THESIS

FORECASTING FED CATTLE PRICES: ERRORS AND PERFORMANCE DURING PERIODS OF HIGH VOLATILITY

Submitted by

Patrick B. Linnell

Department of Agricultural and Resource Economics

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2017

Master's Committee:

Advisor: Stephen R. Koontz

Joleen C. Hadrich
W. Marshall Frasier
Harvey Cutler

ABSTRACT

FORECASTING FED CATTLE PRICES: ERRORS AND PERFORMANCE DURING PERIODS OF HIGH VOLATILITY

Livestock and other commodity prices have displayed considerable volatility in the past ten years. In this environment, price forecasts play a key role in producers' business planning and risk management decisions. The object of this study is to evaluate fed cattle price forecasting performance and errors during this volatile period. Price forecast models are developed using autoregressive, vector autoregressive, and vector error correction frameworks. Forecast performance is compared to the live cattle futures market. Results emphasize the importance of simplicity relative to forecast accuracy. Autoregressive and vector autoregressive methods appear the most useful, with autoregressive models typically being the most accurate of the time series methods. Time series models are significantly more accurate than futures predictions at the one-month horizon. Futures are about as accurate or more accurate at all other horizons, especially as forecast horizon increases, although differences are not significant. Time series methods still provided valuable information relative to futures-based predictions at the two- to six-month horizons. Results suggest forecast errors are related to shocks occurring after the forecast, consistent with market efficiency. Shocks related to market currentness, or the relative supply and demand conditions of the non-storable commodity, appear the most important to fed cattle price forecasting errors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Agricultural commodities have seen considerable price volatility in the past decade. Corn, wheat, hay and cattle prices have all seen volatility levels two-to-three times higher than those levels observed at the end of 2016 and beginning of 2017. Fed cattle prices have become markedly more volatile, as can be seen in Figure 1. Sound business decision-making often depends on an expectation of future prices and producers contend that this market environment has made forming accurate price expectations exponentially more difficult (Gee 2016; Mulvany 2016). Although research has struggled to show alternative forecasting methods to be superior to futures market predictions (Garcia et al. 1988), many producers argue that even the futures market does not provide good forecasts, especially in light of the recent surge in volatility (Meyer 2016; Mulvany 2016).

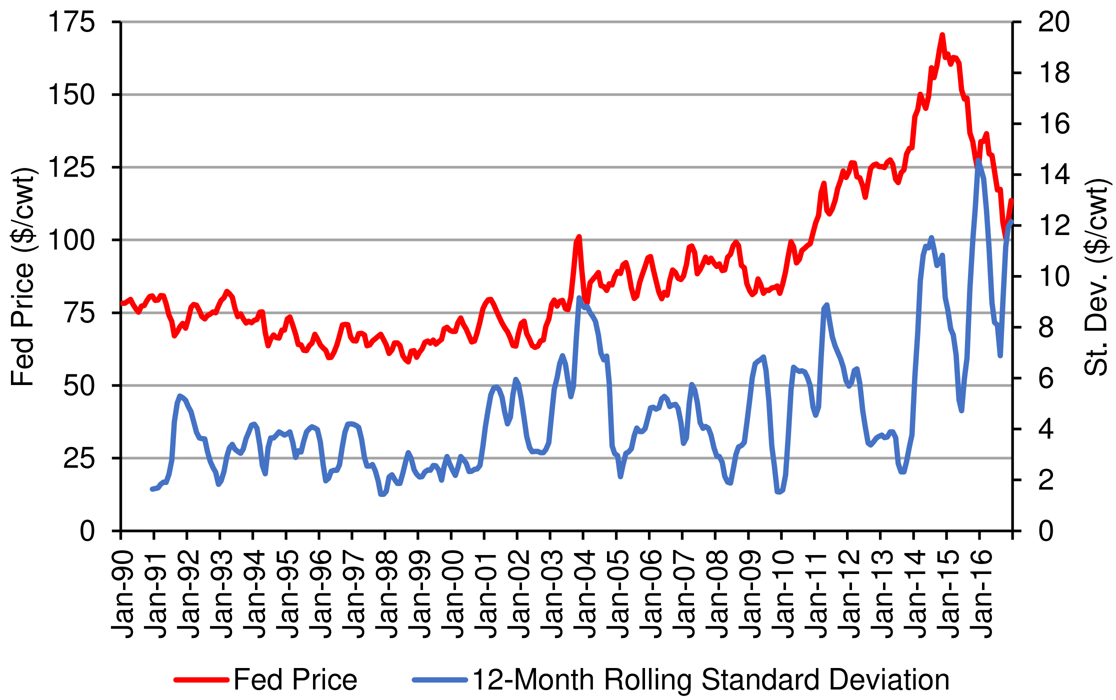

Figure 1: Monthly fed cattle price and standard deviation, 1990-2016

The press has documented the challenges producers have faced in recent years in navigating an environment of highly volatile prices (Meyer 2016). Price movements are expected as market fundamentals change with the cattle cycle. However, measures of cattle price volatility have recently reached levels unseen since the 2003 BSE discovery (Meyer 2016) and prices have regularly surpassed forecasts, in terms of both the magnitude and speed of price movements (e.g., LMIC 2016). Moreover, the cattle futures market has locked at its daily limit atypically often, posing significant planning and risk management challenges (Meyer 2016; Mulvany 2016).

Given this difficult market environment, it is easy to question the usefulness of forecasting prices – at the same time an accurate forecast becomes significantly more valuable. This introduces the question: how well have forecasting procedures performed in this volatile time period? With concerns regarding futures market efficiency, it is important to investigate the forecasting ability of the futures market as well as other forecasting techniques. Moreover, given that forecasts are always inaccurate to a degree, what factors are behind price movements that often exceed forecasts in upswings and downswings? The question becomes whether these forecast errors are consistent with the emergence of new information, consistent with economic theory and market history. This thesis seeks to provide answers to these important questions.

1.1 Objective

The object of this study is to evaluate fed cattle price forecast performance and errors during the recent period of high volatility. The surge in volatility provides a unique and important opportunity to evaluate the forecasting performance of commonly used time series methods and the futures market in a particularly challenging market environment. This work also seeks to evaluate how inaccuracies in price forecasts are related to extreme shocks or surprises to the underlying market fundamental and how they are related to the volatility of the market environment. In other words, are forecast errors persistent and systematic or can they be explained by changes in the market fundamentals, consistent with market history? Understanding the performance of forecast models in this climate has important implications for producers relying on forecasts for risk management and business planning decisions. Furthermore, an investigation of the relationships between forecast errors and changes in the underlying market fundamentals may also provide forecasters a focus for efforts to improve price forecast models.

1.2 Organization of the Thesis

The following chapter provides a review of the literature on cattle price forecasting. This is followed by a discussion of methodology used in this analysis in Chapter 3, including price forecasting model development and evaluation procedures. Chapter 4 outlines the data employed in this analysis, along with how we have accounted for a change in the USDA Cattle on Feed report. Chapter 5 presents the forecast models developed, an evaluation of the forecasting performance of the time series models and the futures market, and an evaluation of forecast errors. Lastly, Chapter 6 contains concluding remarks, implications of the analysis and suggestions for future research.

CHAPTER 2: LITERATURE REVIEW

This chapter is divided into four sections. Section 2.1 provides a background of the volatile market environment in the beef industry. Next, Section 2.2 provides an overview of the literature regarding fed cattle price forecasts, primarily focusing on the use of time series techniques and the live cattle futures market. Techniques used to evaluate forecasts are also discussed. Causes and implications of forecast errors with respect to fed cattle prices are discussed in Section 2.3. Finally, Section 2.4 concerns relevance of this work and its contribution to the literature.

2.1 Volatility in the Beef Industry

The U.S. beef industry has experienced considerable structural change in recent years. This structural change is composed of both sudden, unpredictable events as well as the continuation of gradual, long-term industry trends. Events such as the 2003 BSE case and the 2008 financial crisis and the ensuing recession have impacted consumer demand for beef (Pritchett *et al.* 2007, Darko and Eales 2013). Beef exports have also been impacted by a global recession, large swings in domestic beef prices and changes in the price of the dollar. Consolidation has continued in the cattle feeding and beef packing industries. Increased use of alternative marketing arrangements such as grids and formula pricing have drained the number of cattle trading in the cash market, posing market liquidity and price discovery concerns, especially in some regional markets (Koontz 2016). The cattle market is also not isolated from advances in technology. The liquidity concerns have given rise to online nationwide fed cattle auction platforms designed to provide an additional avenue of price discovery. Meanwhile, another technology—algorithmic futures trading—has been blamed as a contributor to increased volatility in futures markets (Gee 2016). This argument is not without merit, as changes to futures trading activity have been linked to cash market volatility (Yang, Balyeat and Leatham 2005).

In light of these confounding changes in the cattle markets and the price volatility often associated with them, producers have contended that forming accurate price expectations has become more difficult than usual (Mulvany 2016). With large price swings in the cash and futures markets that exceed forecasts, producers may lose confidence in forecasts, impeding the ability to make sound production decisions such as cattle feeding and placement decisions. Numerous articles in the press have documented problems the volatile market environment has posed for market participants. A short list includes Gee 2016, Meyer 2016, and Mulvany 2016. These articles cite hedgers losing confidence in timing of trades due to dramatic intra-week and intra-day price swings, impeding the risk management role of the futures markets. They also contend that large price movements seemingly unrelated to the market fundamentals have dampened the participation of some speculative traders who add important liquidity to the futures market.

2.2 Forecasting and Fed Cattle Prices: Methods and Overview

Forecasting has long been one of the primary tasks and most challenging undertakings of the agricultural economics field. After an initial focus on providing prescriptive solutions to farm business management and profitability, in a review of literature on agricultural forecasting, Allen (1994) described forecasting as the second phase in the progression of agricultural economics research. Forecasts are important for agricultural producers making production and business planning decisions, as well as guiding marketing and risk management strategies, which justifies the resources government and private entities expend on forecasting.

Early forecasts were primarily judgement-based and these expert opinion forecasts still play a significant role in short-term outlooks (Allen 1994). Many commonly cited and long running forecast series are considered outlook forecasts and contain a judgement-based component. Although research is mixed, the findings suggest that outlook series are not optimal and are no more accurate than futures-based forecasts (e.g., Sanders and Manfredo 2003; Colino and Irwin 2007). While not optimal on their own, expert opinion still appears to hold a certain value, especially in the context of combining expert opinion with quantitative forecasting methods and futures market information (Colino *et al.* 2012).

When presented with multiple forecasts, the natural inclination may be to ask which one is the most accurate, but users may be best suited to take some or all into account. Brandt and Bessler (1981) concluded that even if a forecast user had no prior knowledge to judge forecasts by, using a simple average of forecasts can be more accurate than the best individual forecast. Especially when forecast performance is unknown, compositing can take advantage of the strengths of each forecast while lessening the effects of any large mistakes of one forecast. More complex weighting procedures can be used if past performance of forecasts is known, but research has found the benefits to be generally small or inconclusive compared to a simple average in applications to cattle and soybean markets (Park and Tomek 1988) and hog markets (Colino *et al.* 2012). Colino *et al.* (2012) showed that although futures-based forecasts were the most accurate of individual forecasts, a composite of outlook, time series and futures-based forecasts had accuracy improvements, especially at longer horizons. While compositing procedures are not evaluated in this analysis, the prospect of composite forecasting emphasizes the importance of investigating and developing quantitative forecasting methods.

Section 2.2.1: Time Series Methods

Forecast methods have generally grown progressively more complex through the decades. Agricultural forecasting has long employed single equation econometric techniques but has expanded to include multi-sector, multi-equation models (Allen 1994). In recent decades, the focus in quantitative techniques has concentrated on time series techniques for forecasting prices of fed cattle and other agriculture commodities. These methods avoid the need to forecast independent variables prior to obtaining the forecast of interest by relying on only past data and known future data points. For example, time series methods would not first require forecasting of beef supplies or fed cattle slaughter numbers to derive price forecasts.

The Box-Jenkins autoregressive integrated moving average (ARIMA) technique is one of the simplest time series methods and has commonly been applied to cattle markets. ARIMA models have a univariate specification that models price as a function of past values of themselves and a moving average component. Since early application by Oliveira, O'Connor, and Smith (1979), ARIMA methods have often been applied and studied in comparison to other forecast methods due to the technique's favorable combination of a simple specification and relatively accurate forecast performance. As found by Oliveira, O'Connor and Smith, ARIMA models generally perform well at short forecast horizons. Showing the value of simple time series techniques, Sanders and Manfredo (2003) concluded that USDA fed cattle price forecasts could generally be improved by compositing these forecasts with predictions from a simple time series alternative such as the AR(4) model that they compared with USDA forecasts. ARIMA techniques applied to disaggregated time series have also been shown to more accurately forecast prices in more aggregated time periods than models estimated at the higher level of aggregation in time, as shown in recent work by Pena-Levano, Ramirez and Renteria-Pinon (2015). These and other works reinforce the value in a forecasting context of capturing the patterns and characteristics of a series over time, even in a relatively a-theoretic framework. We find it important to evaluate the performance of ARIMA models in forecasting fed cattle prices as part of this analysis.

Multiple equation time series techniques have commonly involved applications of vector autoregressive (VAR) models and vector error correction (VEC) models. Equations estimated in the VAR model set each dependent variable as a function of past values of itself, past values of the other endogenous variables in the system and exogenous variables, which usually includes seasonal dummy variables and a trend. VEC models are specified similarly to VAR models, but take into account long-run equilibria between variables. Many time series are non-stationary and therefore must be estimated in first differences. However, there can be important relationships between the levels of variables and using differenced data can ignore these important long-run relationships in the VAR models (Gujarati and Porter 2009, p. 788). The VEC specification includes these relationships by re-incorporating the data in levels (Johansen 1995).

An important component of multiple equation time series models is determining what variables are endogenous to the systems. The variables are the same as would be included in an econometric model of the market and as a result the VAR and VEC models in the literature often include many of the same variables. Zapata and Garcia (1990) based their VAR and VEC specifications on the econometric model of Garcia et al. (1988), using fed cattle prices, feeder cattle prices and per capita income as endogenous variables and dropping corn price due to insignificance. Park (1990) used fed cattle price, feeder cattle price, beef production and total cattle on feed in evaluating five types of multiple-equation time series models. Goodwin (1992) included prices of fed cattle, hogs and broilers, total cattle on feed, corn prices, and disposable income in multiple-equation models in an analysis of forecasting methods in the presence of structural change. The works listed above and others guided the choice of variables in developing the VAR and VEC models in the present analysis.

Multiple equation time series methods are common in macro-economic applications and in other commodity markets where several variables are expected to influence one another over time. One reason for the widespread application is simple: multiple equation time series models tend to forecast well. For instance, Beckers and Beidas-Strom (2015) developed VAR models that could out-forecast futures market predictions of monthly oil prices out to 24 months. Structural VAR models are frequently used to place *a priori* restrictions on the relationship between variables in the model based on economic theory (Brown and Yucel 1999; Kilian 2009). Multiple-equation models are also useful in analyzing market shocks. Impulse response functions and variance decompositions of VAR's are used to study how shocks to one market are expected to influence other variables in the model and understand the sources of error variance for each variable (Ratti and Vespignani 2016; Kilian 2009; Brown and Yucel 1999). However, impulse response functions and variance decompositions limit error and shock analysis to the variables included in the model. As described later, a key difference with the present research is the use of a different process to evaluate forecast errors against shocks to variables not necessarily included in the forecast model.

