Volatility Estimation in Agricultural Futures Markets: A Microstructure Approach.

by

Xianglin Kong

A Thesis submitted to the Faculty of Graduate Studies
The University of Manitoba
In partial fulfillment of the requirements of the degree of
Master of Science

Agribusiness and Agricultural Economics
University of Manitoba
Winnipeg

Abstract

Agricultural markets are known to be more volatile than the other markets. Understanding volatility movements is important to improve both risk management strategies and market forecasts. In an order-driven electronic trading system, the Limit Order Book (LOB) contains trading information based on market participants expectations. Such information may help explain volatility and utilised to make forecasts. In agricultural market, most previous studies have used daily information to predict volatility. This research uses intraday data to forecast volatility in both lean hog and corn markets. Two models are considered, the well-known GARCH (1,1) and the GARCH-X model which includes LOB information. Intraday forecasts coming from GARCH and GARCH-X are compared with intraday realized volatility (RV). Our findings suggest that GARCH and GARCH-X model forecasts are more in line with each other than with RV, and that GARCH-X forecasts do not outperform GARCH (1,1) forecasts. Reasons associated with these findings, limitations of this study, and future work are outlined.

Table of Contents

List of Tables

List of Figures

List of Acronyms

Autoregressive Conditional Heteroskedasticity (ARCH)

Bid-Ask Spread (BAS)

Central Time (CT)

Chicago Board of Trade (CBOT)

Chicago Mercantile Exchange (CME)

Diebold-Mariano (DM)

Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH)

Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

Limit Order Book (LOB)

Market Microstructural Noise (MMN)

Mean Absolute Error (MAE)

Mean Absolute Error (MAE)

Mean Absolute Percentage Error (MAPE)

Realized Volatility (RV)

Root Mean Square Error (RMSE)

Threshold Generalized Autoregressive Conditional Heteroskedasticity (TGARCH)

United States Department of Agriculture (USDA)

1. Introduction

Summary. Understanding volatility movements is an area of great interest in agricultural markets. The limit order book (LOB) of an exchange contains information that may help improve volatility forecasts. In addition, intraday data is known to contain information that would be missed when using daily data. This section discusses the standard approach of volatility studies vs the changes associated with the adoption of the electronic platform to trade agricultural commodities. On this basis, this section provides the motivation and objective of the research undertaken here.

Understanding volatility movements is important to improve both risk management strategies and market forecasts. Estimating assets' volatility has been the subject of a vast area of research. In particular, agricultural commodity markets are known to be more volatile than other markets (Tomek and Kaiser, 2014), which has sparked great interest as an area of research. Most studies on the volatility of agricultural futures markets refer to prices observed in the traditional open-outcry trading system, which does not capture market dynamics in the current order-driven electronic trading system. In the current system, more information is available at lower latency, which has important implications for volatility modelling. Volatility estimation and forecasting require a good understanding of the trading system in which prices are discovered.

In an order-driven electronic trading system, market participants submit limit orders that are placed in the limit order book (LOB) of the exchange. The steps of the LOB contain all the bids (asks) in descending (ascending) order with their corresponding number of futures contracts that traders are willing to buy (sell). The LOB is updated when an existing limit order is fulfilled, modified, or cancelled, or when a new limit order is submitted. The LOB is also updated when a market order is submitted, which will be matched and filled at either the best bid price or the best ask price. Therefore, the prices and quantities in the LOB should reflect traders' beliefs and market expectations about the underlying commodity, and this information could help predict volatility. Few studies have utilized the information contained in the LOB to study agricultural commodity markets. For example, Arzandeh and Frank (2019) showed that the steps beyond the top of the LOB contain information that contributes to price discovery. He et al. (2020) used the depth beyond the top of the LOB to show market adjustment to large price changes. However, studies using LOB information to model volatility are scarce. Oz (2022) used the best bid and ask quotes and three levels of depth information contained in the LOB to estimate price impact and to improve volatility forecasts in the wheat market. Adjemian and Irwin (2018) used intra-daily data in conjunction with the USDA reports to ascertain the impact of new information on volatility.

In addition to the LOB containing market information at the quote and quantity levels, the LOB also captures very frequent market updates, which have been shown to improve volatility forecasts. Andersen and Bollerslev (1998) showed that using intraday data for volatility estimation results in less noisy estimates. They found that volatility forecast evaluations and measurement errors improved by reducing the time interval from hourly to five-minute intervals. Zhang and Dufour (2019) and Frijns and Margaritis (2008) provided evidence that intraday data contained information that could improve daily volatility forecasts. When the dataset was obtained more frequently, there were short-term fluctuations that could be observed. As soon as the news came to the market, the traders normally responded quickly, which was almost invisible from low-frequency daily data. The benefits of using intraday data in volatility estimation plus the easier access to high-frequency data in the last few years have increased the amount of research in microstructure agricultural markets (some examples are Arzandeh and Frank, 2019; Couleau et al., 2018; Couleau et al., 2020; He et al., 2020).

Using high frequency intraday data and the information contained in the LOB, this research aims to estimate and forecast volatility. The analysis is performed for the lean hogs and corn markets. The time interval between observations is determined by how often trades occur in the market. Two models, one containing LOB information and one without it, are estimated to perform volatility forecasts. The forecasts are evaluated using one measure of "real volatility," the standard measure widely used in the literature.

In market microstructure studies it is common practice to sample observations at regular time intervals. For example, a large number of studies use 5-minute time intervals (Andersen et al., 2001; Barndorff-Nielsen and Shephard, 2002; Liu et al., 2018). Higher-frequency price observations are prone to market microstructural noise (MMN) which complicates volatility estimation. The sources of MMN are the bid-ask bounce, the discreteness of the data (Harris, 1990), and the properties of the trading mechanism (Hansen and Lunde, 2006). The bid-ask bounce is the limited range of price back and forth between the bid price and the ask price, which makes the "real" or efficient price not observable. Aït-Sahalia and Yu (2009) describe a liquid market as a large quantity offer that could be fulfilled soon after it goes into the market, and in which the price will be close to the price from the preceding and following transaction. However, this price is unclear. Discrete prices are driven by market regulations that all prices need to be a multiple of the minimum tick (Harris, 1990), for example, in lean hogs the minimum tick is $10 and in corn is $12.5. The properties of the trading mechanism in the agricultural future market is a clearing house in which traders submit two types of orders – limit and market orders. The single price, which clears the market and it is applied to all executed orders, is the transaction price (Amihud and Mendelson, 1987). The trading mechanism also requires market participants to follow regulations: actions they can take, information available, and matching orders (Foucault et al., 1955). These sources of MMN, also referred to as market frictions, lead to two main consequences for volatility estimation. The first one is that when sampling occurs at very short time intervals, a major component of what is being captured is MMN rather than price volatility (Aït Sahalia et al., 2005). The more often the data is sampled (that is, the shorter the time intervals between observations), the stronger the effect of MMN on estimates (Zhang et al., 2005). The noise is due to the imperfection of pricing observations (Black, 1986). Black further explains that when the market includes more noise trading, it will become more liquid. The reason is more unprofessional or impulsive traders making decision without using information, which will bring more trades to the market. Therefore, high-frequency data include more noise than low-frequency data. The second consequence brought by the frictions described above is that observed prices do not accurately represent the true underlying price of a commodity. Rather, the true, efficient, or fundamental price of a commodity falls in the space between the observed best bid and ask quotes and it is not directly observed. Therefore, it is not clear which price (or return) should be used for the estimation of volatility. The approaches undertaken to tackle these two consequences are described in the next two paragraphs.

Realized volatility (RV) is a widespread measure of real volatility (Andersen et al., 2001; Meddahi, 2002; Barndorff-Nielsen and Shephard, 2002). However, results in previous studies suggest that when data is sampled at very short time intervals, the RV is a biased estimator (due to MMN), and it is not robust to the length of the sample interval. These results prompted many studies to sample data at larger time intervals and discard observations between those time intervals, a procedure that many find conflicting as massively discarding observations is, in most cases, not desirable. To overcome this problem, the average trading frequency (Oz, 2022) is used to define the snapshot.

Most volatility studies on agricultural commodities' future markets are based on observed transaction prices. Earlier studies used daily data as intraday transaction prices for these markets were not publicly available before trading switched to the electronic platform. In most recent studies, intraday transaction prices data have been used in the current electronic trading mechanism. For example, Couleau et al. (2018) used transaction prices to identify the MMN in live cattle future markets. However, as mentioned above, transaction prices may not accurately reflect the unobserved true price of a commodity. In addition, Kavajecz and Odders-White (2001) argue that transaction prices were the result of liquidity demanders and suppliers' actions, which could be driven either by their level of information or simply by their liquidity needs. In the latter case, transaction prices are not informative. The authors further discuss that the midpoint between the best ask and the best bid has been used as a better approximation of the true price of a commodity, which lies somewhere between the best two quotes. Hagströmer (2021) showed that, for S&P500 stocks, the midpoint is not a good proxy for the true price because the true price is a continuous variable, whereas the best quotes used to compute the midpoint are discrete variables. Specifically, the best quotes only move by the tick size and do not capture gradual changes in between. Hagströmer (2021) suggests using the best quotes' quantities in computing a proxy for the true price. Therefore, this research uses the weighted midpoint with the best ask and bid quantities being the weights.

