}$</td>
<td>-</td>
<td>-</td>
<td>0.526***</td>
<td>0.525***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-</td>
<td>-0.026**</td>
<td>-</td>
<td>-0.022**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>-</td>
<td>0.005</td>
<td>-</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-</td>
<td>-0.003</td>
<td>-</td>
<td>-0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>-</td>
<td>-0.009</td>
<td>-</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\gamma_3$</td>
<td>-</td>
<td>-0.011</td>
<td>-</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\delta_3$</td>
<td>-</td>
<td>0.008</td>
<td>-</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\gamma_4$</td>
<td>-</td>
<td>0.009</td>
<td>-</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\delta_4$</td>
<td>-</td>
<td>0.007</td>
<td>-</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.440</td>
<td>0.442</td>
<td>0.560</td>
<td>0.562</td>
</tr>
</tbody>
</table>

**Standardized residuals, p-values**

<table>
<thead>
<tr>
<th></th>
<th>Ljung-Box test, lags 10</th>
<th>Ljung-Box test, lags 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ljung-Box test, lags 10</td>
<td>0.40</td>
<td>0.37</td>
</tr>
<tr>
<td>Ljung-Box test, lags 20</td>
<td>0.47</td>
<td>0.41</td>
</tr>
</tbody>
</table>

**Squared standardized residuals, p-values**

<table>
<thead>
<tr>
<th></th>
<th>Ljung-Box test, lags 10</th>
<th>Ljung-Box test, lags 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ljung-Box test, lags 10</td>
<td>0.76</td>
<td>0.96</td>
</tr>
<tr>
<td>Ljung-Box test, lags 20</td>
<td>0.84</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. Significance levels are represented by symbols *, **, and *** for 10%, 5%, and 1% respectively. The $\log \hat{IV}$ refers to the logarithm form of the estimated annualized integrated volatility.
Table 4.2. (Continued)

(equation 4.1). $D_{e}^{usd}$ equals to 1 on announcement days and 0 otherwise. Terms $\gamma_1$ to $\delta_4$ are seasonal trigonometric variables. The table contains the HAR, S-HAR, HAR-X and S-HAR-X model specifications, where the prefix “S-” corresponds to seasonally adjusted models and the suffix “-X” corresponds to models containing exogenous variables which corresponds to the dummies for announcement days (equals 1 if there is a USDA report release, equals 0 otherwise).
Table 4.3. In-sample evaluation on the full sample, 01/02/2009 - 05/31/2017

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
<th>MAE</th>
<th>RMSPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARFIMA(1,d,1)</td>
<td>0.234</td>
<td>0.162</td>
<td>0.052</td>
</tr>
<tr>
<td>HAR</td>
<td>0.233</td>
<td>0.160</td>
<td>0.053</td>
</tr>
<tr>
<td>S-HAR</td>
<td>0.232</td>
<td>0.160</td>
<td>0.052</td>
</tr>
<tr>
<td>HAR-X</td>
<td>0.195</td>
<td>0.144</td>
<td>0.046</td>
</tr>
<tr>
<td>S-HAR-X</td>
<td>0.194</td>
<td>0.143</td>
<td>0.046</td>
</tr>
<tr>
<td>HAR-ANN</td>
<td>0.228</td>
<td>0.158</td>
<td>0.052</td>
</tr>
<tr>
<td>ANN</td>
<td><strong>0.172</strong></td>
<td><strong>0.113</strong></td>
<td><strong>0.040</strong></td>
</tr>
<tr>
<td>S-ANN</td>
<td>0.192</td>
<td>0.140</td>
<td>0.045</td>
</tr>
<tr>
<td>ANN-X</td>
<td>0.275</td>
<td>0.213</td>
<td>0.084</td>
</tr>
</tbody>
</table>

Notes: The prefix “S-” refers to models augmented for seasonality, and the suffix “-X” refers to the models with dummies for USDA announcement days (equals 1 if there is a USDA crop report is released, equals 0 otherwise). Two hidden layers are used for the ANN and ANN-X, with 2:4 nodes (number of neurons in layer 1: number of neurons in layer 2) which combinations has the lowest RMSE (for numbers of neurons going from 1 to 8). The forecasting evaluation measures are the root mean square errors (RMSE) (equation 4.8), the mean absolute error (MAE) (equation 4.9) and the root mean square percentage error (RMSPE) (equations 4.10 and 4.11). The numbers in bold refer to the smallest.
Table 4.4. Out-of-sample forecasting using rolling approach for window sizes $W = 800$ and $1,200$.