Section 2.2.2: Futures Markets and Forecasting

In addition to applications for risk management and decision-making, price discovery is an important role of commodity futures markets since they can be considered a forward-looking market consensus of what prices will be in the future (Purcell and Koontz 1999, p. 11). In fact, predictions from the live cattle futures market have consistently been found to forecast as well or better than other forecasts. Colino and Irwin (2010) considered futures as the “gold standard” with which to compare other forecasts, subsequently finding futures to perform as well or better than university outlook forecasts for hogs and cattle prices. Kastens, Schroeder and Plain (1998) reached similar conclusions comparing futures with university extension and USDA livestock price forecasts. They concluded that futures provided a reasonable substitute for extension forecasts of livestock prices, although extension forecasts were marginally more accurate for fed cattle prices. Both futures and extension, however, appeared decisively better than USDA forecasts for cattle, broiler and hog prices. While it is not clear how these public forecasts are generated, but it can be assumed that they likely include some combination of econometric supply and demand models, time series methods and expert opinion. Bowman and Husain (2004) compared forecasts of spot prices that included futures in an error-correction framework against univariate and judgement-based forecasts. Although livestock prices series were not evaluated, they found that including futures in the models generated the most accurate forecasts for prices 15 different commodities.

Futures may provide accurate forecasts from a comparative standpoint, but many have argued that futures still do not provide efficient forecasts and are certainly not without their deficiencies. Reviewing a number of works on the use of commodity futures for forecasting, Tomek (1997) concluded that futures-based forecasts provided relatively poor forecasts, although quantitative methods could not generally forecast better than futures. Indeed, multiple studies have found live cattle futures to be inefficient predictors. Martin and Garcia (1981) tested four separate hypotheses about the price forecasting performance of live cattle futures. The questions regarded potential changes in forecasting performance over time, with cyclical price variations, with seasonality, and in unstable compared to stable economic conditions. Live cattle futures failed all four tests of forecast performance and generally forecasted no better than lagged cash prices. Similarly, Leuthold and Hartmann (1981) found the forward-pricing ability of live cattle futures has periodically performed inefficiently and in these periods did not forecast better than lagged cash prices.

On the contrary, Garcia et al. (1988) argued that there was not sufficient evidence to prove inefficiency in the futures market due to lack of abnormal profits earned in application of other, more accurate forecasts. When comparing fed cattle price forecasts from econometric, ARIMA, futures-based, and composite forecast methods, one and often more forecast methods were more accurate than futures market predictions according to statistical measures. However, in a simulated trading application, the more accurate forecasts generated only small and highly variable profits. The authors argue that because the risk-adverse trader would not be willing to accept these large risk-return ratios, market inefficiency could not be concluded.

As summarized above, the literature has reached mixed conclusions on the forecasting efficiency of the live cattle futures market. However, even if futures-based predictions are not good, there appears to be a large consensus that the futures are generally about as good as it gets. Consistent with the bulk of the literature, in this thesis futures are also considered the gold standard and the baseline to which other forecasts are compared.

Section 2.2.3: Evaluating and Comparing Forecasts

Just as one would expect forecast users' definition of usefulness to vary, so does the means to measure and compare forecasts. Commonly used methods involve statistical accuracy criteria, forecast encompassing, forecast efficiency tests and utility measures among others. Since only statistical criteria and forecast encompassing tests will be used in this analysis, this review will be limited to those techniques.

Kastens, Schroeder and Plain (1998) described four types of statistical criteria that provide different information about forecast errors: bias, ratio-type, volume-type and fit. Examples of each type include mean error, mean absolute percent error (MAPE), root mean squared error (RMSE) and squared linear correlation coefficient (R^2), respectively. By measuring accuracy differently, these statistical measures provide different information regarding forecasts. Due to greater penalty imposed on large forecast errors, RMSE is one of the most commonly used measures (Kastens, Schroeder and Plain 1998). RMSE is the primary statistical accuracy measure used in the present analysis due to its penalty on large errors and the availability of statistical tests for differences. MAPE is also used as a secondary criterion in this analysis, and calculation of both measures is outlined later. To test the statistical significance of differences in forecast accuracy, tests such as the modified Diebold-Mariano (MDM) test developed by Harvey, Leybourne, and Newbold (1997) can be used. The MDM test uses a quadratic loss function to test for differences in accuracy between two forecasts. Use of the MDM test is described further in the methodology.

Granger and Newbold (1973) first showed that, given two forecasts, it is possible for the less accurate forecast to still contain valuable information relative to the preferred forecast. Following this concept, Harvey, Leybourne and Newbold (1998) developed a forecast encompassing test to determine if a preferred forecast entirely encompasses all information provided in an alternative, less accurate forecast. The encompassing test is based on the idea that a forecast is encompassed by the preferred forecast if the optimal weight of the alternative forecast is zero in a weighted average of two forecasts. In this way, encompassing tests are useful in determining where compositing separate forecasts may be beneficial to forecast accuracy.

No tests of forecast encompassing between fed cattle price forecasts from time series methods and futures markets appear in the literature, however, encompassing tests have been used to compare outlook forecasts with alternative forecasts. Comparing university outlook forecasts to futures market predictions, Colino and Irwin (2010) concluded that futures market predictions did not encompass all the information in outlook forecasts of fed cattle prices. Sanders and Manfredo (2003) concluded that users of USDA cattle price forecasts may want to supplement them with a time series alternative because USDA forecasts did not encompass an AR(4) alternative. In an application that more closely aligns with this research, Sanders and Manfredo (2005) used multiple forecast encompassing to test the efficiency of the fluid milk futures market by comparing it with two simple time series alternatives and USDA milk price forecasts. Their research concluded that the futures market did not encompass all the information provided in USDA forecasts at a two-quarter forecast horizon.

2.3 Forecast Errors and Market Shocks

An important part of this analysis is the evaluation of forecast errors. Forecast errors can be attributed to two distinct reasons: failure to incorporate all relevant information and changes to the underlying assumptions built into the forecast. The first is a forecast efficiency issue and the second reflects an efficient forecast, but the two are not mutually exclusive. Reasonably, the exclusion of relevant information and changes to important variables included in the model may simultaneously contribute to forecast error. While both are important forecasting efficiency questions, in this analysis we focus on changing information relative to forecast model assumptions.

By making an analogy to finance theory, Nordhaus (1987) argues that a forecast is efficient if it minimizes the loss function of the forecast (i.e. the forecast error) with respect to all information available at the time the forecast is generated. Therefore, forecast errors would reflect the workings of an efficient market if they are related to surprises or unexpected changes in the market compared to the information that forecasts are based on. By definition, a market shock is a random event that cannot largely be anticipated. When shocks occur, a sound forecast based on the best available information can still result in significant errors. Therefore, forecast errors can be directly related to random market shocks. Outlining this concept, Nordhaus (1987) showed that if a forecast is efficient, then successive revisions of a forecast are a random walk as market shocks occur and forecasts are adjusted with the new information. The same concept can be extended to the relationship between changes in futures market prices and changes to the expectations of underlying market conditions.

Shocks to fundamental market conditions are often cited for price forecast errors or prices moving to more extreme levels than anticipated. Analysis in corn markets has shown that shocks to market-specific fundamentals (stocks-to-use ratios) and residual shocks are important to corn price movements and are large sources of forecast error variance (Etienne, Irwin and Garcia 2014). A similar fundamental shock in beef markets would be supplies of beef that were substantially larger than anticipated by forecasters, likely causing realized cattle prices to be lower than initially forecasted. This scenario has been cited for recent price deteriorations that far exceeded expectations (LMIC 2016). Analysts have also commonly cited export demand and supplies of beef and substitute meats as critical assumptions in market outlooks (e.g., Bechtel 2017 and LMIC 2016). Changes to these variables relative to expectations have clear implications to forecast accuracy and forecast error.

Section 2.3.1: Currentness of the Fed Cattle Market

As a non-storable commodity, the "currentness" of the fed cattle market can have an important impact on prices (LMIC 2016). The relationship between currentness and prices in the fed cattle market follows a well-known narrative. As the pace of fed cattle marketings for harvest unexpectedly slow, cattle spend more days on feed. The feedlot operator is limited on the additional time cattle can be held before they need to be marketed. Once the feedlot gets behind on marketings, bargaining position is lost as the non-storable commodity needs to be processed. As a result, the meatpacker gains leverage and the short-term market power allows prices to be pushed downward. Larger carcass weights result from additional days on feed, potentially contributing to packer bargaining position as more pounds of beef per head partially offsets the number of head harvested. Of course, the opposite of this scenario can result in higher than expected prices as the feedyard's bargaining position is improved with increased currentness.

2.4 Contributions to the Literature

The surge in volatility in recent years provides a unique opportunity to evaluate forecasting performance of time series methods and futures markets in a particularly challenging market environment. In this context, the present study provides an update to previous work. The more novel and arguably more significant contribution of this work is the investigation of errors from both the time series forecasts and futures market predictions. This work seeks to understand which market shocks are the most significant drivers of forecast errors. This contributes to the understanding of the nature of the uncertainty related to forecasting cattle prices and how forecast errors are related to shocks to underlying fundamentals in the market. Explanation of forecast errors may also provide forecasters with a direction to focus in improving forecast models. This has important implications to users and producers of cattle price forecasts alike.

CHAPTER 3: METHODOLOGY

Fed cattle price forecast models were developed using time series methods. Forecast models were first constructed in a simple framework and complexity was added to the extent that forecast performance improved. Variables added are grounded in the functionality of the beef markets and found in previous literature. Stationarity tests are discussed in Section 3.1. Section 3.2 describes the process of candidate forecast model development and selection of candidate models for extensive evaluation. Next, Section 3.3 discusses the model specifications used in this analysis. In Section 3.4, the incorporation of futures market information into forecasts is discussed. Forecasts of feedlot placement numbers are needed for some price forecast models. The procedures used to forecast placements are described in Section 3.5. Next, the methods used to evaluate the performance of price forecasts are outlined in Section 3.6. Finally, Section 3.7 discusses the methods for investigating and explaining forecast errors.

3.1 Stationarity Tests

Data series are tested for stationarity prior to model estimation using the Augmented Dickey-Fuller (ADF) test. Stationarity of a series is defined by a constant mean and variance over time. Ensuring stationarity prevents spurious regressions and is important for empirical work like hypothesis testing (Gujarati and Porter 2009, p. 737). If the data is nonstationary, this indicates the series is a random walk and will need to be differenced to achieve stationarity. The general form of the ADF test is:

(1) $\Delta y_t = \beta_0 + \beta_1 t + \delta y_{t-1} + \alpha_1 \Delta y_{t-1} + \cdots + \alpha_p \Delta y_{t-p} + e_t.$

where Δ is the first difference operator, y_t is the series being tested for stationarity, β_0 is a drift term, β_1 is the coefficient on a time trend and p is the lag order of the autoregressive process. A lag order should be that is sufficiently long enough to remove all autocorrelation in the error term, e_t. In this analysis, the lag order is selected by the Akaike information criterion (AIC). The α_i's are the coefficients on the lagged differenced values included to remove autocorrelation and are not tested. Under the null hypothesis of nonstationarity, δ is equal to zero. Failing to reject the null hypothesis indicates a random walk series. By rejecting the null hypothesis, we can conclude that the series is stationary and does not need to be differenced. One criticism of stationarity tests is their low power to distinguish between different forms of nonstationarity (Gujarati and Porter 2009, p. 759). Unless there is an expectation of a trend, the trend coefficient β_1 is restricted to zero. All series that fail to reject nonstationarity at the 5% level are first differenced and retested for stationarity of the transformed series.

This work considers a variety of data series in forecasting fed cattle prices and analyzing forecast errors. However, only the variables considered for use in forecasting models are tested for stationarity and, if necessary, differenced. As we show later, the stationarity of the data series used in the error analysis is not important because they are not used in their original form.

3.2 Candidate Model Development and Selection

Multiple candidate forecast models are developed in this analysis and only the best, most accurate models are selected for extensive performance evaluation. Multiple model specifications are tested and considered because the model that best explains the underlying data generating process may be different than the model that forecasts future values most accurately.

Candidate models are first developed in a simple construct and complexity is added to the extent that forecasting performance is improved. Complexity is introduced by including additional variables in forecast models and by moving to more robust model specifications. Variables included in candidate models are based on the literature and the workings of the market. We experiment with different combinations of variables in different model types to develop the most accurate forecast models. Because the purpose of this work is forecasting, a candidate forecast may be selected for evaluation based on accurate forecasting regardless of the significance of the variables in the model. Model specifications are described thoroughly in the following section but are discussed in general terms here as they relate to the model development process.

First, univariate forecast models are estimated in which fed cattle prices are modeled strictly as a function of past values of themselves. Next, predetermined, exogenous information is introduced in the form of month dummy variables to capture seasonality. Futures market information and lagged cattle on feed placements are also experimented with in candidate models as predetermined, exogenous variables.

Multiple-equation time series models introduce another level of complexity. Fed cattle prices and other variables are estimated as a function of lagged values of themselves and the other variables in the system. Various combinations of these endogenous variables are tested, as well as with combinations of exogenous variables. Multiple-equation models with error correction terms to account for long-run equilibria are also tested.

Model specifications consistent with the properties of the time series are important for accurate forecasting (Zapata and Garcia 1990). When structural change is believed to be present, it is also important to use only the most recent data to model and forecast the relevant underlying data generating process (Clark and McCracken 2009). Fixed-estimation and rolling-estimation schemes were used in this analysis to address these two concerns. Ideally, model specification would be investigated for significant variables and autocorrelation structure with each iteration of forecasting, but this is unrealistic considering our long post-sample period of 120 observations. Forecast models were first developed and compared in a single-estimation framework where the model is estimated once over the in-sample period and used to forecast throughout the post-sample period. Initial models were examined for fit and coefficients were tested for significance to develop models that fit the properties of the data. From these initial models, subsequent candidate models were developed by adding and removing variables to the extent that forecasting performance was improved.

Some candidate models were then selected for further evaluation in a rolling-estimation framework where coefficients and the lag order of the model could change as the estimation window shifted forward through time. Lag order was selected with each iteration of estimation based on the Schwarz information criterion. The updated models were then used for forecasting at each step forward through the post-sample period. These selected candidate models were chosen based on accuracy at one or more forecast horizons. Models selected for re-estimation in a rolling framework included the best model of each model type. Variations to the best models that were also selected subjectively for rolling estimation.

The best candidate models from the rolling framework at each forecast horizon are selected for extensive evaluation described in Section 3.6. These include the best model of each specification type as well as futures market-based forecasts. Criteria used to define accuracy are also described in Section 3.6.

3.3 Forecast Model Specifications

The price forecasting models tested are autoregressive integrated moving average (ARIMA), vector autoregressive (VAR), and vector error correction (VEC) model specifications. Multiple models of each specification are developed and tested for forecasting performance. Futures market predictions are also evaluated for comparative purposes. Model specifications are based on models used in the literature and further models are developed by adding and removing potentially relevant variables to the extent that forecasting performance is improved. Accurate forecasting is dependent on identification and selection of a model consistent with the system's properties (Zapata and Garcia 1990). Since the purpose of this work is forecasting, candidate model specifications that do not necessarily fit the system's characteristics may still be used based on accurate forecasting ability. The following sections outline the basic model specifications employed. A comprehensive list of the exact model specifications of each candidate model developed can be found in Appendix B.