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model has been widely used to estimate and predict volatility in both the financial and agricultural markets. It has been shown to work well when daily data (Cermák et al., 2017; Musunuru et al., 2013) as well as intraday high-frequency data (Andersen and Bollerslev, 1998) are used. In the GARCH model, newest observations have a higher weight than the older ones. Intraday data became available in agricultural markets at a slower pace relative to other markets. Therefore, most research using the GARCH model focused on daily data. However, the use of the GARCH model based on high-frequency data has been proven accurate and favorable. Andersen and Bollerslev (1998) show that GARCH models produce accurate intraday forecasts for volatility. The authors examined various types of GARCH models in exchange rate financial markets based on high-frequency intraday data. Their findings show that the GARCH model is suitable for analyzing the volatility of high-frequency intraday data.

In the electronic trading system, traders incorporate their information in the LOB. Therefore, LOB variables could be included in the GARCH model. This research assesses if those variables help improve volatility prediction. The GARCH-X model is estimated (Francq and Thieu, 2019) using ask and bid quotes and quantities. The estimated model is used to make one-day ahead forecasts (Degiannakis, 2017; Mcmillan and Speight, 2012). Traditional one-day ahead forecast coming from daily data use several days to predict the next one day. In this study, the model using intraday data for each day is used to forecast the next-day, and on the day of the forecast, each observation is used to make the forecast for the next time period (Chen et al., 2016).

To assess if the LOB information improves the accuracy of volatility prediction, the intraday forecasts from the GARCH and GARCH-X models are compared to RV. The RV was shown to perform better when the sample frequency increases (Oomen, 2005; Andersen et al., 2003; Barndorff-Nielsen and Shephard, 2004); making it a suitable measure of volatility when using intraday data. The Diebold-Mariano (DM) test (Harvey et al., 1998) is performed to assess forecasts differences between the GARCH and GARCH-X models. It is expected that the GARCH-X model provides a better prediction of volatility since it incorporates information present in the LOB. In turn, this would suggest that the LOB information is useful for market participants.

2. Literature review

Summary. This section is divided into five parts. The initial part is concerned with the measurement of volatility. It discusses the use of realized volatility (RV) as a means of identifying the underlying "true volatility." The second part examines the potential of LOB information to enhance the forecasting of volatility. The third part presents the findings of a comparative analysis between intra-day and daily data, suggesting that the former may offer superior performance. The fourth part delves into the merits of weighted mid-point return as a superior alternative to other return calculations. Finally, the fifth part outlines how to estimate the volatility forecasting model with LOB information.

The majority of research conducted on market microstructure and volatility forecasting has concentrated on financial markets, while only a limited number of studies have examined agricultural futures markets. In comparison to agricultural future markets, financial markets initiated electronic trading at an earlier point in time, namely in the 1970s. Consequently, there is a greater corpus of literature dedicated to these markets. The advent of electronic trading in agricultural future markets was marked by the introduction of the electronic platform by the Chicago Board of Trade (CBOT) (Irwin and Sanders, 2012; Gousgounis and Onur, 2018). The CBOT was historically the primary venue for trading grain products. At the outset of the introduction of the electronic platform, the predominant focus of trading remained on pit trading, largely due to historical precedent. Between 2000 and 2011, there was a notable surge in electronic trading for grains. A similar trend is evident in the case of hogs and live cattle which were trading on the Chicago Mercantile Exchange (CME). In 2007, both exchanges were merged into the CME Group. Despite the prevalence of electronic trading in both the agricultural and financial markets, notable distinctions between both markets remain. In financial markets, the number of trades is greater, and the frequency of trading is also higher than that observed in the agricultural markets. Research for agricultural market has seldomly incorporated high frequency data and LOB information as of today.

1) Measures of volatility

The concept of “true volatility” is not susceptible to empirical observation, which explains much of the efforts in the literature devoted to its measurement. Andersen and Bollerslev (1998) suggest using the sum of intra-day returns, the RV, as a measure of the “true volatility.” Subsequently, RV has became a widely used measure of volatility (Andersen et al., 2001; Meddahi, 2002; Barndorff-Nielsen and Shephard, 2002). The RV represents a model-free approach to estimating volatility and it is frequently regarded as a proxy for the “true volatility” when assessments are made of estimates derived from disparate models. The RV can be used at varying scales such as daily or intraday. It is important to choose suitable time interval for intraday RV. It should avoid the effect from MMN, and at the same time, it should contains as many observations as possible.

2) LOB information

The LOB contains information such as bid and ask prices, bid and ask depth, and spread, which can be used to improve the estimation of volatility. In financial markets, Jain and Jiang (2014) use LOB information beyond the top of the book to predict future price volatility on the Shanghai Stock Exchange which is an order-driven market. The study focusses on stocks and uses one-minute time intervals. The authors show that the best bid and ask contain information which significantly predicts future price volatility. In the electronic platform system, after the best bid and ask depth match and trades occur, unfilled contracts will move to the second bid and ask price to be filled. Accordingly, the depths at the top and the second level of both bid and ask sides are used in this research to extract the information contained in the LOB. The spread has been shown to include information. Bollerslev and Melvin (1994) found that the bid-ask spread is positively correlated with the exchange rate market volatility through a GARCH-type model. Furthermore, research conducted on agricultural futures markets that included LOB information found that prices as well as quantities present in the LOB are relevant. To illustrate, Arzandeh and Frank (2019) investigated the contribution of LOB information beyond the bid-ask spread (BAS) to the price discovery process in five agricultural products from the CME Group. The results show that the depths of the LOB beyond the BAS contain information and suggests that informed traders use the depths of the LOB beyond the top of the book actively. Nevertheless, the LOB information is not used for the purpose of analyzing volatility. He et al. (2020) used minute-level data and the spread and depth from the CME Group to show that in the corn market, the depth is markedly diminished during substantial price fluctuations, whereas in the hog market, a significantly higher spread is observed during large price movements. The aforementioned studies show that the LOB contains information that can be used to predict volatility. The only study that specifically incorporates the information contained in the LOB to predict volatility in agricultural markets is Oz (2022). The study focuses on the impact of incoming orders on volatility. The above evidence of the information contained in the LOB, in addition to the limited studies performed in agricultural markets warrants further exploring volatility prediction in such markets. In this study, BAS and the depths at the top and second level from the LOB are incorporated into a GARCH-type model to predict volatility.

3) *Daily data vs. intraday data*

Based on market microstructure theory, information is embedded in traders' strategies with each new order that enters the market. Because this information would not necessarily be reflected when using daily data, the focus of this study is on the microstructure level data. The quantity of observations coming from intraday data is contingent upon the frequency with which the observations are sampled. The 5-minute frequency is a commonly used time interval (Andersen et al., 2001; Barndorff-Nielsen and Shephard, 2002). The MMN has an impact on high-frequency data, distorting the distribution of high-frequency intra-daily returns. In a study conducted by Clinet and Potiron (2019), the existence of the MMN was detected at a high frequency in stocks using tick-by-tick time, 15 and 30 seconds. The five-minute interval has been proven to be sufficiently brief to estimate volatility using RV, while also being sufficiently long to avoid the impact of the MMN effect (Meddahi, 2002; Andersen et al., 2001). Most of the literature studying volatility in agricultural markets uses daily price data. For instance, Cermák et al. (2017) used daily data on wheat from the CME Group to analyze the impact of shocks on volatility. In a similar vein, Degiannakis and Filis (2017) used tick-by-tick time intervals to forecast volatility in the oil market over a range of time horizons, from 1-day ahead to 66-days ahead. The results of the study indicate that the incorporation of data pertaining to cross-market assets, stock market indices, currency, and macroeconomic indicators enhances the precision of volatility predictions. Liu et al. (2018) determined that a sampling frequency of 5-minute intervals was optimal for oil futures markets with decomposed jumps, and subsequently forecast real oil price volatility. All of the aforementioned papers used fixed time intervals for all products. However, it is possible that the optimal sampling frequency differs between products with different trading frequencies. Accordingly, this paper employs a variety of snapshot times, contingent on the trading frequency of each product (Arzandeh and Frank, 2019; Oz, 2022).

4) *Computation of returns*

This study will employ an optimal methodology for computing returns. In their 2001 study, Kavajecz and Odders-White compared LOB midpoint returns, quote midpoint returns and transaction returns. Their findings indicate that earnings announcements had the greatest impact on LOB midpoint returns, providing evidence that midpoint returns are more sensitive than transaction prices. Furthermore, Hagströmer (2021) posits that the weighted midpoint price, wherein the best bid and ask quantities serve as the weights, is superior to the midpoint price in addressing the issue of discrete prices. In a different approach, Andersen and Bollerslev (1998) computed the weighted midpoint by using bid and ask quantities as the ask and bid price weights, respectively. The rationale is that bid (ask) quantities are introduced into the market to fulfil ask (bid) orders, and thus the bid (ask) quantities will impact the price set by the sellers (buyers). Cao et al. (2009) also used the bid and ask quantities as the weights. Therefore, since the weighted mid-point returns are a superior method to both transaction returns and midpoint returns, the weighted mid-point returns will be used in this research.