<table>
<thead>
<tr>
<th></th>
<th>$W = 800$</th>
<th></th>
<th>$W = 1200$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>MAE</td>
<td>RMSPE</td>
<td>RMSE</td>
<td>MAE</td>
<td>RMSPE</td>
</tr>
<tr>
<td>$h = 1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARFIMA(1,$d$,1)</td>
<td>0.253</td>
<td>0.167</td>
<td>0.056</td>
<td>0.252</td>
<td>0.166</td>
<td>0.056</td>
</tr>
<tr>
<td>HAR</td>
<td>0.261</td>
<td>0.179</td>
<td>0.058</td>
<td>0.252</td>
<td>0.166</td>
<td>0.056</td>
</tr>
<tr>
<td>S-HAR</td>
<td>0.261</td>
<td>0.182</td>
<td>0.058</td>
<td>0.251</td>
<td>0.165</td>
<td>0.056</td>
</tr>
<tr>
<td>HAR-X</td>
<td>0.207</td>
<td>0.152</td>
<td>0.047</td>
<td>0.191</td>
<td>0.138</td>
<td>0.044</td>
</tr>
<tr>
<td>S-HAR-X</td>
<td>0.207</td>
<td>0.154</td>
<td>0.048</td>
<td>0.191</td>
<td>0.137</td>
<td>0.044</td>
</tr>
<tr>
<td>ANN</td>
<td><strong>0.189</strong></td>
<td><strong>0.129</strong></td>
<td><strong>0.043</strong></td>
<td><strong>0.182</strong></td>
<td><strong>0.123</strong></td>
<td><strong>0.042</strong></td>
</tr>
<tr>
<td>ANN-X</td>
<td>0.213</td>
<td>0.154</td>
<td>0.044</td>
<td>0.263</td>
<td>0.199</td>
<td>0.044</td>
</tr>
<tr>
<td>$h = 5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARFIMA(1,$d$,1)</td>
<td>0.272</td>
<td>0.194</td>
<td>0.061</td>
<td>0.274</td>
<td>0.194</td>
<td>0.061</td>
</tr>
<tr>
<td>HAR</td>
<td>0.272</td>
<td>0.193</td>
<td>0.061</td>
<td>0.273</td>
<td>0.193</td>
<td>0.061</td>
</tr>
<tr>
<td>S-HAR</td>
<td>0.271</td>
<td>0.194</td>
<td>0.060</td>
<td>0.269</td>
<td>0.187</td>
<td>0.059</td>
</tr>
<tr>
<td>HAR-X</td>
<td>0.341</td>
<td>0.225</td>
<td>0.083</td>
<td>0.360</td>
<td>0.233</td>
<td>0.089</td>
</tr>
<tr>
<td>S-HAR-X</td>
<td>0.332</td>
<td>0.223</td>
<td>0.079</td>
<td>0.347</td>
<td>0.226</td>
<td>0.084</td>
</tr>
<tr>
<td>ANN</td>
<td><strong>0.261</strong></td>
<td><strong>0.186</strong></td>
<td><strong>0.059</strong></td>
<td><strong>0.246</strong></td>
<td><strong>0.174</strong></td>
<td><strong>0.055</strong></td>
</tr>
<tr>
<td>ANN-X</td>
<td>0.328</td>
<td>0.252</td>
<td><strong>0.059</strong></td>
<td>0.386</td>
<td>0.313</td>
<td>0.057</td>
</tr>
<tr>
<td>$h = 22$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARFIMA(1,$d$,1)</td>
<td>0.302</td>
<td>0.229</td>
<td>0.068</td>
<td>0.293</td>
<td>0.218</td>
<td>0.065</td>
</tr>
<tr>
<td>HAR</td>
<td>0.304</td>
<td>0.228</td>
<td>0.068</td>
<td>0.295</td>
<td>0.216</td>
<td>0.065</td>
</tr>
<tr>
<td>S-HAR</td>
<td>0.300</td>
<td>0.225</td>
<td>0.068</td>
<td>0.289</td>
<td>0.211</td>
<td>0.064</td>
</tr>
<tr>
<td>HAR-X</td>
<td>0.376</td>
<td>0.257</td>
<td>0.096</td>
<td>0.415</td>
<td>0.259</td>
<td>0.117</td>
</tr>
<tr>
<td>S-HAR-X</td>
<td>0.360</td>
<td>0.249</td>
<td>0.090</td>
<td>0.394</td>
<td>0.249</td>
<td>0.107</td>
</tr>
<tr>
<td>ANN</td>
<td><strong>0.266</strong></td>
<td><strong>0.204</strong></td>
<td><strong>0.061</strong></td>
<td><strong>0.250</strong></td>
<td><strong>0.192</strong></td>
<td><strong>0.057</strong></td>
</tr>
</tbody>
</table>
Table 4.4. (Continued)