Section 3.3.1: Autoregressive Integrated Moving Average Model (ARIMA)

First, ARIMA models were developed. The univariate ARIMA model is the simplest model specification of the models developed and is expressed as:

(2) $y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \, y_{t-i} + \sum_{j=1}^{q} \theta_j \, u_{t-j} + BX_t + u_t$

where y is a stationary series, α_0 is an intercept term, the α_i's are the coefficients on lagged values of y where the autoregressive process is of order p, the θ_j's are coefficients on the lagged error terms of the moving average process of order q, B is a vector of coefficients on the exogenous variables in vector X, and u_t is the error term. The Box-Jenkins methodology is used to determine the values of p and q through the examination of autocorrelation functions and partial autocorrelation functions to identify the autoregressive and moving average properties of the series[1]. If the series y is nonstationary in its original form, it must be differenced until it is stationary. The series is said to be integrated of order d if it must be differenced d times to achieve stationarity. The exogenous variables in X are deterministic predetermined variables that

[1] ARIMA models are referred to in the results section as AR models because autocorrelation and partial autocorrelation functions indicated no moving average (MA) component.

have a value that is known at the time of forecasting. This often includes a trend or observation number and seasonal dummy variables. Futures market information and lagged placement data are considered for inclusion in *X* since this is information known prior to forecasting and can be considered deterministic. Use of futures and placement data is further discussed later. Exogenous information is used to the extent that forecast performance is improved.

Section 3.3.2: Vector Autoregressive Model (VAR)

VAR models are also of considerable interest in forecasting. Variables in the VAR model are specified as linear functions of lagged values of themselves and the other endogenous variables in the system, as well as exogenous variables. The VAR model with *K* endogenous variables is modeled as:

(3) $Y_t = A_0 + \sum_{i=1}^{p} A_i Y_{t-i} + BX_t + u_t$

where Y_t is a vector of stationary time series in the system, A_0 is a vector of constants in each equation, the $\sum_{i=1}^{p} A_i$ are $K \times K$ matrices of parameters on the lagged endogenous variables with lag order *p*, *B* is a $K \times M$ matrix of parameters on the vector of *M* exogenous variables, X_t, and u_t is the error term. The variables used in the candidate VAR models are selected based on previous literature and an *a priori* expectation that the variables are influential to fed cattle prices. Exogenous variables considered for inclusion in X_t are the same as with the ARIMA models described in the previous section.

Section 3.3.3: Vector Error Correction Model (VEC)

In cases where endogenous variables in the VAR model are differenced to achieve stationarity, long-run equilibria between the level forms of the variables may be ignored. The VEC model is a variation on the VAR model that includes an error correction term to account for important long-run relationships. The equilibrium conditions imposed by the VEC can be important to accuracy in long-run forecasts (Zapata and Garcia 1990). The equation for the VEC is expressed similarly to the VAR in Equation (3) except first differenced series are explicitly denoted with the first difference operator, *Δ*. Series not preceded by *Δ* are used in level form. The VEC model is expressed as follows:

(4) $\Delta Y_t = A_0 + \sum_{i=1}^{p} A_i \, \Delta Y_{t-i} + \Pi Y_{t-1} + BX_t + u_t$

where ΔY_t is a vector of differenced time series, A_0 is a vector of constants in each equation, the $\sum_{i=1}^{p} A_i$ are $K \times K$ matrices of parameters on the lagged differences of the endogenous variables with lag order *p*, *Π* is a $K \times K$ matrix of parameters on the lagged levels of the endogenous variables, *B* is a $K \times M$ matrix of parameters on the vector of *M* exogenous variables, X_t, and u_t is the error term. The coefficients in *Π* are derived from $\Pi = \alpha\beta'$, in which *α* and *β* are derived from the cointegrating equation on the levels of the endogenous variables:

(5) $y_{1t} = \alpha + \beta_1 y_{2t} + \ldots + \beta_{k-1} y_{kt} + e_t$

where the y_{kt} are the *K* variables in the Y_t matrix and e_t are the errors from the cointegrating equation (Johansen 1995). The same set of endogenous and exogenous variables are considered in the VEC models as in the VAR specification.

Section 3.3.4: Futures Implied Model

The futures market should provide an unbiased predictor of cash market prices, assuming cash market price is equal to the futures price at the time of contract expiration. Furthermore, since the futures market is widely used by hedgers and speculators alike, it can be considered a market consensus on future price levels (Purcell and Koontz 1999, p. 11). The futures market is used as a baseline with which to compare forecasting models. If the econometric forecasting models consistently out-perform the futures market, this would imply a clear market inefficiency. Although this was not expected to be the case, the futures market predictions can also be investigated alongside other models for causes of forecast error. The futures implied model is represented as:

(6) $pfed_t = LC_{t-p}^t + u_t$

where fed price, *pfed*, at time *t* is equal to the futures price *p* periods ahead for the live cattle contract (*LC*) expiring at time *t* and u_t is the error term. This equation implies that there will be a zero-level basis (i.e. the difference between cash and futures prices). While this assumption could be improved, it is reasonable to expect cash prices be equal to futures during the expiration month. Garcia et al. (1988) used a similar assumption, finding that cash prices and futures prices were not significantly different. Since every other month has a futures contract, if no contract expires during a given month, the price of the next closest futures contract will be used and considered the forecast for that month, also assuming a zero basis.

3.4 Futures Data in Fed Cattle Forecasts

As the market consensus price expectation, futures information may be important to include in times series forecast models to improve forecast performance. However, the nonstationarity of futures prices and fed cattle prices (as shown below in Section 5.1 Stationarity Tests) complicates the inclusion of the future market predictions. Since the two series must be differenced, adding futures price as an independent variable does not tie predictions to the actual futures price because predictions are in differences and not levels. If differenced futures prices were used, we would be modeling changes in fed prices as a function of past changes in futures prices. In other words:

(7) $\Delta pfed_t = f(\Delta LC_{t-h}^t, \dots)$

where *t-h* is the period from which future prices are used to predict fed prices in time period *t*. This ignores the information of interest, which is the future price levels predicted by futures market data. Instead, a variable is constructed based on the difference between cash prices and futures prices. This variable, *fut*, is constructed as:

(8) $fut_t^{t+h} = pfed_t - LC_t^{t+h}$

where t is time, h is forecast horizon, and $t+h$ is the period being forecasted. The futures variable can be interpreted as the level cash prices will need to change to reach the price implied by futures market predictions. This is based on two assumptions. First, that current futures price for a given contract will be equal to the price of that contract when it is the nearby contract. Second, this assumes a zero-basis level during the time that the contract is the nearby marketing contract. These assumptions are summarized by Equation (9):

(9) $fut_t^{t+h} = fut_{t+h}^{t+h} = pfed_{t+h}$.

In other words, for a given futures contract that is the nearby contract for period $t+h$, the price of that futures contract in time t will be equal to its price in time $t+h$. Prices of the given futures contract are assumed to be equal to cash prices in time period $t+h$. It may be helpful to illustrate the use of this variable in practice. Equation (10) shows this variable in a simple AR model:

(10) $\Delta pfed_t = \beta_0 + \beta_1 \Delta pfed_{t-1} + \beta_2 fut_{t-1}^{t} + u_t$

where changes in fed price are a function of a drift term (β_0), the prior period change ($pfed_{t-1}$), the futures variable (fut^t_{t-1}) and an error term. Substituting the right-hand side of Equation (9) for the futures variable yields Equation (11):

(11) $\Delta pfed_t = \beta_0 + \beta_1 \Delta pfed_{t-1} + \beta_2 (pfed_{t-1} - LC_{t-1}^{t}) + u_t$.

Here, it can be seen more clearly that changes in fed price in period t are a function of the level futures prices in period $t-1$ implied that fed price would change into period t. Changes in fed price are modeled as a function of the difference between prior period prices and what futures in the prior period suggested prices would be in period of interest.

3.5 Placements of Cattle on Feed Forecasts

One benefit of time series techniques used in this analysis is the reliance on past values, minimizing the need to use other predicted values as independent variables in the forecast models. However, number of placements of feeder cattle on feed is one independent variable that will require an assumption of future values for long-horizon forecasts of fed cattle prices. The maximum forecast horizon we evaluated is nine months, while feeder cattle are typically placed on feed four to six months on feed before marketed as fed cattle. The prices forecasted at the maximum forecast horizon are for cattle that have primarily not yet been placed on feed. Since this research aims to evaluate the performance of fed cattle price forecasts in a purely ex ante context, forecasts of cattle on feed placements are generated rather than using actual placement data when that data would not yet be known.

Placements are forecast according to the model:

$$plmt_t = \beta_0 + \sum_{i=1}^{p} \beta_i \, plmt_{t-i} + \gamma cows_{t-6} + u_t \quad (12)$$

where $plmt_t$ are placements in time period t, β_0 is the constant, $\sum_{i=1}^{p} \beta_i$ are coefficients on the lagged values of placements, the values of p describe the lag structure of the autoregressive process of placements, γ is the coefficient on lagged beef cow inventory levels[2]. The values of i are determined through examining autocorrelation and partial autocorrelation functions. Beef cow inventories are included to tie placements to the number of beef cows calving. This is an

[2] Lagged corn prices and feeder steer prices are found to be insignificant and did not contribute to forecast accuracy so these variables are dropped from the placement forecast model.

annual number that is an observation as of January 1 each year. A six-month lag is used to reflect the biological lag in the production process where calves enter feedlots at the time of weaning at six to eight months of age or are grown on grass after weaning before being placed on feed. Due to this production lag, cattle placed on feed in the first half of the year are related to the prior year's beef cow inventory and cattle placed during the second half of the year are related to the current year's January 1 beef cow inventory number. Using this model, placements are forecasted one- through six-months ahead for use in fed price forecast models.

3.6 Forecasting and Performance Evaluation

The in-sample period in which candidate forecasting models are fit is from January 1990 through December 2006. Forecasts are evaluated over the post-sample period from January 2007 through December 2016, this being the more volatile time period. Forecasts are generated for one- to nine-month forecast horizons for each month within the post-sample period. Forecasts are estimated in a fixed-estimation and rolling-estimation framework to first determine the appropriate structural model, then to fit the model appropriately to the data over time.

In the first, forecast models are estimated only once over the in-sample period from January 1990 to December 2006. This single estimation of the forecast model is used to generate forecasts over the entirety of the post-sample forecasting period. Hence, parameter estimates are held constant for each iteration of forecasting even as new, more recent observations become available with each step through the post-sample period. A variety of variable combinations are tested to determine what structural forecast models perform best.

Second, models developed in the fixed framework are re-estimated over a rolling estimation window where the oldest observation is removed as a new observation becomes available, holding the estimation period at a constant number of observations. The first iteration of these models will be identical to the fixed estimation models, but may change in subsequent estimations as the estimation window shifts forward in time. The rolling update of forecast models should better fit the data over time and reflect forecast methods employed in practice as practitioners update their models over time.

Once forecasts are generated, forecast performance is evaluated by comparing price forecasts to actual price with root mean square error (RMSE) and mean absolute percent error (MAPE) criteria. The RMSE is equal to the square root of the mean square error (MSE) and is given by the following equation:

$$(13) \quad RMSE = \sqrt{MSE} = \sqrt{\sum_{t=1}^{n} \frac{(y_t - \widehat{y_t})^2}{n}}$$

Like MSE, the RMSE is a measure to compare accuracy across models and penalizes larger errors by squaring the error terms, but the RMSE has that advantage of being in the same units as the data. The MAPE is also evaluated to account for error size relative to differing price levels over time. The MAPE is given by:

$$(14) \quad MAPE = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{y_t - \widehat{y_t}}{y_t} \right|$$

The MAPE penalizes all percentage errors equally and avoids penalizes errors during periods of higher prices more heavily, a commonly cited pitfall of RMSE evaluations (Kastens, Schroeder and Plain 1998).

In this analysis RMSE is used as the primary evaluation measure to select forecast models that minimize large forecast errors. The MAPE is evaluated as a secondary criterion to validate and supplement evaluation by RMSE. Unless there is a large discrepancy, MAPE will only be used to ties between candidate forecasts with similar RMSE's. For each horizon, forecasting models of each specification performing the best according to these metrics will be selected for performance evaluation and error investigation.

It is important to determine is differences in forecasting accuracy are statistically significant. The differences in accuracy can be tested with the modified Diebold-Mariano (MDM) test proposed by Harvey, Leybourne and Newbold (1997). The null hypothesis of the test is equal forecast accuracy; rejecting the null allows us to conclude that differences are statistically significant. Assuming a quadratic loss function, the test is based on the difference between squared errors of two forecasts:

(15) $d_t = u_{1t}^2 - u_{2t}^2$

where u_{1t}^2 and u_{2t}^2 are the squared errors from two separate forecasts at the same forecast horizon. The MDM test, which will not be specifically described here, is conducted on d_t and follows a t-distribution.

Section: 3.6.1 Forecast Encompassing Tests

Forecast encompassing tests are used to determine if an alternative, less accurate forecast contains incremental information not found in a superior forecast. In this analysis, we use the encompassing test proposed by Harvey, Leybourne, and Newbold (1997). This test is based on the idea that if one forecast encompasses another, then the optimal weight of the inferior forecast would be zero in a composite. This concept can be represented in the regression:

(16) $u_{1t} = \lambda(u_{1t} - u_{2t}) + e_t$

where u_{1t} are the errors from superior forecast and u_{2t} are the errors from the alternative forecast and λ is the coefficient to be estimated. The null hypothesis for the encompassing test is that λ = 0, which can be tested with a standard t-test. The rejection of the null hypothesis implies that the alternative forecast contains incremental information not contained in the superior forecast and that greater forecast accuracy could be achieved by a composite of the two forecasts.

Forecast encompassing tests are used to test two concepts. First, encompassing tests are used to determine if the time series techniques provide incremental information to futures predictions. To test this, encompassing tests are conducted pairwise between futures market errors and the errors from each type of time series forecast. Even if futures are more accurate than forecasts from time series methods, there still may be valuable information captured by time series forecasts that is not contained in futures market predictions. In this case, even though the time series methods are inferior forecasts when considered separately, they should not be discarded but rather used in a composite with futures market predictions.

Similarly, encompassing tests are conducted between the most accurate time series forecast and the other two time series forecasts. If a simpler time series model is more accurate, is there incremental information in more complex models? Or if a more complex model is more accurate, does a simpler time series model offer additional forecasting value in a composite forecast? The important question in this context is whether different time series methods offer incremental value to one another or if all the time series techniques we evaluate capture the same information.

Section: 3.6.2 Forecasting with Regional Fed Cattle Prices

Differing geographical, seasonal and market structure characteristics create regional differences in fed cattle prices. The USDA Agricultural Marketing Service defines five primary markets for which it reports fed cattle prices: Colorado, Iowa-Minnesota, Nebraska, Kansas, and Texas-Oklahoma. These regional prices have been found to have distinct differences but cointegrating interrelationships as the markets have spatial linkages through transportation and competing demands for resources (Bailey and Brorsen 1985). In the context of forecasting, we investigate two questions: the performance of intraregional price forecasts and the performance of national price forecasts using regionally disaggregated data.[3] Can disaggregating prices into the separate regions improve forecasting by capitalizing on differences in regional price patterns and market conditions? To forecast regional prices, the best forecast models from the rolling analysis are re-estimated on each set of regional data. In this way, we apply the structural form of the most accurate models to the regional series to allow coefficients to change to fit the new data. Regional cattle on feed and placement data are substituted for national data where applicable to best model the workings of the individual markets.