5) *GARCH model vs. GARCH-X model*

A substantial corpus of literature has used the GARCH model to estimate volatility and/or to facilitate comparisons between disparate estimation methodologies. To illustrate, Musunuru et al. (2013) used the TGARCH and EGARCH models to forecast the volatility returns of corn and found that volatility exhibited distinct reactions to positive and negative news. Benavides (2004) compared ARCH models, univariate GARCH and multivariate ARCH

(BEKK model), and highlighted the challenges associated with accurate forecasting of asset price return volatility. The LOB information can be incorporated into the GARCH model via the GARCH-X model (Francq and Thieu, 2019). In a similar vein, Shah et al. (2009) used the GARCH-X model to analyze exchange rate volatility in Pakistan. The exogenous variables are the short-term interest rate, the capacity for intervention by the State Bank of Pakistan and expectations regarding the exchange rate. Cermák et al. (2017) used the high-frequency data and a GARCH model for forecasting the volatility of wheat futures prices. It was suggested that the wheat producers could hedge their production with short-term futures contracts, given the capacity of a fitted model to predict price fluctuations in the short term. Manfredo et al. (2001) used a GARCH model to forecast the volatility of cattle and corn markets. They further proposed that the GARCH model facilitates a more comprehensive understanding of volatility forecasting based on specific data and time horizons. The aforementioned research in agricultural futures markets does not incorporate LOB information for the purpose of forecasting volatility. It may therefore be posited that the incorporation of LOB information into the GARCH-X model will result in a more accurate forecast of market volatility.

3. Data

Summary. This section describes the hogs and corn data used in this research. Specifically, it outlines how the data is sampled from the LOB, and it provides some intuition for the selection of the variables that will be included in the model. The last part of the section specifies how the return series is computed by incorporating LOB information.

The data comprises futures contracts for lean hog and corn trading in the CME Group, spanning the period from November 23, 2018 to December 27, 2019. The lean hogs trading session commences at 8:30 am and concludes at 1:05 pm CT. For corn, only the data from the morning session data is employed, extending from 8:30 am to 1:20 pm CT, given that trading activity is more prevalent during that period. It is standard practice to exclude the first and last five minutes of trading from the data set, as this period is characterised by a different set of behaviours than those observed during the main trading session. In order to ensure that the data is representative, all dates with low trading volume, shorter trading times and holidays have been excluded. The contract unit for lean hog is 40,000 pounds, which is equivalent to approximately 18 metric tons. The contract unit for corn is 5,000 bushels. The price is quoted in US dollars. Contracts with different maturities are traded on a daily basis. The highest-volume trading contract for each day is used. The data are presented as snapshots taken at regular intervals. Five-minute snapshots are a common approach (for example, Luo et al., 2019; Luo and Chen, 2019). However, with high frequency data, this would result in a large amount of data being omitted between snapshots. A smaller time interval would result in an increase of the effect of MMN. Furthermore, the MMN has been demonstrated to be excessively sensitive to ultra-high-frequency data (Zhou, 1996). A data-driven sampling method is used in this research.

The LOB is a list of bid and ask orders at a specific price. The LOB is updated in accordance with the submission of new orders, the cancellation, amendment or trading of existing orders. The orders remain in the book until a match is made and the trade is executed. Two distinct types of limit orders exist. One such order is the bid order, which is an order to purchase at the specified price. The alternative is the ask order, which is the order to sell at a specified price. Upon submission of the order, the pertinent information is duly recorded in the LOB and subsequently organised according to price level. An order may be understood as a reaction on the part of traders to information. Traders analyse market information and submit orders based on their predictions. It should be noted that LOB updates resulting from trades occur less frequently than other updates. Consequently, the series of returns based on transaction prices undergoes less frequent changes relative to the time series containing all LOB updates.

To circumvent the issue of limited change in the return data based on transaction prices, a series constructed using snapshots at the average trading frequency is used (Arzandeh and Frank, 2019). During the time period between November 23, 2018 and December 27, 2019 the average trading frequency for lean hogs was 6 seconds, and for corn was 9 seconds. The corn market is known to be more active relative to the lean hogs market. However, a high proportion of transactions in corn occur at the same price and therefore not counted in the computation of the average trading frequency. The spread is calculated using the best bid and ask prices; while returns are calculated using the best bid and ask price in conjunction with the bid and ask quantities. The weighted mid-point price (W_t) is computed as follows:

$$W_t = \frac{A_t \times QA_t + B_t \times QB_t}{QA_t + QB_t} \qquad (1)$$

where A_t is the best ask price at time t, B_t is the best bid price at time t, QA_t is the quantity for the best ask price at time t, and QB_t is the quantity for the best bid price at time t.

The weighted mid-point returns (RW_t) are computed as follows:

$$RW_t = \frac{\log(W_t) - \log(W_{t-1})}{\log(W_{t-1})} \tag{2}$$

4. Methods

Summary. This section describes a measure of real volatility (RV), and two models (GARCH and GARCH-X) used for the prediction of volatility. It also presents the criteria used to evaluate forecasting performance, as well as a formal test to assess if competing models' predictions have significant differences.

Volatility is estimated using three methods: realized volatility (RV), GARCH and GARCH-X models.

The GARCH model is the most commonly used volatility model. The original specification of the GARCH (1,1) model (Engle, 1982) is:

$$RW_t = \mu + \theta RW_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim (0, h_t) \tag{3}$$

$$h_t = \omega + \alpha\, \varepsilon^2_{t-1} + \beta\, h_{t-1} \tag{4}$$

where $\omega > 0$, $\alpha > 0$, $\beta > 0$, and $\alpha + \beta < 1$

The variance of the error term is a function of the squared of the previous innovations and on its own past values. The second model being used in this research incorporates the high-frequency data from the LOB (Visser, 2011; Deng et al, 2020). In the GARCH-X model exogenous variables are used in the variance equation (Engle, 2002). Foucault et al. (2007) argue that limit order traders may possess private information on the occurrence of future events leading to price changes, and therefore adjust their order submission strategies to the level of risk perceived. In this vein, a wide bid-ask spread would signal that limit order traders expected high volatility. Aidov and Daigler (2015) argue that depth demonstrates the level of interest among traders at a particular price level and also provides insight into the cost of trading. Market depth is utilised by traders as a means of gauging the probable trajectory of price. Furthermore, it demonstrates the willingness of traders to engage in buying and selling activities. In particular, Pascual and Veredas (2010) find that the depth at the best quotes and the depth beyond the best quotes play different roles in signaling changes in the efficient price volatility. Accordingly, the best bid and ask depths as well as the second level bid and ask depths are incorporated into the model. The GARCH-X model is:

$$RW_t = \mu + \theta RW_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim (0, h_t) \tag{5}$$

$$h_t = \omega + \alpha\, \varepsilon^2_{t-1} + \beta\, h_{t-1} + \gamma Z_t \tag{6}$$

where the matrix Z_t contains the exogenous variables including the best ask quantity, the best bid quantity, and the spread. One-day-ahead volatility forecasts are estimated using the GARCH model (3)-(4) and the GARCH-X model (5)-(6). In order to produce a forecast, data from day *i* is used to estimate the forecasting model. On day *i*+1, volatility is predicted for each *t* using the observed data at *t*-1.

The “real volatility” is measured using RV. The RV was shown to perform better when intra-daily data is used instead of daily data (Oomen, 2005). RV for observation *t* is computed as follows (Andersen and Bollerslev, 1998):

$$RV_t = [\log(W_t) - \log(W_{t-1})]^2 \tag{7}$$

The different RV intraday time interval is chosen based on the average trading frequency (Arzandeh and Frank, 2019) of each commodity. It should be noted that the average trading frequency does not include repeating trade information and instead captures snapshots of the data based on transaction price changes. This approach is preferable to the use of tick data, as it mitigates the MMN problem while retaining the trading information. The average trading frequencies for each commodity are as follows: 6 seconds for hogs and 9 seconds for corn.

Forecasts are evaluated using several methods. The Diebold-Mariano (DM) test based on MSE, MAE, and Mean Absolute Percentage Error (MAPE) will be used to compare the forecast error from the GARCH and GARCH-X models with RV. A significant result would show differences between GARCH and GARCH-X models, and together with lower forecast errors for GARCH-X it would suggest that the LOB information included in the GARCH-X model is useful to predict volatility. The MSE, RMSE and MAE are used to compare GARCH and GARCH-X with RV. These methods compare the prediction results with the real volatility. A smaller error indicates a greater proximity to the actual volatility, suggesting superior forecasting capabilities. MSE, RMSE and MAE are computed as follows,

$$\text{MSE} = \frac{1}{T}\sum_{t=1}^{T}(RV - predicted)^2 \tag{8}$$

$$RMSE = \sqrt{MSE} \tag{9}$$

$$\text{MAE} = \frac{1}{T}\sum_{t=1}^{T}|RV - predicted| \tag{10}$$

where predicted are the predictions coming from the GARCH and GARCH-X models. Two distinct MSE, RMSE and MAE outcomes will be generated: one for RV and GARCH model prediction, and one for RV and GARCH-X model prediction. The outcome showing the smaller error is considered the better prediction of volatility. The comparison of the results from RV with GARCH and GARCH-X models will provide evidence of whether the incorporation of additional information enhances the forecasting of volatility in the GARCH-X model.