<table>
<thead>
<tr>
<th></th>
<th>ANN-X</th>
<th>0.379</th>
<th>0.305</th>
<th>0.062</th>
<th>0.452</th>
<th>0.375</th>
<th>0.059</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of out-of-sample forecasts for each horizon</td>
<td></td>
<td>1290</td>
<td></td>
<td></td>
<td></td>
<td>907</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The prefix “S-” refers to models augmented for seasonality, and the suffix “-X” refers to the models with dummies for announcement days (equals 1 if there is a USDA crop report released, equals 0 otherwise). The different window sizes ($W = 800$ and $1,200$) correspond to fixed window sizes used in the rolling approach for out-of-sample forecasting. The forecasted values produced by each model are point forecasts daily volatility at horizons, $h$, of one day ($h = 1$), 5 days ($h = 5$), and 22 days ($h = 22$). The forecasting evaluation measures are the root mean square errors (RMSE) (equation 4.8), the mean absolute error (MAE) (equation 4.9) and the root mean square percentage error (RMSPE) (equations 4.10 and 4.11). The numbers in bold refer to the smallest.
Table 4.5. Modified Diebold-Mariano (MDM) test results for 1-day, 5-day, and 22-day horizons forecasts for $\log \hat{V}_t$ for RV-based models.

**W = 800**

<table>
<thead>
<tr>
<th></th>
<th>HAR</th>
<th>S-HAR</th>
<th>HAR-X</th>
<th>S-HAR-X</th>
<th>ANN-X</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARFIMA</td>
<td>h = 1</td>
<td>0.8993</td>
<td>0.0591</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0016</td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td><strong>0.0433</strong></td>
<td>0.0551</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>0.6524</td>
<td>0.4311</td>
<td>0.0001</td>
<td><strong>0.0028</strong></td>
<td><strong>0.0001</strong></td>
</tr>
<tr>
<td>HAR</td>
<td>h = 1</td>
<td>-</td>
<td><strong>0.0103</strong></td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td>-</td>
<td>0.5004</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>-</td>
<td>0.4371</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0000</td>
</tr>
<tr>
<td>S-HAR</td>
<td>h = 1</td>
<td>-</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td>-</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>-</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>HAR-X</td>
<td>h = 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0625</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>0.0283</strong></td>
<td><strong>0.0164</strong></td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>0.0231</strong></td>
<td>0.0896</td>
</tr>
<tr>
<td>S-HAR-X</td>
<td>h = 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>0.003</strong></td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0061</td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0329</td>
</tr>
<tr>
<td>ANN-X</td>
<td>h = 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**W = 1200**

<table>
<thead>
<tr>
<th></th>
<th>HAR</th>
<th>S-HAR</th>
<th>HAR-X</th>
<th>S-HAR-X</th>
<th>ANN-X</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARFIMA</td>
<td>h = 1</td>
<td>0.828</td>
<td>0.6145</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td>0.4901</td>
<td><strong>0.0235</strong></td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>0.7739</td>
<td>0.4522</td>
<td><strong>0.0012</strong></td>
<td><strong>0.0153</strong></td>
<td>0.0000</td>
</tr>
<tr>
<td>HAR</td>
<td>h = 1</td>
<td>-</td>
<td>0.3641</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td>-</td>
<td><strong>0.0195</strong></td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>-</td>
<td>0.2354</td>
<td>0.0000</td>
<td>0.0010</td>
<td>0.0000</td>
</tr>
<tr>
<td>S-HAR</td>
<td>h = 1</td>
<td>-</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td>-</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>-</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>HAR-X</td>
<td>h = 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0266</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>0.0001</strong></td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>0.0230</strong></td>
<td>0.0005</td>
</tr>
<tr>
<td>S-HAR-X</td>
<td>h = 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0002</td>
</tr>
<tr>
<td>ANN-X</td>
<td>h = 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>h = 5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>h = 22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes: The p-values correspond to the Modified Diebold-Mariano test for the null hypothesis that the column and row models perform equally well in terms of mean square errors. The bold p-values correspond to p-values < 0.05.