Forecasting performance of the regional models are compared with the performance of the national price forecast models by examining differences in RMSE. First, the forecast accuracy of each regional model is evaluated to understand how well prices can be forecasted within each region. Forecast accuracy is compared between regions and compared to the national price series as a baseline. The average RMSE and the minimum RMSE of the forecast models for each region are compared to the average and minimum RMSE of national forecasts. Comparing the average RMSE of the forecast model set allows us to understand how forecasts of each regional price series compare in general. Examining the minimum RMSE reveals any differences in the maximum achievable accuracy of forecasts of each region's prices. This is the real measure of interest because it is the maximum accuracy that matters most to forecast users.

Second, we evaluate the effectiveness of using separate regional models to forecast national prices. For each structural model, a simple average of price forecasts from each region is taken to obtain a national price forecast. This aggregation is given by Equation (17):

[3]Here forward term "national prices" will be considered synonymous with the 5-market price series since these five markets account for the majority of negotiated fed cattle transactions in the U.S. and this is a widely followed price series.

$$(17) \quad \widehat{pfed}_t^{national} = \frac{\widehat{pfed}_t^{region\_1} + \widehat{pfed}_t^{region\_2} + \cdots + \widehat{pfed}_t^{region\_n}}{(\#\ of\ regions)}$$

where the left side of the equality is the national forecast and $\widehat{pfed}_t^{region\_1}$ through $\widehat{pfed}_t^{region\_n}$ are the regional forecasts. Following this procedure, national price forecasts are generated from each forecast model with the regionally disaggregated data. These aggregated national forecasts are then compared with the forecasts from the direct national models with the same structural forms. Our assumption of equal-weighting of regional forecasts is likely an over-simplification since the 5-market price data is not a simple average of prices but a weighted average based on number of head sold from each region. However, regional prices have been shown to have cointegrating relationships over time (Bailey and Brorsen 1985) and equal-weighting procedures have been found just as accurate as more complicated forecast compositing methods in other applications (Colino et al. 2012). Therefore, a simple average is considered adequate for our purposes; if regionally disaggregated forecasting shows little potential with equal-weighting, then more complex weighting procedures would probably be no better. Again, the minimum RMSE and average RMSE are compared between the direct national price forecasts and the regionally disaggregated forecasts to evaluate the implications of regional disaggregation to maximum accuracy potential and to accuracy of the model set in general.

3.7 Error Evaluation

Forecast errors are examined to determine if they are random or can be explained by fundamental changes in the market. Forecast errors would reflect the workings of an efficient market if they are related to surprises or unexpected changes in the market compared to assumptions the forecasts are based on. For instance, if supplies of beef were substantially larger than anticipated by the model, realized prices would be expected to be lower than forecasts that anticipated lower levels of beef supplies. The same concept applies to futures market predictions and market consensus fundamental assumptions. The relationship between forecast errors and shocks to market fundamentals will be investigated via correlations and regression analysis.

Errors are compared against shocks to variables that may or may not be directly encompassed in the forecast models. These shocks are defined by the applying a simple univariate econometric model with a trend and monthly dummy variables to the series of interest. An autoregressive component is added for models with adjusted R^2 value less than 0.80 to improve fit while avoiding over-fitting the data[4]. A twelve-month lag is used for the autoregressive variable to tie predictions to year-ago levels and to better reflect an expectation developed farther in advance as compared to a one-month lag. For example, one univariate model could appear as:

$$(18) \qquad y_t = \beta_0 + \beta_1 Jan_t + \cdots + \beta_{11} Nov_t + \beta_{12} t + \beta_{13} y_{t-12} + e_t$$

where y_t is the series being analyzed for shocks, the β's are the parameters to be estimated, Jan_t through Nov_t are monthly dummy variables, t is time, y_{t-12} is a twelve-month lag of y_t and e_t is the error term. Solving for e_t yields:

$$(19) \qquad e_t = y_t - \beta_0 + \beta_1 Jan_t + \cdots + \beta_{11} Nov_t + \beta_{12} Trend_t + \beta_{13} y_{t-12} = y_t - \widehat{y_t}$$

where $\widehat{y_t}$ are the fitted values from the equation. The errors, e_t, from these univariate models are considered "shocks" in this analysis as they are deviations from the normal pattern of the series and will be referred to as shock variables. Correlations between forecast errors and shock variables indicate a relationship between forecast errors and underlying fundamental shocks in the market. Forecast errors are later regressed on shock variables to determine which

---

[4] Several R^2 thresholds were tested. Below 0.80, the univariate models did not fit the data well enough without the AR component. Adding the AR component to univariate models above this threshold resulted in R^2 in excess of 0.95. These models posed a problem since they had virtually no errors to use to explain forecast error.

fundamental shocks significantly contribute to errors and how much of the forecast error can be explained.

In addition to fundamental factors, the relationship between price forecasts errors and the volatility in the market place is also investigated. Two measures are considered to represent this behavioral component of the market: the momentum of price movements and a rolling standard deviation of price. Momentum will be defined as the difference between a two-month and six-month moving average. Momentum in time period *t* is calculated as:

$$(20) \quad momentum_t = \frac{1}{m}\sum_{t-m+1}^{t} pfed_t - \frac{1}{n}\sum_{t-n+1}^{t} pfed_t$$

where m=2 and n=6 such that the first term gives a two-period moving average and the second term gives a six-period moving average of price. The standard deviation will be calculated for the most recent 12 observations and will serve as a general proxy for market volatility.

Correlations between momentum and standard deviation with forecast errors indicate a relationship between the shorter-term behavior of prices and forecast performance. It is important to note that some of this price behavior may be the direct result of shocks to other fundamental factors. Large shocks to fundamentals could logically cause increases in momentum and volatility of prices. However, the relationship between forecast error and price volatility is important to investigate due to the potential for volatility resulting from fundamental shocks to compound forecast error.

Following correlation analysis, forecast errors are regressed on a combination of the shock variables to obtain more robust results than the initial inferences drawn above. Most importantly, conclusions can be drawn regarding the statistical significance of shock variables on forecast error, holding the other shocks constant. Consideration is given to avoid using highly correlated independent variables to avoid multicollinearity issues (Gujarati and Porter 2009, p.

344). Since multiple shock variables investigated contain some of the same information, this allows flexibility to select shock variable combinations that minimize correlation between these variables while ensuring all relevant information is represented in the regression. In this way, multicollinearity issues can be minimized while also avoiding specification bias associated dropping relevant variables. The resulting error regression models take the following form:

$$(21) \quad u_{t-h}^{t} = \beta_0 + \beta_1 shock\_a_t + \beta_2 shock\_b_t + \cdots + \beta_{\mathrm{n}} shock\_n_t + v_t$$

where u_{t-h}^{t} are the errors from a price forecast for time period *t* generated in time period *t-h*, $shock\_a_t$ through $shock\_n_t$ are the shock variables (the errors from the univariate shock regressions), and v_t are the residuals. In other words, forecast errors are modeled as a function of shocks or surprises to each of the various fundamental factors. Coefficients are interpreted as the expected change in forecast error given a one-unit change in error from the respective univariate shock model, holding all else constant. They measure how deviations from the expected pattern of a given fundamental factor contributes to the difference between forecasted and realized fed cattle prices. While the numerical value of the coefficients cannot be interpreted in meaningful units, the sign on the coefficients denotes a positive or negative relationship between shocks and forecast errors.

CHAPTER 4: DATA

Data used in this analysis are monthly series from January 1990 to December 2016. Data are divided into two sets: data used in developing fed cattle price forecasts and data used to evaluate price forecast errors. Some data series are used in both applications, but other are used strictly in either forecasting or evaluation. The data series considered for use in forecasting are weighted average negotiated fed cattle prices, per capita consumption of beef, cattle on feed numbers, national farm corn prices, Oklahoma City feeder cattle prices and disposable income. These will be considered for use as endogenous variables in forecasting models. Additionally, feedlot placement data and futures market prices are considered for exogenous information. The description and variable symbols for each of these series are provided in Table 1. The variable symbols will be used to refer to data in the results section.

The other data collected and considered for explanation of forecast error are carcass weights, fed slaughter numbers, beef production, beef exports, net beef trade, cattle on feed over 150 days, and disposable income. Per capita disappearances of meat are also considered, including beef, the major substitutes for beef (combined pork and broilers), and all three major proteins (combined beef, pork and broilers). Table 2 provides a description of the data used in analyzing price forecast errors.

Table 1: Description of data used for fed cattle price forecasts.

Variable	Description	Symbol	Units
Beef consumption	Disappearance (production plus net trade)	beefcons	Lbs./capita
Cattle on feed	Cattle on feed inventory, 1000+ head capacity		
US	US total	cof	1,000 head
Colorado	Colorado	cof_CO	1,000 head
Iowa	Iowa	cof_IA	1,000 head
Kansas	Kansas	cof_KS	1,000 head
Nebraska	Nebraska	cof_NE	1,000 head
Texas	Texas	cof_TX	1,000 head
Corn price	US grain corn, price received	pcorn	$/bu
Fed price	Fed steer price, negotiated, live basis		
US	5-Market weighted average	pfed	$/cwt
Colorado	Colorado	pfedCO	$/cwt
Iowa-Minnesota	Iowa/Minnesota	pfedIA	$/cwt
Kansas	Kansas	pfedKS	$/cwt
Nebraska	Nebraska	pfedNE	$/cwt
Texas-Oklahoma	Texas/Oklahoma	pfedTX	$/cwt
Feeder price	Steer price, 700-800 lbs., Oklahoma City	pfeeder	$/cwt
Futures price	CME live cattle futures price, monthly close	LC	$/cwt
Futures variable	Fed price minus futures price for each contract	fut	$/cwt
Income	Personal disposable income	income	$/capita
Placements	Feedlot placements, 1000+ head capacity		
US	US total	plmt	1,000 head
Colorado	Colorado	plmt_CO	1,000 head
Iowa	Iowa	plmt_IA	1,000 head
Kansas	Kansas	plmt_KS	1,000 head
Nebraska	Nebraska	plmt_NE	1,000 head
Texas	Texas	plmt_TX	1,000 head

Table 2: Description of data used in forecast error evaluation.

Variable	Description	Symbol	Units
Beef consumption	Disappearance (production plus net trade)	beefcons	Lbs./capita
Beef exports	Total beef exports	beefexports	Million lbs.
Beef production	Commercial beef production	beefprod	Million lbs.
Carcass weights	Weighted average dressed steer and heifer weights	cxwgt	Lbs.
Cattle on feed greater than 150 days	Five-months lagged cattle on feed minus five-month cumulative marketings and disappearance	cof150	1,000 head
Income	Disposable income per capita	income	$/capita
Meat consumption	Beef, pork and broiler disappearance per capita (production plus net trade)	meatcon	Lbs./capita
Net beef trade	Total beef imports minus exports	netbeeftrade	Million lbs.
Slaughter	Federally inspected fed cattle slaughter numbers	fedsltr	1,000 head
Substitute consumption	Pork and broiler disappearance per capita (production plus net trade)	subcons	Lbs./capita

Fed cattle prices are United States Department of Agriculture (USDA) prices and obtained from the Livestock Market Information Center (LMIC). Feedlot placements, cattle on feed numbers, feeder cattle price, slaughter numbers, carcass weights, trade data and data used in calculating per capita meat consumption, are also USDA data obtained from LMIC. Futures prices monthly closing prices from the Chicago Mercantile Exchange via LMIC. Personal disposable income and population data (for per capita calculations) are from the Federal Reserve Bank. Throughout this work, “5-market price” and “national price” will be considered interchangeable terms with respect to fed cattle prices since the majority of negotiated fed cattle sales are included within these five markets (Colorado, Iowa/Minnesota, Nebraska, Kansas and Texas/Oklahoma).

4.1 USDA Cattle on Feed Report Change

The reporting of the monthly USDA Cattle on Feed report changed beginning in December 1996, switching from a seven-state report to a national report. A mean shifting dummy variable is used to account for this change when cattle on feed is included endogenously and visual analysis confirms that a mean shift accounts for the report change. Below, Figure 2 shows the fitted values from regressing the cattle on feed series on a trend, monthly dummy variables and a mean shift after the report format changed. Results from this regression can be found in Table A1 of the appendix. The mean shift is significant at the 1% level and the R^2 from the regression is 81%. This confirms the need for a mean shift to allow the series to be integrated. A trend shift is insignificant at the 5% level and is therefore not used in integrating the report change.

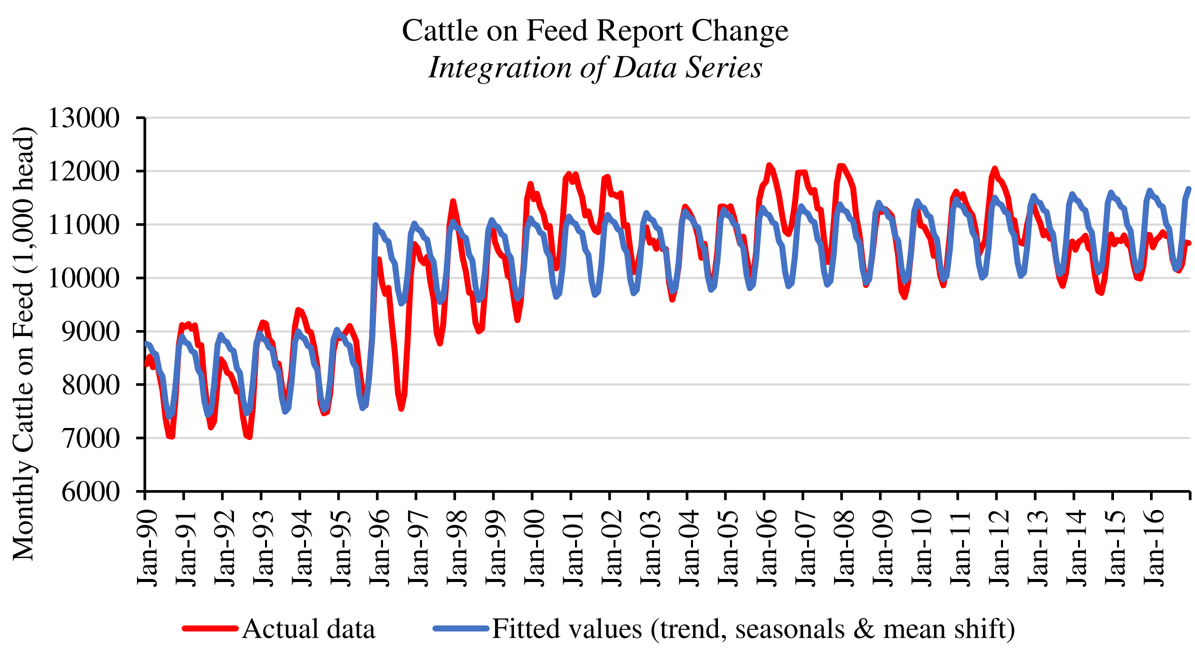

Figure 2: Integration of data across Cattle on Feed report change with a mean shift.

CHAPTER 5: RESULTS AND DISCUSSION

The previous sections have outlined the motivation for this research and the procedures used to obtain the results that will be discussed in this chapter. This chapter begins with the stationarity tests on variables and the necessary transformations. Next, Section 5.2 presents the model used to derive forecasts of feedyard placements that are needed as exogenous variables in some time series forecast models. Section 5.3 presents the forecasts developed and the results of the forecast performance evaluations. This includes a discussion of accuracy of models estimated in both the single-estimation and rolling-estimation frameworks, forecast encompassing tests, and an analysis of forecasting with regional fed cattle data. Next, the results from the forecast error evaluation are presented in section 5.4.