5. Results

Summary. This section reports three main findings. First, the summary of the estimates of both the GARCH(1,1) and GARCH-X models for lean hogs and corn markets. Second, the comparison of the forecasts coming from the estimated model with the realized volatility. Third, the assessment of which model performs better relative to the realized volatility.

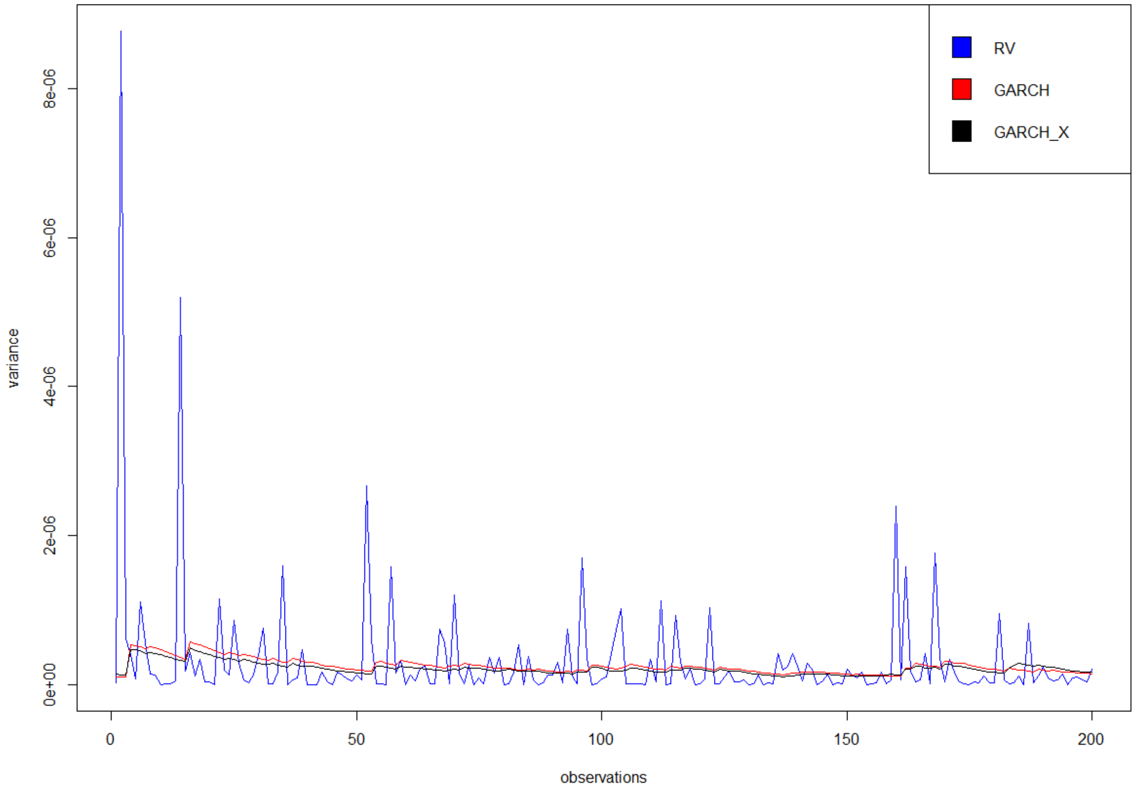
The LM test for autoregressive conditional heteroskedasticity (ARCH) is a statistical tool employed to evaluate the suitability of the GARCH and GARCH-X models. The results demonstrate that for HE, there are 198 out of 222 days with statistically significant ARCH effects in lag one; and more than 200 out of 222 days in lag two to ten. In the case of the ZC data, 236 out of 252 days exhibited significant ARCH effects in lag one, while in lag two to nine, the significant ARCH effects persisted in more than 240 out of 252 days; and 238 out of 252 days exhibited significant ARCH effects in lag ten. Consequently, the GARCH and GARCH-X models are deemed an appropriate choice.

The GARCH-X model was estimated for a period of 218 days for HE data. The p-value for the best ask quantity derived from the GARCH-X model is less than 0.05 for 196 out of 218 days. The p-value for the best bid quantity from the GARCH-X model is less than 0.05 for 194 out of 218 days. The second level of ask quantity derived from the GARCH-X model is statistically significant at 5% level for 172 out of 218 days. The second level of bid quantity from the GARCH-X model is statistically significant at 5% for 171 out of 218 days. The spread in the GARCH-X model is significant at 5% level for 190 out of the 218 days under consideration. The spread appears to be significant over the greatest number of days. For the sign of the exogenous variables, the expectation would be that when higher risk is perceived by informed traders, their strategy would be to submit less aggressive orders, that is, more second level and less first level orders (Ahn et al., 2001). The findings show that the sign of the best ask volume is positive for 131 out of the 218 days under consideration, and the sign of the best bid volume is positive for 131 out of 218 days. In contrast, for the second level, the sign for the ask volume is positive for 112 out of the 218 days under consideration, and the sign for bid volume is positive for 117 out of 218 days. The sign of the spread is positive for 180 out of 218 days.

The results for the ZC data exhibit a somewhat distinct pattern. The GARCH-X model is estimated for a period of 223 days. The p-value for the best ask quantity in the GARCH-X model is less than 0.05 for 222 out of 223 days. The p-value for the best bid quantity from the GARCH-X model is less than 0.05 for 220 out of 223 days. The second level of ask quantity in the GARCH-X model is statistically significant at 5% level for 203 out of 223 days. The second level of bid quantity in the GARCH-X model is significant at 5% level for 203 out of 223 days. The spread in the GARCH-X model is significant at 5% level for 192 out of 223 days. In comparison to the HE data, the greatest number of days on which the results are statistically significant is for the best bid quantity variable. The sign of the best ask volume is positive for 134 out of 223 days. The sign of the best bid volume is positive for 146 out of 223 days. In contrast, the sign of the second level ask volume is positive for only 98 out of 223 days, and for the second level bid volume is positive for 77 out of 223 days. The sign of the spread is positive for 63 out of 223 days, indicating that the negative sign is observed on approximately more than half of the total days.

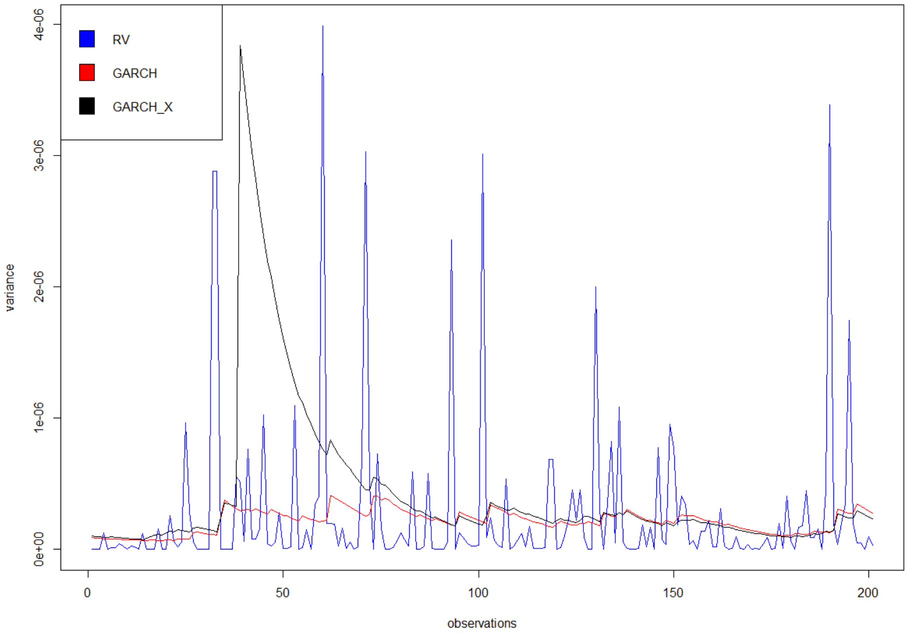
Figures 1 to 3 show the RV and GARCH and GARCH-X forecasts for the different intraday time periods on June 5, 2019 for hogs. One representative day (June 5, 2019) was selected and divided into three intraday periods to make each figure readable. The first period is from 8:36 am to 8:56 am; the second period is from 10:26 am to 10:46 am, which is the middle of trading time; and the last period is from 12:36 pm to 12:56 pm. It is evident from the figures that the RV fluctuates greatly. A notable peak is observed in the midday trading time period for GARCH-X when compared to the other time periods and days, which shows its sensibility to occasional large changes in quantities. Despite minor fluctuations, the GARCH model consistently yields higher predictions than the GARCH-X model, making the GARCH forecasts somehow closer to RV. For the figures 4 to 6, it is observed that throughout all time periods, the GARCH and GARCH-X models exhibit similar movements, with the GARCH model above the GARCH-X model. Concurrently, the RV fluctuates largely throughout the entire day.