W corresponds to the window size.
Table 4.6. Model Confidence Set (MCS) results for 1-day, 5-day, and 22-day horizons forecasts for realized volatility-based models (HARs, and ANNs)

<table>
<thead>
<tr>
<th></th>
<th>$h = 1$</th>
<th></th>
<th>$h = 5$</th>
<th></th>
<th>$h = 22$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_R$</td>
<td>$T_{SQ}$</td>
<td>$T_R$</td>
<td>$T_{SQ}$</td>
<td>$T_R$</td>
<td>$T_{SQ}$</td>
</tr>
<tr>
<td>$W = 800$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARFIMA</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0018</td>
<td>0.0032</td>
</tr>
<tr>
<td>HAR</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0026</td>
<td>0.0032</td>
</tr>
<tr>
<td>S-HAR</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0052</td>
<td>0.0032</td>
</tr>
<tr>
<td>HAR-X</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0052</td>
<td>0.0032</td>
</tr>
<tr>
<td>S-HAR-X</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0052</td>
<td>0.0032</td>
</tr>
<tr>
<td>ANN-X</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0052</td>
<td>0.0040</td>
</tr>
<tr>
<td>ANN</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

| $W = 1200$ |         |         |         |         |         |         |
| ARFIMA | 0.0000  | 0.0002 | 0.0000  | 0.0000  | 0.0000  | 0.0000 |
| HAR    | 0.0000  | 0.0002 | 0.0000  | 0.0000  | 0.0000  | 0.0000 |
| S-HAR  | 0.0000  | 0.0002 | 0.0000  | 0.0000  | 0.0000  | 0.0000 |
| HAR-X  | 0.0000  | 0.0002 | 0.0000  | 0.0000  | 0.0000  | 0.0000 |
| S-HAR-X| 0.0000  | 0.0002 | 0.0000  | 0.0000  | 0.0000  | 0.0000 |
| ANN-X  | 0.0000  | 0.0002 | 0.0000  | 0.0000  | 0.0000  | 0.0000 |
| ANN    | 0.9999  | 0.9999 | 0.9999  | 0.9999  | 0.9999  | 0.9999 |

Notes: We report the p-values from the MCS procedure for window sizes $W = 800$ and $W = 1,200$. In bold, the p-values $> 0.10$ meaning the models are included in the set $\mathcal{M}_{0.90}$. $T_R$ refers to the range statistics, and $T_{SQ}$ to the semi-quadratic statistics.
4.7. References


frequency volatility on large indices.” Quantitative Finance 18(5):737–748.


### 4.8. Supplementary Information

#### 4.8.1. Summary of relevant literature on artificial neural networks for volatility forecasting

Table 4.7. Summary of relevant literature on artificial neural networks and volatility forecasting