5.1 Stationarity Tests

All variables considered for use in forecasting models were tested for stationarity with the Augmented Dickey-Fuller (ADF) test. The trend coefficient is restricted to zero for all series except disposable income, which shows a clear upward trend over time. Results show that all price series are nonstationary since these series have ADF test statistics that are less in absolute value than the 5% critical values for the tests. These nonstationary variables are all fed cattle price series, feeder prices, corn prices, live cattle futures prices and income. All cattle on feed and feedlot placement series are stationary over time, along with the constructed futures variable and per capita consumption of beef. Results from the ADF tests on the data series are presented in Table 3.

Table 3: Augmented Dickey-Fuller test results on data series.

Variable	Symbol	ADF Test Statistic	5% Critical Value	Stationarity
Beef consumption	beefcons	-4.65	-2.87	Stationary
Cattle on feed - National	cof	-5.30	-2.87	Stationary
Colorado	cof_CO	-8.19	-2.87	Stationary
Iowa	cof_IA	-3.41	-2.87	Stationary
Kansas	cof_KS	-4.88	-2.87	Stationary
Nebraska	cof_NE	-10.01	-2.87	Stationary
Texas	cof_TX	-3.91	-2.87	Stationary
Corn price	pcorn	-1.90	-2.87	Nonstationary
Fed price - 5 Market	pfed	-1.43	-2.87	Nonstationary
Colorado	pfedCO	-1.40	-2.87	Nonstationary
Iowa-Minnesota	pfedIA	-1.50	-2.87	Nonstationary
Kansas	pfedKS	-1.43	-2.87	Nonstationary
Nebraska	pfedNE	-1.48	-2.87	Nonstationary
Texas-Oklahoma	pfedTX	-1.35	-2.87	Nonstationary
Feeder price	pfeeder	-1.60	-2.87	Nonstationary
Futures price	LC	-1.37	-2.87	Nonstationary
Futures variable	fut	-7.99	-2.87	Stationary
Income	income	-3.26	-3.42	Nonstationary
Placements - National	plmt	-9.53	-2.87	Stationary
Colorado	plmt_CO	-11.89	-2.87	Stationary
Iowa	plmt_IA	-6.91	-2.87	Stationary
Kansas	plmt_KS	-7.66	-2.87	Stationary
Nebraska	plmt_NE	-10.52	-2.87	Stationary
Texas	plmt_TX	-10.80	-2.87	Stationary

Data series that are nonstationary at the 5% level are first differenced. ADF tests are conducted on the differenced series to check that the transformation made series stationary. As shown in Table 4, all first differenced series are stationary at the 5% level. For these data series, the stationary, first differenced series is used in all subsequent analysis.

Table 4: Augmented Dickey-Fuller test results on first differenced data series.

Variable	Symbol	ADF Test Statistic	5% Critical Value	Stationarity
Corn price	Dpcorn	-9.69	-2.87	Stationary
Fed price - 5 Market	Dpfed	-11.66	-2.87	Stationary
Colorado	DpfedCO	-11.69	-2.87	Stationary
Iowa-Minnesota	DpfedIA	-11.39	-2.87	Stationary
Kansas	DpfedKS	-11.70	-2.87	Stationary
Nebraska	DpfedNE	-11.65	-2.87	Stationary
Texas-Oklahoma	DpfedTX	-11.72	-2.87	Stationary
Feeder price	Dpfeeder	-9.51	-2.87	Stationary
Futures price	DLC	-11.77	-2.87	Stationary
Income	Dincome	-15.26	-3.42	Stationary

5.2 Placement Forecasts

Number of placements of feeder cattle into feedlots from previous periods is an important variable to consider in fed cattle price forecasting models. The maximum forecasting horizon in this analysis is nine months, however feeder cattle placed into feedlots typically spend only four to six months on feed before being marketed as fed cattle. Forecasts of fed cattle prices beyond the standard feeding timeframe are price forecasts for cattle that have not yet been placed on feed. Forecasts of feedlot placement numbers are needed for forecasts of fed cattle prices at horizons longer than this timeframe. Examination of cross-correlation functions and typical production practices indicate a five-month lag of placements is appropriate to use in models of fed prices. In other words:

(22) $\Delta pfed_t = f(plmts_{t-5}, \dots)$

Following this relationship, forecasts of fed cattle prices at six- to nine-month forecast horizons are a function of future levels of feedlot placements:

(23) $\Delta \widehat{pfed}_{t+6} = f(\widehat{plmts}_{t+1}, \dots)$ to $\Delta \widehat{pfed}_{t+9} = f(\widehat{plmts}_{t+4}, \dots)$

Procedures for the development of the placement forecast model are described in Section 3.5. The final placement forecast model is specified as:

$$(24) \quad plmt_t = \beta_0 + \beta_1 plmt_{t-1} + \beta_{11} plmt_{t-11} + \beta_{12} plmt_{t-12} + \gamma cows_{t-6} + \alpha oldcof_t + u_t$$

where $oldcof_t$ is a binary variable to allow for a mean shift with the change in the reporting of data in the USDA Cattle on Feed report in which placement data is reported and all the other terms have previously been defined. The autocorrelation structure of the data series indicates placements are driven by one-, eleven-, and twelve-month lags of itself. Six-month lagged beef cow inventory numbers are also included, as previously discussed.

The placement model is estimated on data monthly data from January 1990 through December 2006. The R^2 value of this equation is 0.8513, indicating that the model fits the data reasonably well. A summary of the results from the model estimation are provided in Table A2 of the appendix. Forecasts are generated over the post-sample period of interest for fed cattle prices, January 2007 through December 2016. The forecast accuracy of this model according to RMSE and MAPE is provided in Table 5 below.

Table 5: Out of sample accuracy of monthly cattle on feed placement forecasts, January 2007- December 2016.

Horizon:	1-Month	2-Months	3-Months	4-Months	5-Months
RMSE	120.29	129.26	129.51	129.80	129.75
MAPE	5.1%	5.5%	5.6%	5.6%	5.6%

5.3 Forecast Performance

Candidate forecast models were developed by the methods outlined in Section 3.2. All models were first developed in a single-estimation framework. This section begins by presenting results from the single-estimation framework in Section 5.3.1. Some candidate models were selected for re-estimation and forecasting in a rolling-estimation framework as presented in

Section 5.3.2. The most accurate rolling-estimation models were selected for accuracy evaluations and comparisons in Section 5.3.3. Forecast encompassing tests and regional forecasting analysis are presented in Section 5.3.4 and 5.3.5, respectively.

Specifications and names for all candidate forecast models can be found in Appendix B. All models (48) were first estimated and evaluated in the single-estimation framework. The forecast models that were selected for re-estimation in a rolling framework (20 models) are denoted by model names in bold font. Of these rolling models, ten were selected for the regional forecasting analysis, and these forecasts are denoted by bold and italicized model names.

Section 5.3.1: Single Estimation

A total of 48 candidate time series models were developed in the single-estimation framework. Forecasts were generated from each model and accuracy was compared between models. The present section focuses on general observations of forecast performance and differences in performance of differing structural models and model specifications.

Examining price forecasts against realized prices over time shows that forecasts generally follow the trend of actual prices but with variable and sometimes substantial errors. As an illustration, Figure 3 compares six-month horizon forecasts from the lowest RMSE model of each specification type (AR, VAR and VEC) over time against actual prices. Price forecasts from different models tend to have the largest errors during the same time periods, some with larger errors than others. Notably, most forecast models missed or were slow to anticipate swift price movements during the price run up in late 2013 and 2014 and during the crash in 2015.

Fed Cattle Price vs. Six-Month Ahead Forecasts

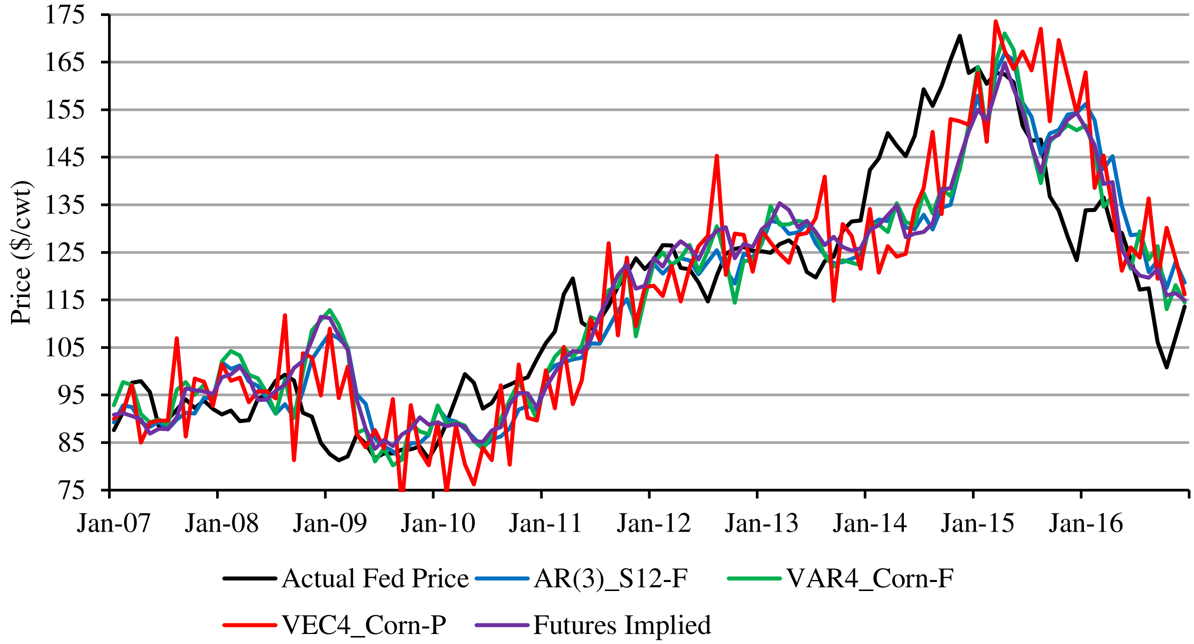

Figure 3: Actual fed cattle prices and six-month horizon forecasts from most accurate single-estimation models and over time, January 2007-December 2016

Figure 4 below shows the forecast RMSE at all nine horizons for selected models. These models were the most accurate of each specification for at least one forecast horizon. Several observations are apparent. Not surprisingly, forecast errors increase with forecast horizon, but the rate of the increase in RMSE appears to slow and plateau at more distant horizons for most models. Similar patterns are found for MAPE and are not presented for brevity. RMSE and MAPE of all forecasts are presented in Table A3 in the appendix.

Different model specifications are more accurate at different horizons. AR models are out-performed at closer horizons but perform well at extended horizons. AR models have similar RMSE values to VAR models beginning at about the four-month horizon and have consistently lower RMSE than VEC models at horizons of five months or more. Comparing VAR to VEC models, VEC models are generally more accurate at closer horizons and especially at a two-month horizon. VAR's are generally better at intermediate and longer horizons of about four months or greater, becoming much more accurate than VEC models at the most distant horizons. An exception is that the VAR model with the lowest RMSE at close horizons is even less accurate than the VEC models at extended horizons.

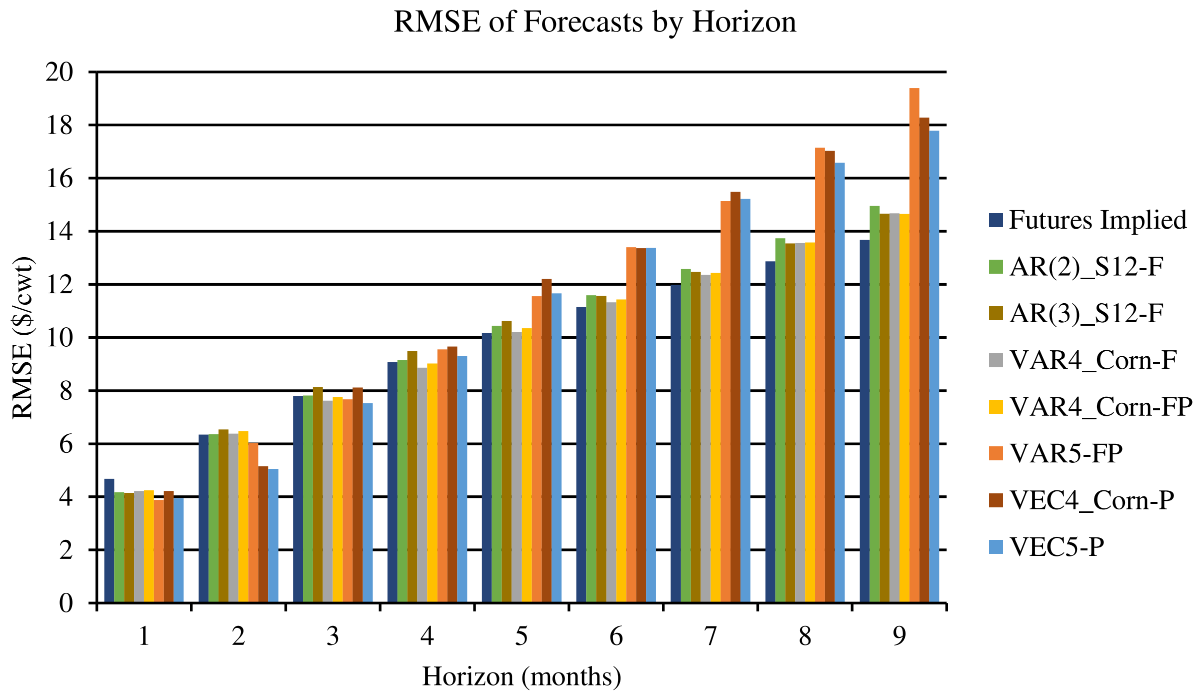

Figure 4: Out of sample RMSE of selected single estimation models by forecast horizon.

Consistent with prior research, the futures market performs generally well from a comparative standpoint, especially at longer horizons. Time series methods can out-perform futures market forecasts at three months ahead and nearer. Most notably, VEC models substantially beat the futures at one and two-months but futures become far inferior at the most extended horizons. All selected time series models, except the VEC models, include futures as an exogenous variable, indicating the value of including futures information in forecasting models.

Previous work has generally found error-correcting terms to be important, especially at longer forecast horizons, which appears contrary to the present findings. However, these previous studies have not included futures prices as an independent variable in forecast models of cash prices. This appears to contribute significantly to forecasting performance. Adding either futures or error-correcting terms improves forecasting compared to the base VAR model, with models that include futures usually being more accurate than the error-correcting specification. However, incorporating both an error-correcting specification and futures as an exogenous variable performs worse than either of these methods alone but somewhat better than the base VAR. This pattern is shown for two base VAR forecast models in Figure 5 below. Within the two groups of four similar models, the left-most is the base VAR, next is the VEC with futures, the base VEC, and the VAR with futures is on the right.

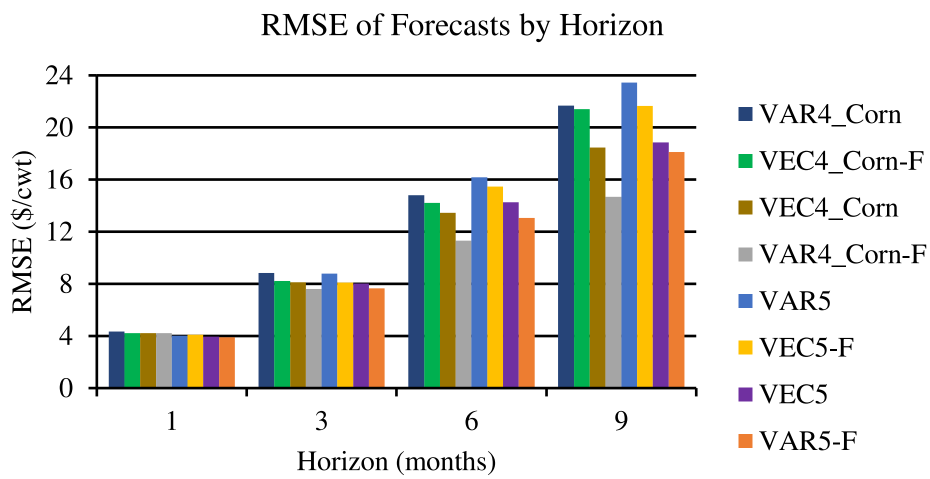

Figure 5: Out of sample forecast accuracy of VAR and VEC models with and without incorporating futures prices.