Figure 1. Realized volatility and two volatility forecasts, lean hogs, morning hours



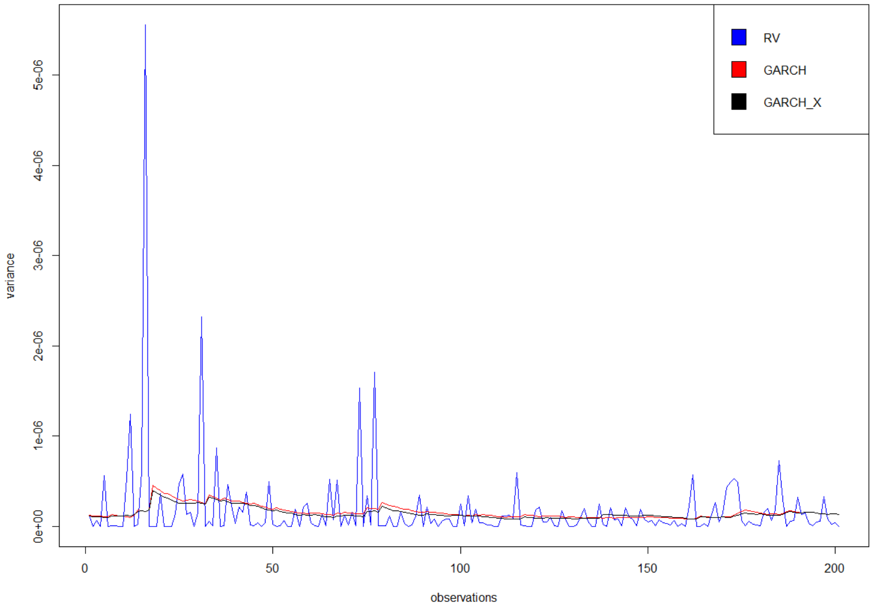
Note: The realized volatility and forecasts are plotted for June 5, 2019, 8:36 am to 8:56 am

Figure 2. Realized volatility and two volatility forecasts, lean hogs, midday hours



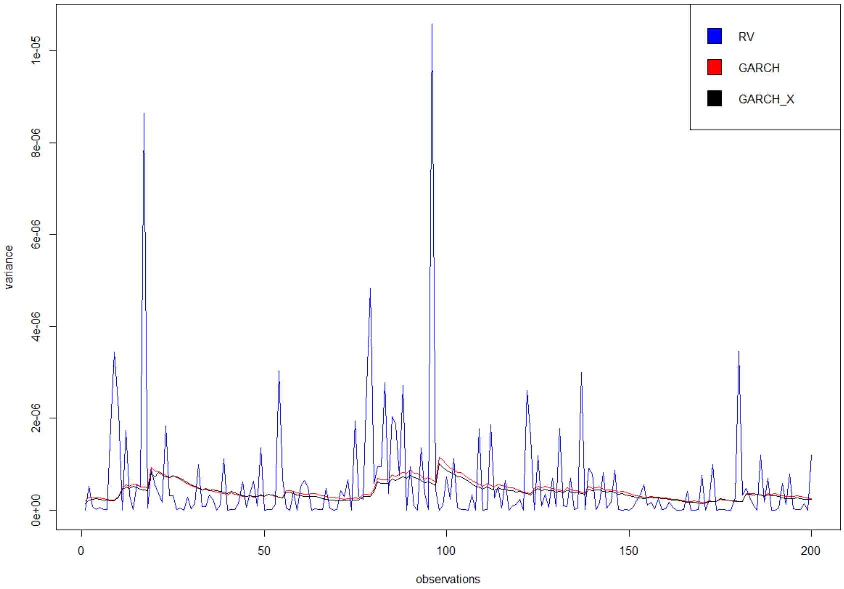
Note: The realized volatility and forecasts are plotted for June 5, 2019, 10:26 am to 10:46 am

Figure 3. Realized volatility and two volatility forecasts, lean hogs, last portion before close



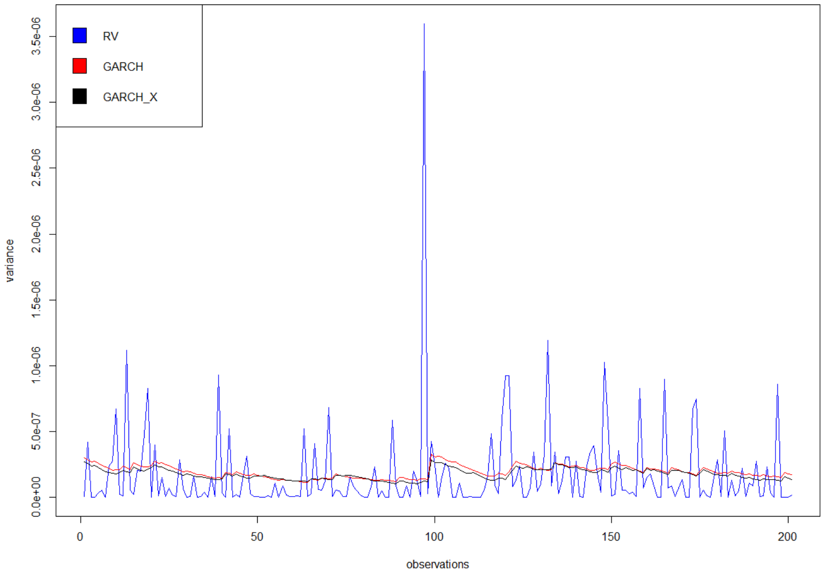
Note: The realized volatility and forecasts are plotted for June 5, 2019, 12:36 pm to 12:56 pm

Figure 4. Realized volatility and two volatility forecasts, lean hogs, morning hours



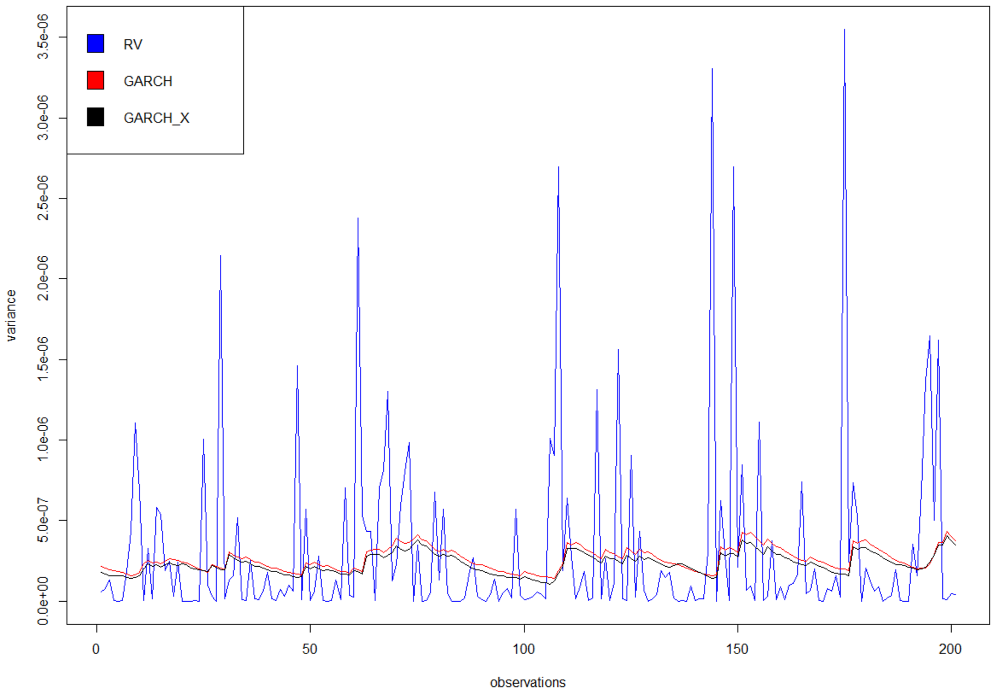
Note: The realized volatility and forecasts are plotted for September 9, 2019, 8:36 am to 8:56 am

Figure 5. Realized volatility and two volatility forecasts, lean hogs, midday hours



Note: The realized volatility and forecasts are plotted for September 9, 2019, 10:26 to 10:46 am