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Period</th>
<th>Sample Size</th>
<th>Market(s)</th>
<th>Series</th>
<th>Model(s)/Method(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamid and Iqbal</td>
<td>2004</td>
<td>February 1, 1984, to January 31, 1994</td>
<td>2531</td>
<td>S&amp;P 500 Index futures (and 20 explanatory variables)</td>
<td>Daily transaction prices</td>
<td>Volatility forecast from ANN model are compared to implied volatility from option prices and to a 55-day realized standard deviation (RSD) from daily log-prices</td>
</tr>
<tr>
<td>Baruník and Krehlík</td>
<td>2016</td>
<td>January 5, 2004 to through December 31, 2012</td>
<td>2231</td>
<td>NYSE crude oil, heating oil, and natural gas traded</td>
<td>Intraday transaction prices</td>
<td>ARFIMA, GARCH, HAR, ANN, HAR-ANN</td>
</tr>
<tr>
<td>Liu, Pantelous and von Mettenheim</td>
<td>2018</td>
<td>March, 20th, 1996, January, 2nd, 1996, and January, 30th, 2009 (for respective markets) and end in 2016</td>
<td>~20 years of data</td>
<td>SPY ETF, the VIX index, and the VXX ETN</td>
<td>Intraday prices</td>
<td>Model 1: linear (HAR-X), model 2: recurrent neural network (RNN), and model 3: Hybrid (RNN that uses the linear model estimates as an input)</td>
</tr>
<tr>
<td>Authors</td>
<td>Year</td>
<td>Structure of the ANN model</td>
<td>Nodes/Layers</td>
<td>Bias nodes?</td>
<td>Activation function</td>
<td>Data normalization</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------</td>
<td>-----------------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>---------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Donaldson and Kamstra</td>
<td>1997</td>
<td>3-layer structure</td>
<td>1:4:1</td>
<td>Since ANN is estimated with MLE, a single intercept is estimated</td>
<td>logistic function</td>
<td>Yes</td>
</tr>
<tr>
<td>Hamid and Iqbal</td>
<td>2004</td>
<td>3-layer structure</td>
<td>13:26:1</td>
<td>Not discussed</td>
<td>logistic function</td>
<td>Yes</td>
</tr>
<tr>
<td>Fernandes, Medeiros, Scharth</td>
<td>2014</td>
<td>3-layer structure</td>
<td>15:3:1</td>
<td>Not used</td>
<td>logistic function</td>
<td>not discussed</td>
</tr>
<tr>
<td>Baruník and Krehlík</td>
<td>2016</td>
<td>3-layer structure</td>
<td>between 7 and 15 neurons in the hidden layer</td>
<td>Not discussed</td>
<td>logistic function</td>
<td>not discussed</td>
</tr>
<tr>
<td>Liu, Pantelous and von Mettenheim</td>
<td>2018</td>
<td>3-layer structure</td>
<td>Input neurons (same number as in the HAR, 3), hidden neurons (use the formula $h = 2 \times \sqrt{(i \times o)}$, where $i$ and $o$ refer to the size of the input and output layer, and 1 output neuron</td>
<td>yes</td>
<td>logistic function</td>
<td>Yes</td>
</tr>
<tr>
<td>Authors</td>
<td>Year</td>
<td>Initial weights/parameters selection</td>
<td>In-sample sample size</td>
<td>Out-of-sample sample size, approach</td>
<td>Out-of-sample evaluation/Horizons forecast for out-of-sample</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------</td>
<td>--------------------------------------</td>
<td>-----------------------</td>
<td>-------------------------------------</td>
<td>----------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Donaldson and Kamstra</td>
<td>1997</td>
<td>not used</td>
<td>50% of sample size</td>
<td>50% of sample size/rolling approach</td>
<td>1-day ahead</td>
<td></td>
</tr>
<tr>
<td>Hamid and Iqbal</td>
<td>2004</td>
<td>no detail about the training of the neural network</td>
<td>500 observations</td>
<td>2031 observations</td>
<td>RMSE/MAE/3 horizons forecast: 15-, 35-, and 55-day ahead</td>
<td></td>
</tr>
<tr>
<td>Fernandes, Medeiros, Scharth</td>
<td>2014</td>
<td>-</td>
<td>Full sample</td>
<td>2500 rolling window to estimate the models, remaining of the sample for out-of-sample</td>
<td>Use seven forecasting error assessments/4 horizons forecast: 1-, 5-, 10-, and 22-day ahead</td>
<td></td>
</tr>
<tr>
<td>Baruník and Krehlík</td>
<td>2016</td>
<td>not discussed</td>
<td>600 observations for the in-sample fits</td>
<td>1631 observations to evaluate the out-of-sample forecasting performance/rolling approach</td>
<td>RMSE/MAE/ 1-, 5- and 10-day ahead horizons forecast</td>
<td></td>
</tr>
<tr>
<td>Liu, Pantelous and von Mettenheim</td>
<td>2018</td>
<td>not discussed</td>
<td>size varies between 22 and 504 observations</td>
<td>to train the RNN and produce out-of-sample forecast on the test sample</td>
<td>RMSE/MAE/MAPE/ 3 forecast horizons: 1-day, 2-day, and 5-day</td>
<td></td>
</tr>
</tbody>
</table>
4.8.2. Statistical appendix for the ANN model.

The 63 parameters obtained from the most often selected model, the ANN model by the Model Confidence Set approach are presented below. Table 4.8 presents the ANN model’s optimal parameters from the in-sample analysis.

Table 4.8. Optimal parameters from the ANN neural network model.