Section 5.3.2: Rolling Estimation

Of the 48 candidate models developed, 20 were selected for re-estimation and forecasting in a rolling estimation framework. This includes all forecast models that were the most accurate by RMSE for one or more horizons, variations of those models with different combinations of exogenous variables, and other models subjectively selected. No forecast models with income as a variable were selected; income neither contributed nor detracted from forecasting performance compared to otherwise identically specified models and was insignificant in all equations. Therefore, models with income were excluded on the principle of parsimony.

Contrary to expectation, forecasting performance generally declined by RMSE and MAPE criteria compared to a single estimation. The differences in RMSE between rolling- and single-estimation forecast are shown in Table 6 where positive values indicate an increase in error with the rolling estimation framework as compared to a fixed estimation. Of the 180 pairwise comparisons, RMSE increased 62% of the time. However, there were some trends in accuracy improvements. The AR models improved at shorter horizons and the AR model with only seasonal dummy variables improved out to the five-month horizon. VAR models became slightly more accurate at the shortest horizons and the two VAR models with specifications that did not include futures or placement information exogenously became more accurate at all horizons. Lastly, the VEC models improved at the one-month horizon and generally became more accurate at the five-month horizon and longer. All improvements in accuracy are denoted by bold font in Table 6.

Table 6: Difference in RMSE between rolling- and single-estimation forecasts by forecast horizon.

Model	H1	H2	H3	H4	H5	H6	H7	H8	H9
AR(3)_S12	**-0.35**	**-0.31**	**-0.23**	**-0.18**	**-0.02**	0.16	0.22	0.21	0.18
AR(2)-F	**-0.41**	**-0.23**	0.18	0.66	1.18	1.71	2.31	2.85	3.37
AR(3)-F	**-0.32**	**-0.25**	0.07	0.47	1.00	1.59	2.18	2.69	3.15
AR(2)_S6-F	**-0.43**	**-0.25**	0.16	0.60	1.09	1.63	2.22	2.80	3.39
AR(2)_S12-F	**-0.41**	**-0.18**	0.18	0.64	1.21	1.85	2.54	3.31	4.10
AR(3)_S12-F	**-0.34**	**-0.27**	**-0.04**	0.30	0.89	1.56	2.19	2.84	3.52
VAR4_Corn	**-0.46**	**-0.41**	**-0.62**	**-1.36**	**-1.99**	**-2.10**	**-3.04**	**-3.89**	**-5.33**
VAR4_Corn-F	**-0.43**	**-0.08**	0.54	1.07	1.59	2.32	3.00	3.70	4.38
VAR4_Corn-FP	**-0.41**	**-0.14**	0.43	0.89	1.38	2.06	2.77	3.52	4.29
VAR4_Corn-P	**-0.47**	**-0.21**	0.17	0.37	0.53	0.82	0.97	1.11	1.30
VAR5	**-0.18**	**-0.10**	**-0.64**	**-1.75**	**-2.79**	**-3.31**	**-4.35**	**-5.16**	**-6.54**
VAR5-F	**-0.15**	0.22	0.50	0.64	1.00	1.57	2.16	2.83	3.50
VAR5-FP	**-0.11**	0.25	0.41	0.41	0.67	1.23	1.86	2.54	3.12
VAR5-P	**-0.14**	0.28	0.43	0.29	0.24	0.34	0.36	0.36	0.28
VEC4_Corn	**-0.33**	1.43	0.42	0.59	**-0.29**	0.10	**-0.58**	**-0.63**	**-0.62**
VEC4_Corn-FP	**-0.29**	1.01	0.28	0.07	0.33	0.80	0.67	0.16	0.16
VEC4_Corn-P	**-0.28**	1.51	0.52	0.73	**-0.09**	0.36	**-0.26**	**-0.26**	**-0.23**
VEC5	**-0.12**	1.28	0.30	**-0.03**	**-0.96**	**-0.90**	**-1.29**	**-0.81**	**-0.59**
VEC5-FP	**-0.17**	1.40	0.83	0.57	0.03	**-0.05**	**-0.32**	0.21	1.03
VEC5-P	**-0.08**	1.45	0.81	0.61	**-0.13**	**-0.28**	**-0.51**	**-0.18**	0.07

Note: Differences are calculated as RMSE of rolling model minus RMSE of single-estimation model. Negative values are bolded and indicate where the rolling-estimation model had a lower RMSE.

A clear intuitive reason for the accuracy decline is not apparent. Assuming structural change has occurred in the post-sample period, rolling updates would be expected to generate parameters that are more unbiased and generate more accurate forecasts. One possible explanation appears to be a change in the drift term. Since fed prices are estimated in first differences, the constant in the regression equation can be interpreted as the drift term. Figure 6 shows the coefficients over time from one VAR model that declines the most dramatically in forecast accuracy at intermediate and longer horizons. Most coefficients do not vary much over time except for the constant, which visually resembles the movements of fed cattle prices over time (plotted in Figure 6 for reference). It appears that the coefficients in the model do not fully explain the changes in price over time, leaving the constant to capture the drifting mean of the price series. Since the constant has the greatest variation over time, it is suspected that declines in accuracy may be related to the constant. Having the correct constant in forecast models appears vital to forecast accuracy and warrants further research beyond the scope of this analysis.

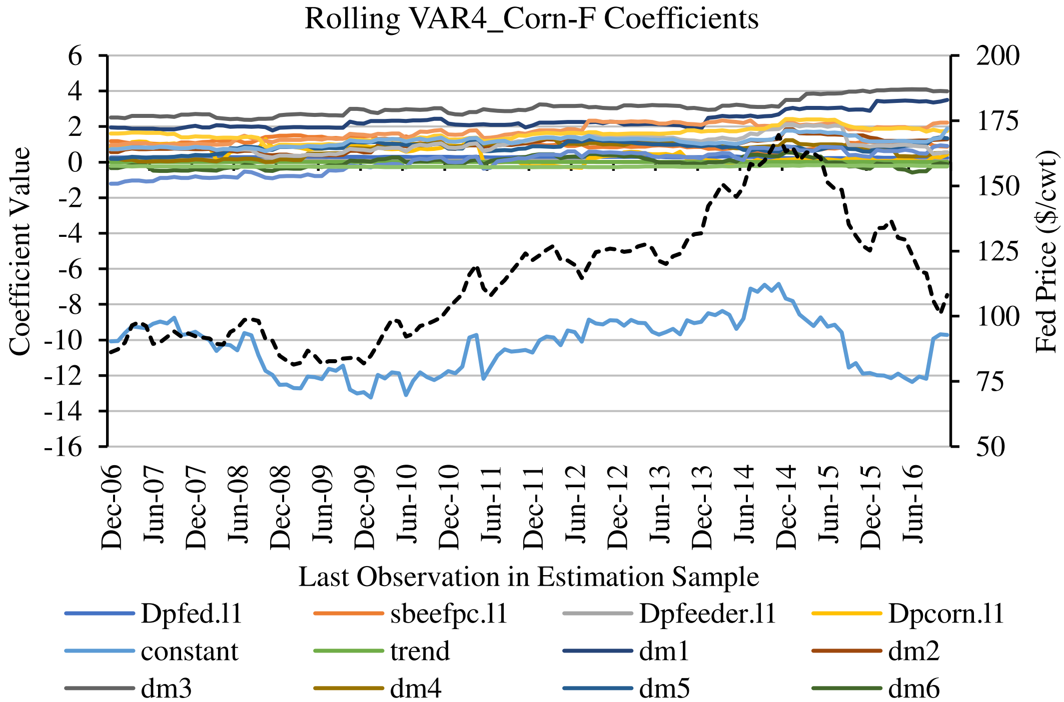

Figure 6: Coefficients from the fed cattle price equation of a VAR model over the post-sample period, Jan. 2007-Dec. 2016.

Although forecast accuracy declined for many forecast models in the rolling framework, even when accuracy declined, the declines were generally not large. Graphical comparisons over time show that forecasts from the same model in the two estimation frameworks do not vary substantially, as seen with six-month horizon forecasts in Figure 7. As with the single estimation forecasts, forecasts tend to move with prices over time with what appears to be a lagging effect.

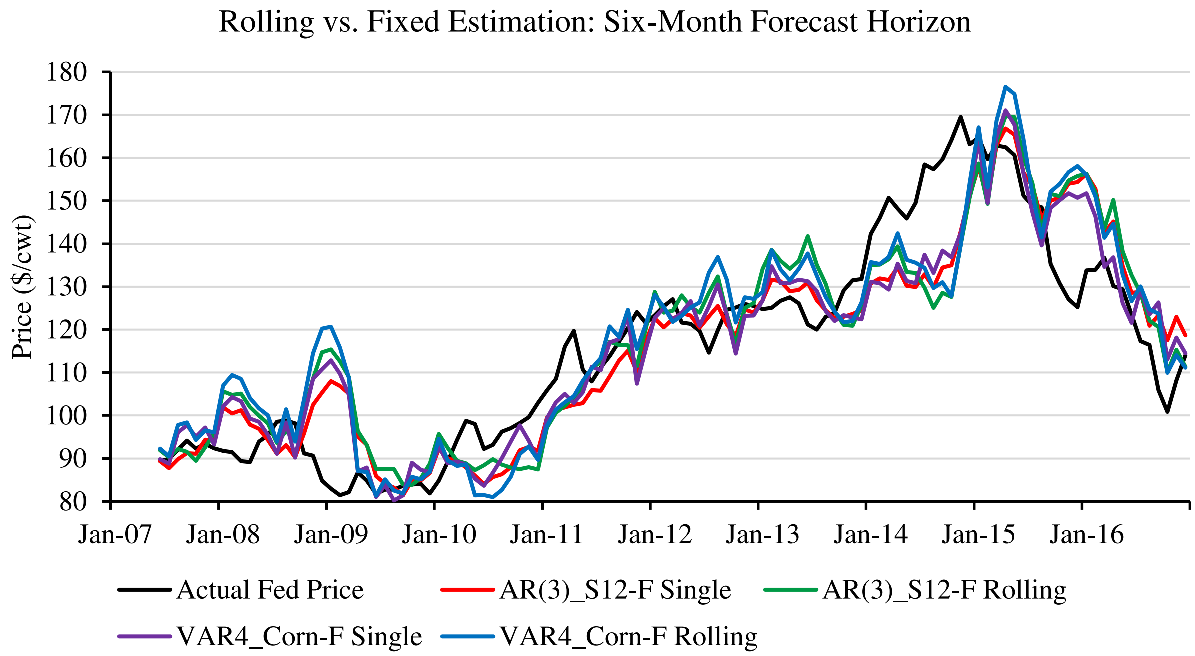

Figure 7: Six-month ahead forecasts from an AR and VAR model in rolling- and fixed-estimation methods.

Comparing forecast accuracy across horizons and model types, we find similar results with the rolling framework as with the fixed estimation previously discussed. Forecast accuracy declines with forecast horizon and different models have relative advantages at different forecast horizons. Figure 8 shows the out of RMSE of selected AR, VAR and VEC models that are the most accurate at the one, three, six or nine-month forecast horizon. Accuracy based on MAPE shows similar results and is used to select between models of the same specification type with similar RMSE.

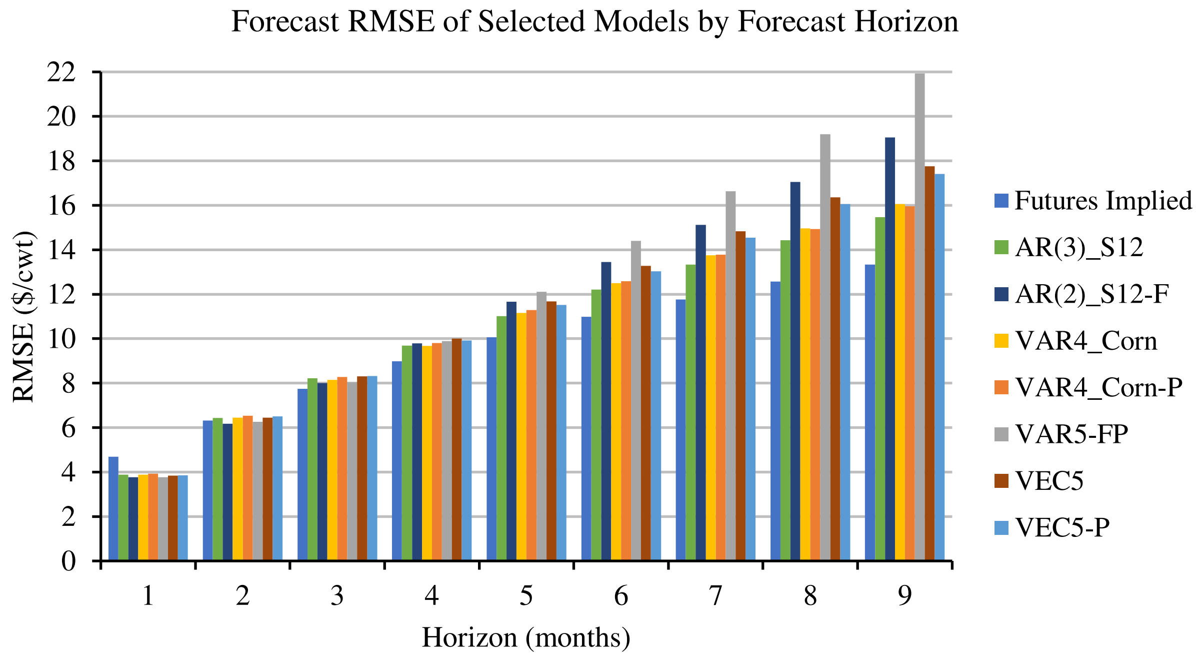

Figure 8: Out of sample RMSE of rolling estimation forecast models by forecast horizon, Jan. 2007-Dec. 2016.

Section 5.3.3: Accuracy Evaluation

The most accurate candidate models at forecast horizons of one, two, three, six and nine months were selected for accuracy evaluations and comparisons. The best AR, VAR, and VEC model based were selected at each of these horizons. Futures-implied predictions are also compared at these horizons.

We focused on these horizons for several reasons. The one-month horizon is important for evaluating short-term forecasting performance. However, values of monthly data are not known until that month ends and the next begins so a one-month forecast cannot be generated until the month being forecasted has begun. A two-month horizon is the more relevant next-month forecast that would be useful to the producer. Three- and six-month horizons were evaluated as intermediate horizons where the producer has some degree of decision-making flexibility. The nine-month horizon is the long-term forecast evaluated in this analysis.

Candidate models that were selected at these horizons are shown in Table 7. The model names, types and variables that compose each model are also presented. The lag structure multiple-equation VAR and VEC models is determined by the Schwarz information criterion (SC) as denoted. Forecast models were often selected as the most accurate models for several horizons.

Table 7: Description and variable composition of most accurate selected time series forecast models.