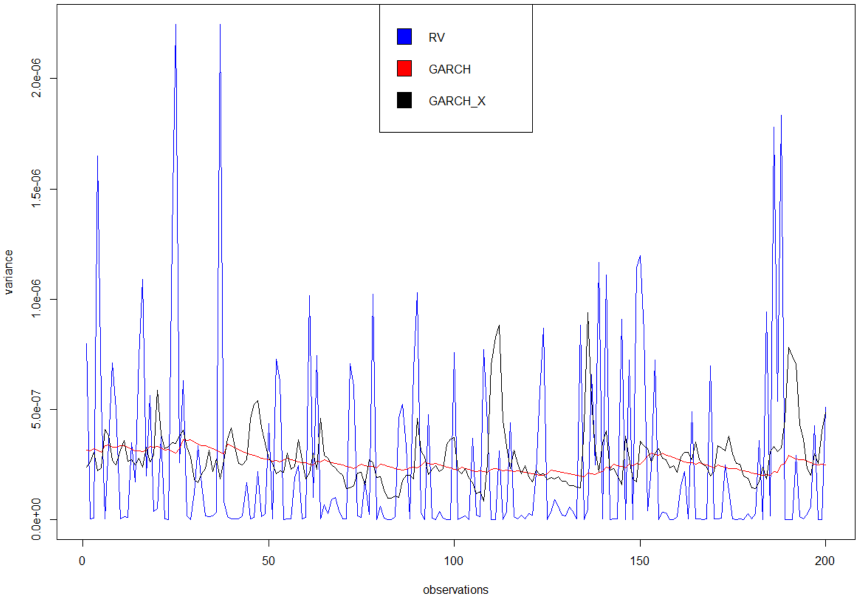
Figure 6. Realized volatility and two volatility forecasts, lean hogs, last portion before close



Note: The realized volatility and forecasts are plotted for September 9, 2019, 12:36pm to 12:56pm

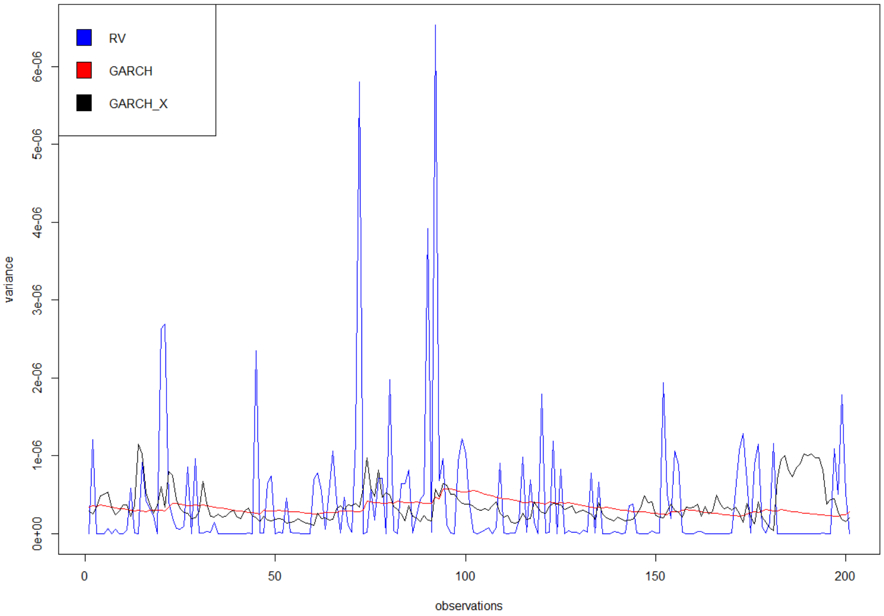
Figures 7 to 12 show the RV, GARCH and GARCH-X forecasts for corn on June 5, 2019 and September 9, 2019. The three time periods are 8:36 am to 9:06 am (morning); 10:41 am to 11:01am (midday); and 12:36pm to 1:06 pm (last portion before close). A comparison of the ZC data with the HE data reveals some differences. Both the GARCH and GARCH-X models for corn show higher values than those estimated for hogs' data. The GARCH model forecasts seem to be smoother over time while the GARCH-X model exhibits higher fluctuations, likely due to changes in quantities observed in the first two levels of the book. Compared to lean hogs, RV in corn exhibits a greater prevalence of small fluctuations. However, RV continues to exhibit the largest fluctuations compared to GARCH and GARCH-X models.

Figure 7. Realized volatility and two volatility forecasts, corn, morning hours



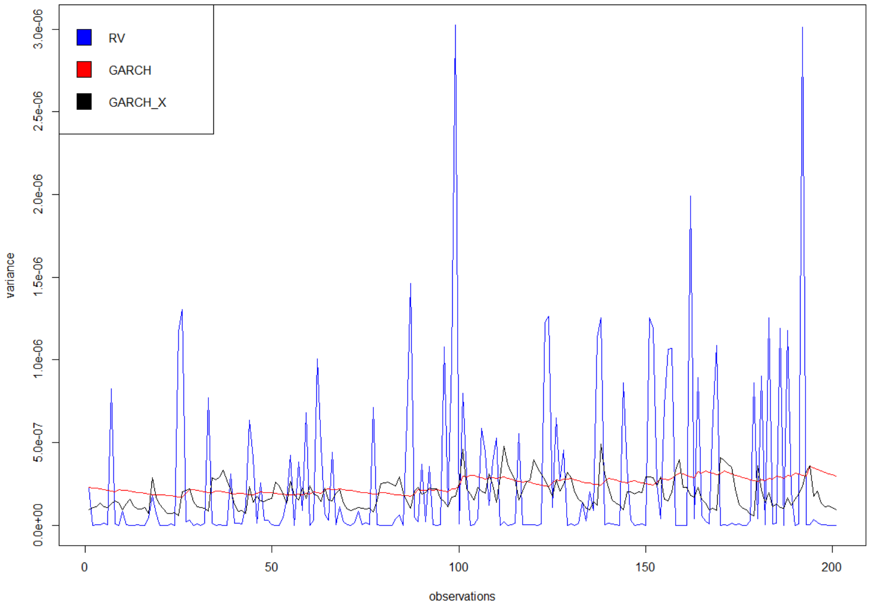
Note: The realized volatility and forecasts are plotted for June 5, 2019, 8:36 am to 9:06 am

Figure 8. Realized volatility and two volatility forecasts, corn, midday hours



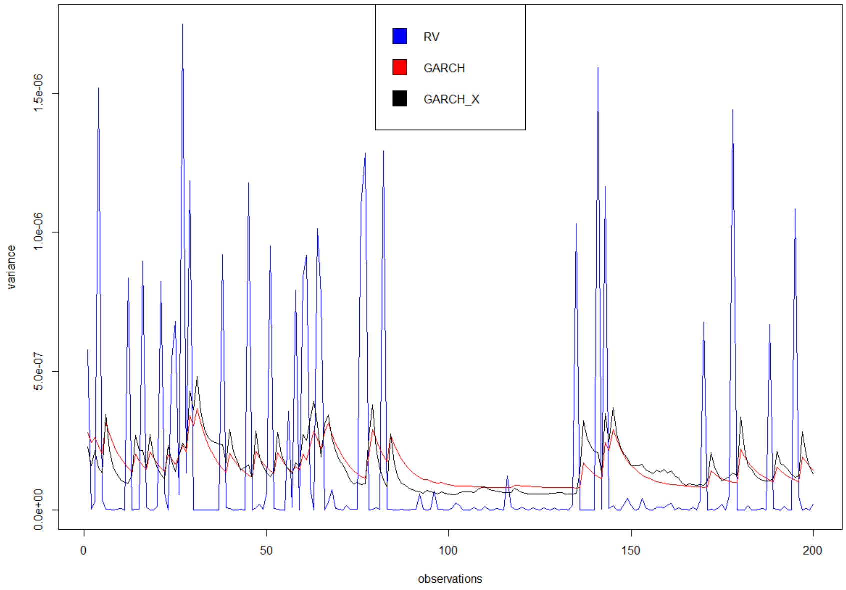
Note: The realized volatility and forecasts are plotted for June 5, 2019, 10:41 am to 11:01 am

Figure 9. Realized volatility and two volatility forecasts, corn, last portion before close hours



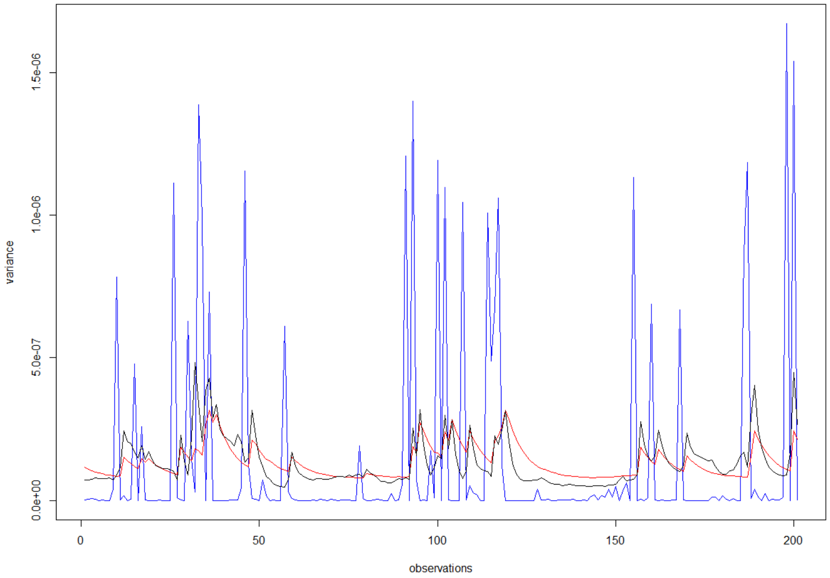
Note: The realized volatility and forecasts are plotted for June 5, 2019, 12:36 pm to 1:06 pm