<table>
<thead>
<tr>
<th></th>
<th>Hidden layer1.node1</th>
<th>Hidden layer1.node2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias node 1</td>
<td>-3.084</td>
<td>0.457</td>
</tr>
<tr>
<td>log(IV_{t-1})</td>
<td>0.250</td>
<td>6.681</td>
</tr>
<tr>
<td>log(IV_{t-2})</td>
<td>3.066</td>
<td>-5.626</td>
</tr>
<tr>
<td>log(IV_{t-3})</td>
<td>0.479</td>
<td>-0.539</td>
</tr>
<tr>
<td>log(IV_{t-4})</td>
<td>0.682</td>
<td>-0.673</td>
</tr>
<tr>
<td>log(IV_{t-5})</td>
<td>0.637</td>
<td>-1.004</td>
</tr>
<tr>
<td>log(IV_{t-6})</td>
<td>1.483</td>
<td>-2.444</td>
</tr>
<tr>
<td>log(IV_{t-7})</td>
<td>-0.115</td>
<td>1.443</td>
</tr>
<tr>
<td>log(IV_{t-8})</td>
<td>-0.197</td>
<td>0.730</td>
</tr>
<tr>
<td>log(IV_{t-9})</td>
<td>0.688</td>
<td>-1.674</td>
</tr>
<tr>
<td>log(IV_{t-10})</td>
<td>-0.176</td>
<td>0.530</td>
</tr>
<tr>
<td>log(IV_{t-11})</td>
<td>-0.062</td>
<td>-0.331</td>
</tr>
<tr>
<td>log(IV_{t-12})</td>
<td>0.462</td>
<td>-0.953</td>
</tr>
<tr>
<td>log(IV_{t-13})</td>
<td>0.582</td>
<td>-1.155</td>
</tr>
<tr>
<td>log(IV_{t-14})</td>
<td>-0.227</td>
<td>0.942</td>
</tr>
<tr>
<td>log(IV_{t-15})</td>
<td>-0.133</td>
<td>0.361</td>
</tr>
<tr>
<td>log(IV_{t-16})</td>
<td>-1.040</td>
<td>2.933</td>
</tr>
<tr>
<td>log(IV_{t-17})</td>
<td>-0.662</td>
<td>1.032</td>
</tr>
<tr>
<td>log(IV_{t-18})</td>
<td>0.074</td>
<td>0.043</td>
</tr>
<tr>
<td>log(IV_{t-19})</td>
<td>-0.389</td>
<td>1.229</td>
</tr>
<tr>
<td>log(IV_{t-20})</td>
<td>-0.375</td>
<td>1.117</td>
</tr>
<tr>
<td>log(IV_{t-21})</td>
<td>0.475</td>
<td>-0.072</td>
</tr>
<tr>
<td>log(IV_{t-22})</td>
<td>0.413</td>
<td>-0.264</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Hidden layer2.node1</th>
<th>Hidden layer2.node2</th>
<th>Hidden layer2.node3</th>
<th>Hidden layer2.node4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias node 2</td>
<td>-0.455</td>
<td>0.099</td>
<td>-0.243</td>
<td>1.265</td>
</tr>
<tr>
<td>Hidden layer1.node1</td>
<td>-0.343</td>
<td>-2.433</td>
<td>1.623</td>
<td>5.720</td>
</tr>
<tr>
<td>Hidden layer 2.node2</td>
<td>-1.921</td>
<td>-0.487</td>
<td>0.923</td>
<td>-1.611</td>
</tr>
</tbody>
</table>
Table 4.8 (continued)

<table>
<thead>
<tr>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias node 3</td>
</tr>
<tr>
<td>Hidden layer2.node1</td>
</tr>
<tr>
<td>Hidden layer2.node2</td>
</tr>
<tr>
<td>Hidden layer2.node3</td>
</tr>
<tr>
<td>Hidden layer2.node4</td>
</tr>
</tbody>
</table>
CHAPTER 5
CONCLUSIONS

Agricultural futures markets have experienced several structural changes since the first contracts started to be traded in 1877 at the Chicago Board of Trade. Irwin and Sanders (2012) identify three main structural changes in the agricultural futures markets in the U.S. in the last decade: predominance of electronic trading, easier access to trading on futures markets, and the entry of new markets participants. Most of these changes occurred in the period from 2006 to 2008 and have substantially modified the landscape of futures trading environment of today. In fact, another major structural change has emerged since then, facilitated by the electronic trading environment, the automated trading, which relies on computerized algorithms. The rise of computerized algorithms has rapidly spread to various industries which experienced automation, including the futures market industry (Meyer 2019). However, this automation also has led to a race for gains from speed and the emergence of low latency trading in agricultural futures markets. This race concerns policy makers and producers using these markets for hedging purposes and whom are now calling for a slowing down of futures markets (Stafford 2019). This context raises concerns on the role of automation in agricultural futures markets. In this dissertation, I investigate three aspects of intraday market dynamics in this new trading environment.