Model name	Type	Endogenous variables	Lag order	Monthly dummy variables	Exogenous variables	Error correction	Horizon(s) evaluated
AR(2)_S12-F	AR	Fed price	2	x	Futures variable		1, 2, 3
AR(3)_S12	AR	Fed price	3	x			6, 9
VAR5-FP	VAR	Fed price, feeder price, corn price, beef consumption, cattle on feed	Selected by SC	x	Futures variable, placements		1, 2, 3
VAR4_Corn	VAR	Fed price, feeder price, corn price, beef consumption	Selected by SC	x			6
VAR4_Corn-P	VAR	Fed price, feeder price, corn price, beef consumption	Selected by SC	x	Placements		9
VEC5	VEC	Fed price, feeder price, corn price, beef consumption, cattle on feed	Selected by SC	x		x	1, 2, 3
VEC5-P	VEC	Fed price, feeder price, corn price, beef consumption, cattle on feed	Selected by SC	x	Placements	x	6, 9

To compare forecast accuracy, Table 8 gives the differences in RMSE between the best forecast model of each of the four specification types by forecast horizon with significance based on the modified Diebold-Mariano (MDM) test. Differences are calculated as the RMSE of forecasts across the columns minus the forecast RMSE of the models labeled by row. For each horizon, the forecast model names are given in the row names. Results from forecast evaluations are presented at forecast horizons of one, two, three, six and nine months.

Table 8: Differences in RMSE between fed cattle price forecasting models.

Horizon	Forecast model	Futures RMSE diff.	AR RMSE diff.	VAR RMSE diff.
1-month	AR(2)_S12-F	0.92 ***		
	VAR5-FP	0.92 ***	0.00	
	VEC5	0.85 **	-0.07	-0.07
2-months	AR(2)_S12-F	0.15		
	VAR5-FP	0.07	-0.08	
	VEC5	-0.12	-0.27	-0.28
3-months	AR(2)_S12-F	-0.26		
	VAR5-FP	-0.28	-0.03	
	VEC5	-0.56	-0.30	-0.28
6-months	AR(3)_S12	-1.21		
	VAR4_Corn	-1.51	-0.30	
	VEC5-P	-2.05	-0.83	-0.54
9-months	AR(3)_S12	-2.13		
	VAR4_Corn-P	-2.61	-0.48	
	VEC5-P	-4.07	-1.93	-1.46

Note: Differences are RMSE of column model minus RMSE of models in the corresponding rows. Single, double and triple asterisk (*) denote significance differences at the 10%, 5% and 1% levels by the MDM test.

At the one-month horizon, the AR and VAR models had a very similar RMSE and the p-value of the MDM test approaches one, showing that despite the added complexity of the VAR, forecast accuracies are virtually identical. The more robust error correcting specification of the VEC model shows small and insignificant decrease in accuracy compared to the AR and VAR models. All time series models are more accurate than futures market predictions at the one-month horizon. The differences between RMSE of futures market predictions and the best AR, VAR and VEC models at the one-month horizon are the only significant differences in RMSE between the selected models shown in Table 8. At the two-month horizon, the AR and VAR models are more accurate than futures but the differences are not significant.

Inclusion of futures information appears to important to accuracy of time series models in the shorter term, since the best AR and VAR models at the one- through three-month horizons are models that include the futures variables (denoted by an “F” in the model name). The same three forecast models were selected as the most accurate at the one-month through three-month horizons, indicating that the same model specifications are preferred throughout this window of forecast horizons. The most accurate forecasts begin to change at the four-month forecast horizon. This can be seen by the full list forecast of RMSE and MAPE values of models estimated in the rolling estimation framework provided in the Appendix in Table A4.

Futures market predictions have the lowest RMSE at a three-month forecast horizon and beyond although the differences in accuracy are not significant between futures and any of the most accurate time series methods. Interestingly, at intermediate and more extended forecast horizons, including futures information no longer improves time series models, although the best models in the single-estimation framework did include futures information at these horizons. An explanation for the change in the benefits of including futures information between the single and rolling estimation frameworks is not immediately apparent, even when examining the coefficient estimates over time from the VAR model shown in Figure 6 above. Since futures by itself is the most accurate forecast, we would expect incorporating futures to improve time series models as well.

At the intermediate and longer horizons, simpler models appear to forecast more accurately, although differences in RMSE are not significantly different according to MDM tests. The AR model with only monthly dummy variables (AR(3)-S) has a lower RMSE value than all other time series models from the five-month forecast horizon and beyond. VAR models also have lower RMSE than the more complex VEC models at intermediate and longer horizons.

Placement information appears to be important to accuracy of multiple-equation forecast models at longer horizons by providing an estimate of future levels of fed cattle supplies. The most accurate VEC model at the six- and nine-month horizon and the VAR at the nine-month horizon include five month lagged cattle on feed placements exogenously. The importance of placements is especially surprising at the nine-month horizon since these models use forecasted values of placements. Improved placement forecasts may further improve accuracy of these fed cattle price forecast models.

Section 5.3.4: Forecast Encompassing

Forecast encompassing tests were conducted on the most accurate forecast model of each specification. Similar to the MDM test above, encompassing tests were conducted on forecast errors at the one-, two-, three-, six- and nine-month horizons. Two types of comparisons were made: futures predictions compared to time series forecasts and the most accuracy time series forecast compared to other forecasts. Forecast encompassing tests are used to determine if less accurate forecasts contain additional information not contained in a superior forecast. For this evaluation, the superior forecast in each comparison is selected based on the RMSE evaluation described earlier. Since futures is the superior forecast at the three, six and nine-month forecast horizons, all time series forecasts are compared to futures at these horizons. Encompassing tests are conducted between AR forecasts and the other time series forecasts at all horizons as well as with future predictions at the one- and two-month horizon, because the AR model is the superior forecast in these comparisons.

Results from the encompassing tests are presented in Table 9. The values from the pairwise regressions can be interpreted as the optimal weight of the alternative forecast in a composite with the superior forecast. Where futures predictions are more accurate, time series forecasts generally have incremental information within the two- to six-month window. The optimal weight of the VEC in a composite with futures is nearly 50% and significant at the two-month horizon. The optimal weights for the AR, VAR and VEC are approximately 40% at the three-month horizon and the weights on the VAR and VEC are significant. At the six-month horizon, the AR and VAR are significant at the 10% level and their optimal weights decline to about 25%. None of the time series forecasts contain significant incremental information at the nine-month horizon and notably, the coefficient on the VEC is negative, indicating that this forecast offers no additional value. No encompassing tests are conducted with futures at the one-month horizon since futures is not the superior forecast to any time series forecasts at this horizon.

Table 9: Forecast Encompassing Tests between Futures and Time Series Models and between AR Forecasts and Other Forecasts.

Alternate model	Forecast horizon 1-month	2-months	3-months	6-months	9-months
	- Optimal weight in composite with futures -				
AR	-	-	0.38	0.26 *	0.10
VAR	-	-	0.42 **	0.24 *	0.13
VEC	-	0.49 ***	0.38 ***	0.09	-0.18
	- Optimal weight in composite with AR -				
FUT	0.14	0.37 **	-	-	-
VAR	0.50	0.36	0.47 *	0.24	0.12
VEC	0.35	0.28	0.36 **	0.09	-0.42 *

Note: Single, double and triple asterisk (*) denote significance of alternate model in a composite forecast at the 10%, 5% and 1% levels by t-tests. AR, VAR and VEC models at each horizon are the same as in Table 8.

Encompassing tests show little incremental information in VAR and VEC forecast models compared to AR forecasts. At the one-month horizon, weights indicate that an equally weighted composite between AR and VAR forecasts would be optimal, but the weight is insignificant. This is consistent with the suggestion in the MDM test that these forecasts are virtually identical (see Table 8). Results also indicate that the one-month AR forecasts fully encompass VEC forecasts. Futures market predictions contain no incremental information to the AR forecasts at the one-month horizon, but contain valuable information at the two-month horizon with a significant weight of 37%. At the three-month horizon, the more complex VAR and VEC specifications appear to have incremental information to the AR forecasts. However, the optimal weights of the VAR and VEC are not significantly different than zero at the six and nine-month forecast horizon. As with the comparison with futures, the optimal weight of the VEC is negative at the nine-month horizon and is significant at the 10% level. This result is surprising. A possible explanation is compositing the AR forecasts with any weight on the VEC would be highly suboptimal.

Section 5.3.5: Regional Fed Cattle Prices and Forecasting

Regional fed cattle price data was used to investigate intra-regional forecasting accuracy and the forecasting efficiency of using regionally disaggregated data to forecast national prices. Ten forecast models were estimated on the regional data: the seven models that were selected as the most accurate models from the rolling evaluation above and one additional AR, VAR and VEC model each.

Regional forecasts were evaluated by comparing the average and minimum RMSE from the set of ten models between regions. The average and minimum RMSE from the same set of models on the national prices was used as the baseline for comparisons. The average RMSE is useful for understanding how the forecast models perform regionally in general, while the minimum RMSE shows the maximum accuracy this set of models can achieve by the best model for each region.

Forecasts for regional prices generally show only minor differences in accuracy as compared to the national price forecasts. For the model set, the average RMSE for any region at any horizon does not increase more than 10 percent or decrease more than 4 percent compared to national forecasts. The changes in average RMSE for forecasts at the one, three, six and nine-month horizons are shown in Figure 9. Kansas and Texas prices are generally forecasted more accurately at longer horizons. Forecasts of prices at nearby horizons had accuracy declines of approximately 6 to 9 percent for Colorado, Nebraska and Texas prices.

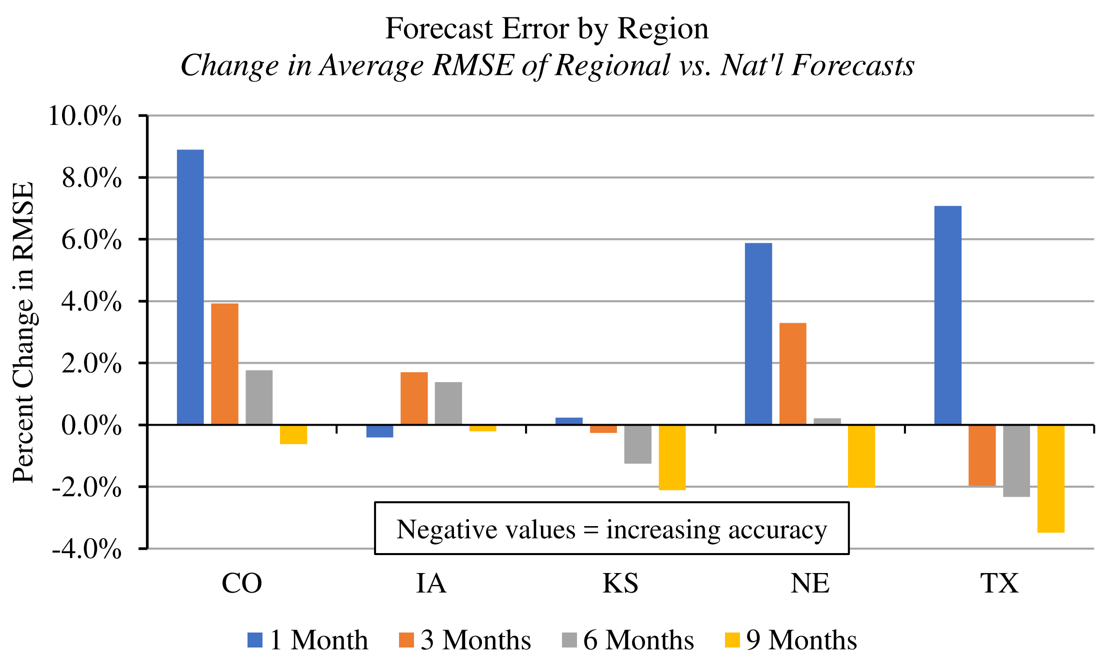

Figure 9: Percent change in average RMSE of regional price forecasts compared to average RMSE of national price forecasts from the model set.

The accuracy changes of the most accurate forecasts are shown in Figure 10 by comparing minimum RMSE of each region. Again, accuracy differences are small. All increases in minimum RMSE were less than 7 percent and decreases were less than 3 percent. Accuracy improved slightly at intermediate horizons for Kansas and at forecast horizons of three-months and longer for Texas prices. The accuracy of the best forecasts of Colorado and Nebraska prices declined slightly at all horizons. Iowa-Minnesota prices show virtually no differences in forecasting accuracy compared to forecasts of national prices by average RMSE or minimum RMSE.

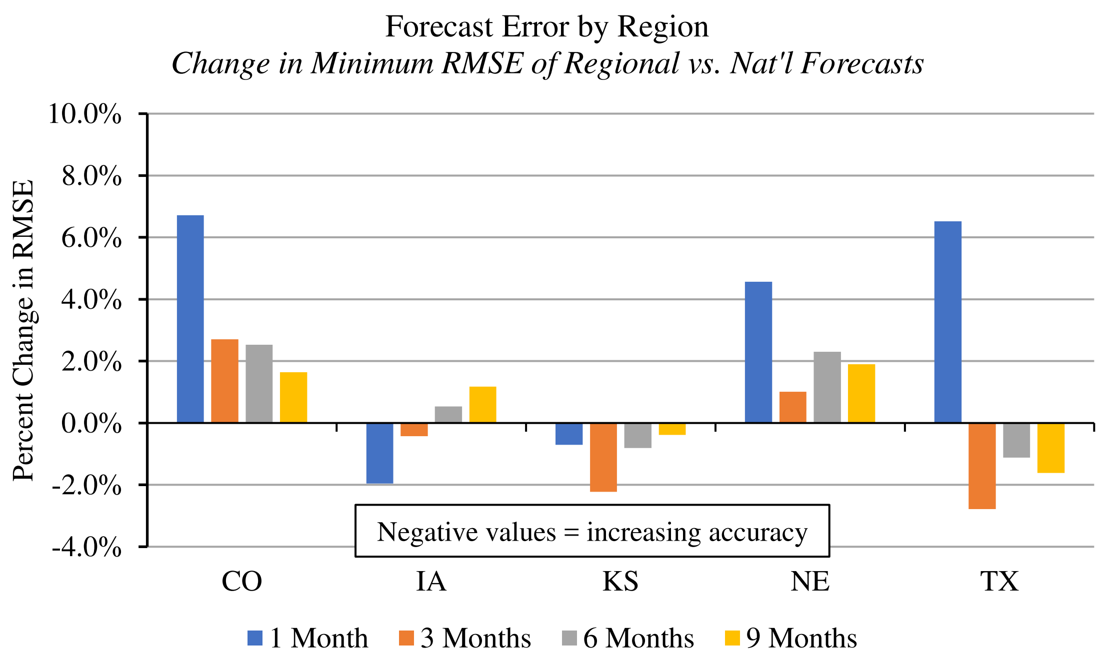

Figure 10: Percent change in minimum RMSE of regional price forecasts compared to minimum RMSE of national price forecasts from the model set.

Forecasts of national prices are generated by taking a simple average of the five regional forecasts to aggregate them into a national forecast. This aggregation procedure is done for each of the ten regional forecast models evaluated in this section. The accuracy of regionally disaggregated forecasts is then compared with direct forecasts of national prices by the same set of forecast models.