Figure 10. Realized volatility and two volatility forecasts, corn, morning hours



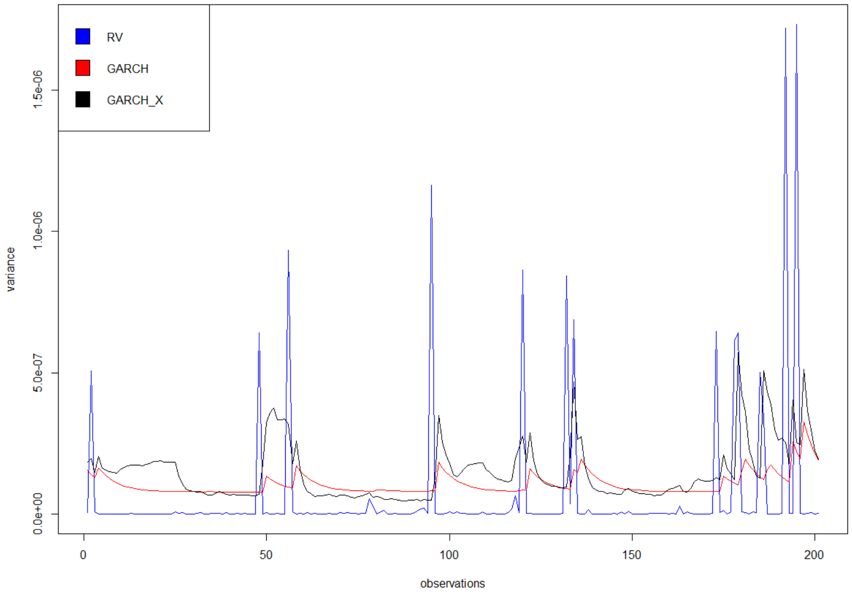
Note: The realized volatility and forecasts are plotted for September 9, 2019, 8:36 am to 9:06 am

Figure 11. Realized volatility and two volatility forecasts, corn, midday hours



Note. RV is blue, GARCH is red, and GARCH-X is black. The realized volatility and forecasts are plotted for September 9, 2019, 10:41 am to 11:01 am

Figure 12. Realized volatility and two volatility forecasts, corn, last portion before close hours



Note: The realized volatility and forecasts are plotted for September 9, 2019, 12:36 to 1:06pm

Forecast evaluation

The DM test is employed to ascertain the degree of discrepancy between both the forecast error of the GARCH and GARCH-X models and real volatility. The findings presented in table 1 utilise the DM test, based on MSE, MAE, and MAPE to access the forecasting performance of competing models GARCH and GARCH-X relative to RV for HE data. The ZC data presented in table 2 indicates a similar outcome. The results in tables 1 and 2 show that on most days, the GARCH and GARCH-X models are statistically indistinguishable, suggesting that there is no significant difference between them. Both tables show a low percentage number of days in which the GARCH-X model performs better than the GARCH model, suggesting that results vary for different days and commodities.

Table 1. DM test for lean hogs.

	p-value < 0.05	GARCH-X performs better than GARCH
MSE	31.52%	11.41%
MAE	42.93%	21.74%
MAPE	16.30%	36.96%

Note. There are a total of 184 days. The numbers in the table are are the percentage days in which the DM test is significant (column 2) and in which the GARCH-X model exhibits a lower forecast error than the GARCH model (column 3).

Table 2. DM test for corn.

	p-value <0.05	GARCH-X performs better than GARCH
MSE	47.89%	6.32%
MAE	66.32%	13.68%
MAPE	23.68%	17.39%

Note. There are a total of 190 days. The numbers in the table are the percentage days in which the DM test is significant (column 2) and in which the GARCH-X model exhibits a lower forecast error than the GARCH model (column 3).

The MSE, RMSE, and MAE are computed to assess how close the forecast is to the "real volatility" (RV). As illustrated in table 3, for HE data, the GARCH-X model exhibited a smaller error of a mere 11.41% of the total days, thereby suggesting that relative to RV, its forecasting performance is lower than that from the GARCH model. A similar result is shown in table 4, where only 6.32% of total days GARCH-X has a smaller error.

Table 3. *MSE, RMSE, and MAE for lean hogs.*

	RV
GARCH model	88.59%
GARCH-X model	11.41%

Note. The results are the same for MSE, RMSE, and MAE. There are 184 total days. The numbers in the table show the percentage days in which each model has a smaller error.

Table 4. *MSE, RMSE, and MAE for corn.*

	RV
GARCH model	93.68%
GARCH-X model	6.32%

Note: The result are the same for MSE, RMSE, and MAE. There are 190 total days. The numbers in the table show the percentage days in which each model has a smaller error.

6. Conclusions

Summary. This section delineates the findings from the study and its limitations. It also provides some explanations for less expected outcomes and suggests future work that may overcome some of the limitations found here.

This study employs intraday data from the CME Group to forecast volatility in the lean hogs and corn futures markets, incorporating microstructure features of the data. The full LOBs for both markets were reconstructed and market data was sampled at appropriate frequencies to effectively address the MMN issue. Two models were utilized in this study to forecast volatility. Firstly, the GARCH model, which is frequently used in both agricultural and financial markets to model and forecast volatility. Secondly, the GARCH-X model, which includes LOB variables such as bid-ask spread and bid and ask quantities observed in the first two levels of the book. Realized volatility is also computed and GARCH and GARCH-X forecasts are evaluated relative to the RV. The forecasting performance were evaluated using the Diebold-Mariano test.

The results suggest that both GARCH and GARCH-X models fit the data. In particular, the LOB variables included in the GARCH-X model appear to be significant in explaining the returns variance. However, somehow unanticipated, the GARCH-X model does not outperform the GARCH model's forecasting performance when using RV as a measure of real volatility. Both GARCH and GARCH-X forecasts appear to be more in line with each other and apart from RV, with GARCH-X exhibiting higher fluctuations. These results could be due to several reasons.

First, both models are compared to RV, which is constructed using similar data than the one used in the GARCH model, and which does not incorporate LOB information (other than the weights used to construct the return series). Future work could use, or develop, a different measure of realized volatility to use as a benchmark when assessing the forecasting performance of the estimated models, which takes into account more information from the LOB. For example, a measure of realized volatility using weights coming from steps beyond the best bid and ask could be constructed. The two-scale realized volatility proposed by Zhang et al. (2005) accommodates some of the issues that are inherent to high frequency data.

Second, the higher variability of the GARCH-X forecasts can be explained by the more frequent changes in variables such as bid and ask quantities. The signs of the coefficients for quantities are more in line with expectations in hogs relative to corn, when the expectation is based on the assumption that trader strategies target either the first level or the second level when their perception of risk changes. Strategies could also target both levels at the same time, or both levels could be targeted by traders with different expectations, in which case the LOB would show increases or decreases in the quantities of both steps. Traders' strategies could also involve steps further away from the first two steps. Future work could use different variables such as quantities in further levels of the LOB or the ratio of quantities (however the ratio wouldn't account for the case when quantities in different levels increase or decrease by the same proportion).

Third, the forecasting performance of the estimated models could also be improved by updating the estimated models more often and including data from the same day for which the forecast is being made. Furthermore, the specification of each model could change for different days based on both goodness of fit measures and market information which may be known in specific days (for example, when public reports with specific information for each commodity are released).

Finally, another area to be explored in future research is the frequency of data sampling. This study followed one criteria based on transaction price changes. The snapshots duration used here may change when a different time period is used. Valuable information may be overlooked or discarded when sampling based on transaction prices. Realized volatility computed using more frequent snapshots or the full data could provide a different benchmark for forecasting performance evaluation. Future research could address in more detail the sampling frequency question.

References

1. Adjemian Michael, and Scott Irwin. 2018. USDA Announcement Effects in Real-Time. *American Journal of Agricultural Economics*, 100:1151--1171.
2. Ahn Hee-Joon, Kee-Hong Bae, and Kalok Chan. 2001. Limit Orders, Depth, and Volatility: Evidence from the Stock Exchange of Hong Kong. *The Journal of Finance*, 56 (2):767--788.
3. Aidov Alexandre, and Robert T. Daigler. 2015. Depth Characteristics for the Electronic Futures Limit Order Book. *The Journal of Futures Markets*, 35 (6):542--560.
4. Aït-Sahalia Yacine, Mykland Per A, and Lan Zhang. 2005. How Often to Sample a Continuous-time Process in the Present of Microstructure Noise. *The Review of Financial Studies*, 18 (2):351--416.
5. Aït-Sahalia Yacine, and Jialin Yu. 2009. High Frequency Market Microstructure Noise Estimates and Liquidity Measures. *The Annals of Applied Statistics*, 3 (1):422--457.
6. Amihud Yakov, and Haim Mendelson. 1987. Trading Mechanisms and Stock Returns: An Empirical Investigation. *The Journal of Finance*, 42 (3): 53--553.
7. Andersen Torben, Tim Bollerslev, Francis X. Diebold, and Paul Labys. 2003. Modeling and Forecasting Realized Volatility. *Econometrica*, 71 (2):579--625.
8. Andersen Torben G. and Tim Bollerslev.1998. Answering the Skeptics: Yes, Standard Volatility Models Do Prove Accurate Forecasts. *International Economic Review (Philadelphia)*, 39 (4):885--905.
9. Andersen Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens. 2001. The Distribution of Realized Stock Return Volatility. *Journal of Financial Economics*, 61 (1):43--76.