In the first essay, I examine the microstructure noise and integrated variance components of the price variance in the live cattle futures market. In 2015 and 2016, intraday live cattle futures prices experienced heightened variance. The U.S. beef producers’ association attributed this event to high frequency trading activities that might increase order execution risk. The analysis substantiates the notion that the high 2015-2016 variance was strongly influenced by market participants incorporating information about fundamentals, captured by integrated variance. In contrast,
microstructure noise is found to be economically small and to dissipate within 3 to 4 minutes. The magnitude of noise also varies with the magnitude of changes in demand and cyclical supply. Examinations of CME actions to reduce price variance in this market suggest that they had little or no effect on the market volatility. Overall, this research demonstrates the importance of accounting for noise when one is concerned by the informational volatility. Additionally, it is the first effort to identify the economic value of noise in the live cattle futures market which is found to be less than 1% of transaction prices.

In the second essay, I document that intraday jumps in corn futures prices have increased with electronic trading and the shift to real-time announcement of USDA reports. Using intraday prices from 2008 to 2015, we employ a nonparametric test to detect jumps and variance analysis to estimate jump or execution risk. Real-time trading of major USDA reports has substantially increased the frequency and clustering of price jumps, and results in higher market liquidity costs. In contrast, while the presence of jumps on non-announcement days has doubled recently, their magnitude has declined as have transactions costs during their occurrence. Traders with higher frequency trading activities can experience the largest jump risk due to heightened microstructure noise during price jumps. This result is consistent with the situation examined by Budish, Cramton and Shim (2015) that high-speed traders through their dual wasteful arms race are influencing price discovery. Overall, this research identifies the impact of recent structural changes on price behavior throughout the identification of changing jumps behavior that leads to a higher level of risk for market participants. However, while there appears to be rather substantial jump risk around USDA market releases, the overall magnitude of the jump risk in the most recent volatile period is not large. We find that the weighted average jump risk share of efficient volatility only reaches
1.36% for all trading days in the period, a finding similar to that encountered by Christensen et al. (2014) in financial markets.

The last essay investigates volatility forecasting. This is the first effort to assess the value of HAR and ANN models in forecasting U.S. agricultural market volatility. With the increase in volatility and the availability of improved high frequency data due to the electronic trading, volatility forecasting is more relevant but challenging. Two characteristics of appear to dominate price volatility. The first is the presence of heterogeneous volatility expectations from a diversity of traders that can result in long-memory process, and second is the nonlinearities such as seasonality or public information shocks. The main findings document that long-memory is a strong pattern of the realized volatility in corn futures markets and should be taken into account in forecasting volatility at different horizons. Allowing for nonlinearities through artificial neural networks model improves forecasting of volatility in corn futures market when realized volatility (RV)-based models are employed. The ANN model is particularly useful to forecast volatility at longer-horizons compared to the other models which suggests that accounting for complex nonlinearities improves long-horizons forecasts.

Several contributions of the thesis can be summarized in general terms. This dissertation provides one of the first efforts in agricultural economics to use high frequency data to analyze specific research questions that have become important with recent structural changes in futures markets. For instance, it provides an evaluation of microstructure noise and its economic value in the live cattle market. Second, it is the first attempt to identify price jumps behavior in the corn futures market and to relate it to the jump variation component of volatility. Third, in the context of volatility forecasting in corn futures, this dissertation highlights the importance of accounting for nonlinearities through artificial neural networks which are especially useful on long-horizons
forecasts. It also documents the importance for using realized volatility for forecasting purposes. The informative ability of the RV in forecasting suggests that it may be more useful in describing risk in other modeling situations as well.

Many issues remain to be answered with regards to how high-speed traders interact with conventional traders. This dissertation has focused on the market microstructure changes, informational volatility dynamics, and forecasting a new trading environment marked by automation. Further aspects such as how different categories of traders interact during price jumps in agricultural futures markets and to which extent this affects the hedging role of futures contract for commodities markets is still unknown. In addition, further efforts could be conducted to revisit the performance of the implied volatility based on option prices, that is usually employed as a good predictor of volatility in corn markets, compared to the realized-based models.

5.1. References