10. Arzandeh Mehdi, and Julieta Frank. 2019. Price Discovery in Agricultural Futures Markets: Should We Look beyond the Best Bid-ask Spread? *American Journal of Agricultural Economics*, 101 (5):1482--1498.
11. Barndorff-Nielsen Ole E., and Neil Shephard. 2002. Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society, Series B, Statistical Methodology*, 64 (2):253--280.
12. Barndorff-Nielsen, Ole E., and Neil Shephard. 2004. Econometric Analysis of Realised Covariation: High Frequency Based Covariance, Regression and Correlation in Financial Economics. *Econometrica*, 72:885--925.
13. Benavides Guillermo. 2004. Price Volatility Forecasts for Agricultural Commodities: An Application of Historical Volatility Models, Option Implieds and Composite Approaches for Futures Prices of Corn and Wheat. *SSRN*.
14. Black Fischer. 1986. Noise. *The Journal of Finance*, 41 (3):529--543.
15. Bollerslev Tim, and Melvin Michael. 1994. Bid-ask Spreads and Volatility in the Foreign Exchange Market: An Empirical Analysis. *Journal of International Economics*, 26(3):355-372.
16. Cao Charles, Oliver Hansch, and Xiaoxin Wang. 2008. The Information Content of an Open Limit Order Book. *The Journal of Futures Markets*, 29 (1):16--41.
17. Cermák Michal, Karel Malec, and Mansoor Maitah. 2017. Price Volatility Modelling-Wheat: GARCH Model Application. *AGRIS On-line Papers in Economics and Informatics*, 9 (4):15--24.

18. Chen Jian, Fuwei Jiang, Yangshu Liu, and Jun Tu. 2016. International Volatility Risk and Chinese Stock Return Predictability. *Journal of International Money and Finance*, 70:183--203.

19. Clinet Simon, and Yoann Potiron. 2019. Testing if the Market Microstructure Noise Is Fully Explained by the Informational Content of Some Variables from the Limit Order Book. *Journal of Econometrics*, 209:289--377.

20. Couleau Anabelle, Teresa Serra, and Philip Garcia. 2018. Microstructure Noise and Realized Variance in the Live Cattle Futures Market. *American Journal of Agricultural Economics*, 101 (2):563--578.

21. Couleau Anabelle, Serra Teresa, and Philip Garcia. 2020. Are Corn Futures Prices Getting "Jumpy"? *American Journal of Agricultural Economics*, 102 (2):569--588.

22. Degiannakis Stavros. 2017. The One-trading-day-ahead Forecast Errors of Intra-day Realized Volatility. *Research in International Business and Finance*, 42: 1298--1314.

23. Degiannakis Stavros, and George Filis. 2017. Forecasting Oil Price Realized Volatility Using Information Channels from Other Asset Classes. *Journal of International Money and Finance*, 76:28--49.

24. Deng Chunliang, Xingfa Zhang, Yuan Li, and Qiang Xiong. 2020. GARCH Model Test Using High-frequency Data. *Mathematics (Basel)*, 8:1922.

25. Engle Robert F. 1982. Autogressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50 (4): 987--1007.

26. Engle Robert F. 2002. New Frontiers for ARCH Models. *Journal of Applied Econometrics*, 17 (5): 425--446.

27. Foucault Thierry, Pagano Marco, Roell Ailsa. 1955. Market Liquidity: Theory, Evidence, and Policy. *Oxford Academic*.
28. Foucault Thierry, Sophie Moinas, and Erik Theissen. 2007. Does Anonymity Matter in Electronic Limit Order Markets? *The Review of Financial Studies,* 20 (5):1707--1747.
29. Francq Christian, and Thieu Le Quyen. 2019. Qml Inference for Volatility Models with Covariates. *Econometric Theory,* 35 (1):37--72.
30. Frijns, Bart, and Dimitris Margaritis. 2008. Forecasting Daily Volatility with Intraday Data. *European Journal of Finance,* 14 (6): 523--540.
31. Gousgounis, Eleni, and Esen Onur. 2018. The Effect of Pit Closure on Futures Trading. *Journal of Commodity Markets*, 10: 69--90.
32. Hagströmer Björn. 2021. Bias in the Effective Bid-ask Spread. *Journal of Financial Economics,* 142 (1):314-337.
33. Hansen Peter, and Asger Lunde. 2006. Realized Variance and Market Microstructure Noise. *Journal of Business & Economic Statistics*, 24 (2):127--161.
34. Harris Lawrence. 1990. Estimation of Stock Price Variances and Serial Covariances from Discrete Observations. *Journal of Financial and Quantitative Analysis*, 25 (3):291--306.
35. Harvey David, Stephen Leybourne, and Paul Newbold. 1998. Tests for Forecast Encompassing. *Journal of Business & Economic Statistics,* 16 (2):254.
36. He Xinyue, Teresa Serra, and Philip Garcia. 2020. Resilience in “Flash Events” in the Corn and Lean Hog Futures Markets. *American Journal of Agricultural Economics*, 103 (2):743--764.

37. Irwin, Scott, and Dwight Sanders. 2012. Financialization and Structural Change in Commodity Futures Markets. *Journal of Agricultural and Applied Economics*, 44(3):37--396.

38. Jain Pawan, and Christine Jiang. 2014. Predicting Future Price Volatility: Empirical Evidence from An Emerging Limit Order Market. *Pacific-Basin Finance Journal*, 27:7--93.

39. Kavajecz Kenneth A, and Elizabeth R. Odders-White. 2001. Volatility and Market Structure. *Journal of Financial Markets (Amsterdam, Netherlands)*, 4 (4):359--384.

40. Liu Jing, Feng Ma, Yang Ke, and Yaojie Zhang. 2018. Forecasting the Oil Futures Price Volatility: Large Jumps and Small Jumps. *Energy Economics*, 72:321--330.

41. Luo Jiawen, and Langnan Chen. 2019. Multivariate Realized Volatility Forecasts of Agricultural Commodity Futures. *Journal of Futures Markets*, 39 (12):1565--1586.

42. Luo Jiawen, Klein Tony, Ji Qiang, and Chenghan Hou. 2019. Forecasting Realized Volatility of Agricultural Commodity Futures with Infinite Hidden Markov HAR Models. *International Journal of Forecasting*, 4:1--23.

43. Manfredo Mark R, Raymond M Leuthold, and Scott H. Irwin. 2001. Forecasting Fed Cattle, Feeder Cattle, and Corn Cash Price Volatility: The Accuracy of Time Series, Implied Volatility, and Composite Approaches. *Journal of Agricultural and Applied Economics*, 33 (3):523--538.

44. Mcmillan David G, and Alan E. H. Speight. 2012. Daily FX Volatility Forecasts: Can the GARCH (1,1) Model Be Beaten Using High-frequency Data? *Journal of Forecasting*, 31 (4):330--343.

45. Meddahi Nour. 2002. A Theoretical Comparison Between Integrated and Realized Volatility. *Journal of Applied Econometrics (Chichester, England)*, 17 (5):479--508.
46. Musunuru Naveen, Mark Yu, and Arley Larson. 2013. Forecasting Volatility of Returns for Corn Using GARCH Models. *The Texas Journal of Agriculture and Natural Resources*, 26:42--45.
47. Oomen Roel C.A. 2005. Properties of Bias-Corrected Realized Variance Under Alternative Sampling Schemes. *Journal of Financial Econometric*, 3(4):555--577.
48. Oz Damla. 2022. Using the Information Contained in the Limit Order Book to Predict Volatility. *UofM Space*.
49. Pascual Roberto, and David Veredas. 2010. Does the Open Limit Order Book Matter in Explaining Informational Volatility? *Journal of Financial Econometrics*, 8 (1):57--87.
50. Shah Muhammad Kashif Ali, Zulfiqar Hyder, and Muhammad Khalid. 2009. Central Bank Intervention and Exchange Rate Volatility in Pakistan: An Analysis Using GARCH-X Model. *Applied Financial Economics*, 19 (18):1497--1508.
51. Tomek William, and Kaiser Harry. 2014. Agricultural Product Prices. *United States: Cornell University Press*.
52. Visser Marcel P. 2011. GARCH Parameter Estimation Using High-Frequency Data. *Journal of Financial Econometrics*, 9 (1):162--197.
53. Zhang Hanyu, and Alfonso Dufour. 2019. Modeling Intraday Volatility of European Bond Markets: A Data Filtering Application. *International Review of Financial Analysis*, 63:131--146.

54. Zhang Lan, Mykland Per A, and Aït-Sahalia Yacine. 2005. A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data. *Journal of the American Statistical Association*, 100 (472):1394--1411.

55. Zhou Bin. 1996. High-frequency Data and Volatility in Foreign-exchange Rates. *Journal of Business & Economic Statistics*, 14 (1):45--52.