Accuracy of national price forecasts does not change significantly with regional disaggregation as compared to national models, although there are some differences. Figure 11 shows a box-and-whisker plot comparing the RMSE of the national forecast models and the disaggregated national forecasts. On average across the model set, accuracy is virtually identical at closer horizons and generally shows small improvements with regional disaggregation at intermediate and extended forecast horizons, as indicated by a comparison of the mean RMSE of the national and disaggregated forecasts. The most noticeable difference is the tightening of the spread in forecast errors with regional data, especially at more distant horizons. One possible explanation is that large inaccuracies of one regional model can be diluted by averaging the individual regional price forecasts, reducing large national price forecast errors. The result is that the worst models improve dramatically as the maximum errors being reduced substantially and the RMSE spread tightens.

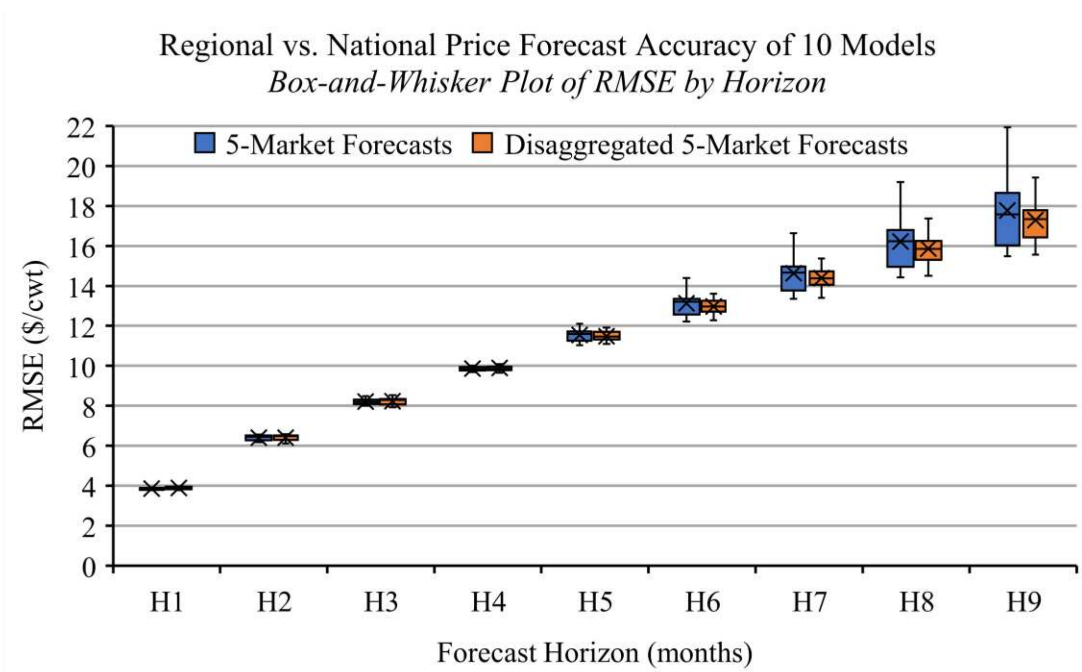

Figure 11: Box-and-whisker plot comparing forecast accuracy of regionally disaggregated forecasts with to national forecast models.

While regional disaggregation slightly improves forecast accuracy of the time series model set in general, it appears to not have any advantages in improving the maximum accuracy capabilities of these time series models. Figure 12 shows a closer examination of differences in RMSE between regionally disaggregated national forecasts and direct national forecasts. As shown in the box-and-whisker diagram above, Figure 12 also shows that the average forecast RMSE improves slightly at the five-month horizon and longer. Comparing the minimum RMSE of the forecasts indicates that the regionally disaggregated forecasts show minor accuracy improvements at closer horizons and minor accuracy declines at five months and beyond. Changes in maximum accuracy are negligible as all increases or decreases in minimum RMSE were approximately 1 percent or less.

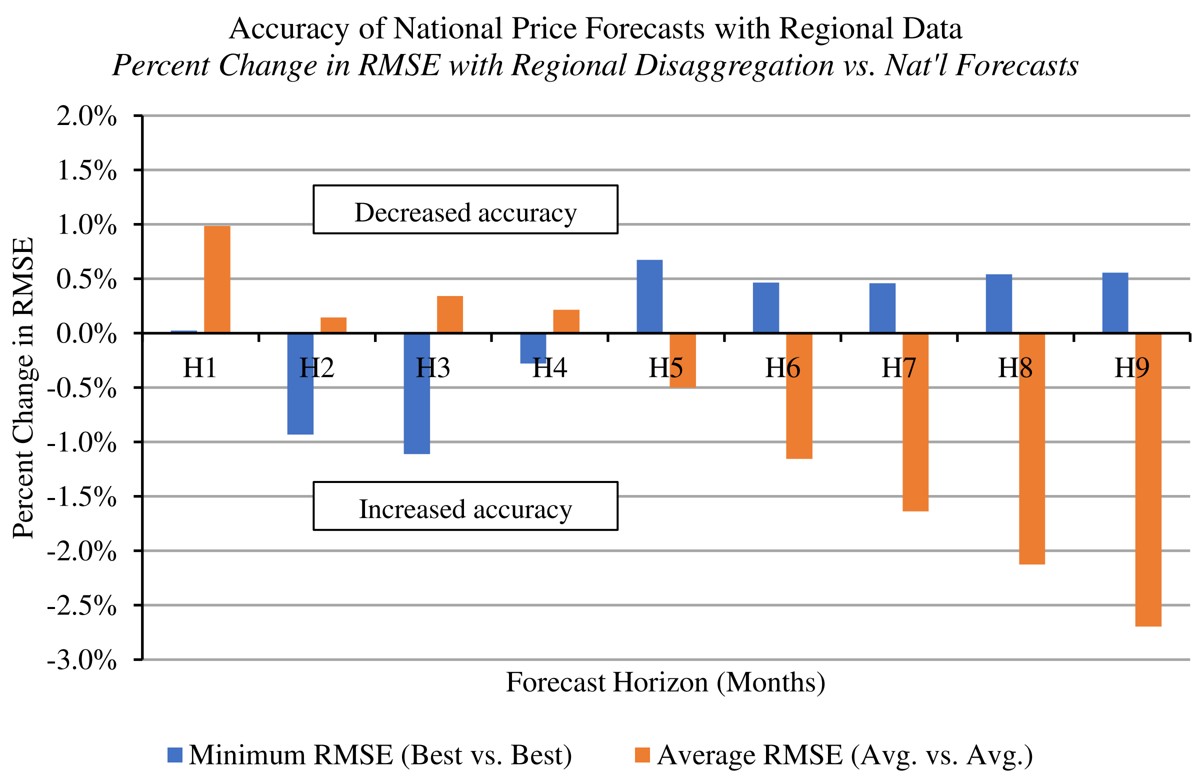

Figure 12: Forecast accuracy of regionally disaggregated forecasts compared to national forecast models.

5.4 Error Evaluation

The relationships between forecast errors and shocks to fundamental factors are examined through simple correlations and by regressing forecast errors on the shock factor variables. The shock variables are defined in Equation (19) and are discussed in Section 3.7. Correlations are also calculated with forecast errors and price momentum and volatility measures as purely technical measures of price behavior. Only the errors from the selected models of each specification for the one, three, six and nine-month horizons are examined. Analysis only performed on rolling-estimation forecasts since these models allow flexibility to structural changes over time and are expected to better reflect forecast models used in practice compared to the single-estimation framework.

Section 5.4.1: Correlation Analysis

Correlations are used as the initial step to identify potential relationships between forecast errors and market shocks. It is acknowledged that simple correlations do not inform us on statistical significance, however, they are useful in indicating potential relationships and in developing the error regressions that will follow.

Most correlations between errors and shocks are generally low, as can be seen in Table 10. All correlations greater than 0.20 in absolute value in bold for illustrative purposes. Errors from one-month ahead forecasts show the least relation to shocks, with most correlations less than 0.20 in absolute value and many less than 0.10. Correlations with momentum and volatility are also weak. Lack of correlation at the one-month horizon may reflect limited time for shocks to occur between forecasting and price realization.

The technical measures clearly indicate relationships with forecast error and these correlations also increase with forecast horizon. Price momentum and volatility have moderately strong and moderate correlations with errors, respectively. Correlations with price momentum reach 0.60 at the nine-month horizon and correlations with price volatility exceed 0.80. Interestingly, correlations with the 12-month standard deviation of price become markedly stronger at the nine-month horizon compared to the all shorter horizons. Futures markets errors show distinctly lower correlations with price momentum and somewhat lower correlation with price volatility than errors from time series models, possibly indicating that futures predictions are less effected during volatile price environment than time series forecasts. Overall, while correlations must be taken with caution, clearly the sharper the price move, the more forecast models tend to miss the extent of the price shift.

A possible explanation for the generally low correlations is the periodic importance of shock variables to forecast errors. A relatively long period was evaluated here and each of these shocks may only influence forecast error occasionally throughout the sample. However, the correlations still provide some level of insight into potential drivers of error. Experience tells that shocks or deviations from the normal pattern in carcass weights and cattle on feed over 150 days are often related to currentness in the fed cattle market. According to these results, unexpected changes to currentness clearly appear to be related to forecast errors. Consumption of beef, substitute meats and all meats (beef plus main substitutes) also appear important. Interestingly, neither beef production or fed slaughter numbers show much relationship with forecast errors.

Table 10: Correlations between forecast errors and market shock variables at each horizon.

	1-Month Horizon Models				3-Month Horizon Models			
Market Shock Factor	AR2_SF	VAR5_FP	VEC5	FUT	AR2_SF	VAR5_FP	VEC4_Cn	FUT
Fed slaughter	0.11	0.11	**0.21**	-0.12	-0.08	-0.07	0.11	-0.17
Carcass weights	**-0.25**	**-0.23**	**-0.24**	**-0.22**	**-0.48**	**-0.41**	**-0.56**	**-0.38**
Beef production	0.05	0.07	0.17	**-0.21**	-0.16	-0.13	0.05	**-0.27**
Net beef trade	-0.15	-0.15	**-0.23**	0.17	-0.15	-0.12	**-0.36**	0.05
Beef exports	0.15	0.14	0.09	**0.32**	**0.25**	**0.20**	0.13	**0.31**
Beef consumption	-0.02	0.00	0.04	-0.13	**-0.22**	-0.18	-0.13	**-0.24**
Substitute consumption	-0.11	-0.13	-0.09	-0.12	**-0.22**	**-0.24**	**-0.25**	-0.16
Meat consumption	-0.05	-0.05	-0.02	-0.11	**-0.22**	**-0.22**	**-0.24**	-0.17
COF > 150 days	-0.14	-0.10	-0.08	**-0.51**	**-0.33**	**-0.28**	**-0.22**	**-0.48**
Disposable income	-0.11	-0.13	-0.08	**-0.23**	-0.20	-0.19	-0.13	**-0.23**
Momentum 2-6	0.19	0.19	0.18	0.11	**0.30**	**0.26**	**0.38**	0.19
St. dev. 12-mo	**0.39**	**0.35**	**0.31**	**0.46**	**0.40**	**0.34**	**0.41**	**0.41**

	6-Month Horizon Models				9-Month Horizon Models			
Market Shock Factor	AR3_SF	VAR4_CnFP	VEC4_Cn	FUT	AR3_S	VAR4_Cn	VEC4_Cn	FUT
Fed slaughter	-0.06	0.02	0.03	-0.14	-0.11	-0.04	-0.05	-0.19
Carcass weights	**-0.67**	**-0.66**	**-0.66**	**-0.43**	**-0.68**	**-0.72**	**-0.68**	**-0.42**
Beef production	-0.10	-0.03	-0.05	**-0.25**	-0.15	-0.10	-0.12	**-0.28**
Net beef trade	**-0.36**	**-0.37**	**-0.35**	0.02	**-0.32**	**-0.37**	**-0.33**	0.00
Beef exports	0.20	0.17	**0.30**	**0.34**	**0.20**	0.19	**0.31**	**0.34**
Beef consumption	**-0.25**	-0.19	**-0.20**	**-0.22**	**-0.25**	**-0.25**	**-0.26**	**-0.26**
Substitute consumption	**-0.23**	**-0.20**	**-0.24**	-0.03	**-0.26**	**-0.26**	**-0.27**	-0.06
Meat consumption	-0.11	-0.05	**-0.20**	-0.07	**-0.27**	**-0.24**	**-0.26**	-0.11
COF > 150 days	**-0.26**	**-0.30**	**-0.36**	**-0.47**	**-0.24**	**-0.30**	**-0.30**	**-0.37**
Disposable income	-0.16	-0.20	-0.19	-0.18	-0.18	**-0.24**	**-0.22**	-0.16
Momentum 2-6	**0.50**	**0.50**	**0.49**	0.20	**0.54**	**0.63**	**0.53**	0.20
St. dev. 12-mo	**0.66**	**0.53**	**0.57**	**0.50**	**0.89**	**0.78**	**0.78**	**0.66**

Note: Bolded values indicate correlations greater than 0.20 in absolute value.

Section 5.4.2: Regression Analysis

Forecast errors are regressed on a combination of the shock variables to evaluate the degree to which forecast error can be explained and which shock variables significantly contribute to forecast errors. Forecast error regressions take the general form of Equation (21). Many of the shock variables analyzed in the previous section contain some of the same information by construction. To reduce collinearity in forecast error regression, consideration was taken to select combinations of shock variables that are unique but contain all relevant information. Selection of market shock variables was made based on the correlations between market shocks and forecast error in the previous section and correlations between market shock variables, provided in Table 11 below.

Table 11: Correlations between shock variables.

	fedsltr	cxwgt	beefprod	netbeeftrade	beefexports	beefcons	subcons	meatcon	cof150	income
fedsltr	1	-0.21	0.96	-0.39	0.23	0.79	0.60	0.47	0.14	0.11
cxwgt	-0.21	1	-0.14	0.63	-0.70	0.15	0.23	0.25	0.11	0.17
beefprod	0.96	-0.14	1	-0.44	0.29	0.81	0.58	0.44	0.26	0.20
netbeeftrade	-0.39	0.63	-0.44	1	-0.86	0.11	0.27	0.30	-0.12	-0.03
beefexports	0.23	-0.70	0.29	-0.86	1	-0.14	-0.30	-0.34	0.03	-0.05
beefcons	0.79	0.15	0.81	0.11	-0.14	1	0.81	0.66	0.30	0.20
subcons	0.60	0.23	0.58	0.27	-0.30	0.81	1	0.97	0.16	0.12
meatcon	0.47	0.25	0.44	0.30	-0.34	0.66	0.97	1	0.09	0.09
cof150	0.14	0.11	0.26	-0.12	0.03	0.30	0.16	0.09	1	0.65
income	0.11	0.17	0.20	-0.03	-0.05	0.20	0.12	0.09	0.65	1

In the regressions, beef production is dropped because it should be fully explained by fed slaughter and carcass weights. Net beef trade is dropped in favor of beef exports, since it is slightly less correlated with other variables and has a similar degree of correlation with forecast errors. Beef consumption and total meat consumption are also dropped because they are explained by other variables. The final specification of the error regression models takes the form of the following equation:

$$(25) \quad u_{t-h}^{t} = \beta_0 + \beta_1 fedsltr_t + \beta_2 cxwgt_t + \beta_3 beefexports_t + \beta_4 subcon_t + \beta_5 cof150_t + \beta_6 income_t + v_t$$

Coefficients describe the expected change in forecast error given a one-unit change in each shock variable, all else held constant. Meaningful conclusions can be drawn on the direction of the relationship between shocks and forecast error from the signs on the coefficients, describing whether the relationship is positive or negative. Regression results are presented in the tables below for analysis of errors at the one, three, six and nine month forecast horizons, in that order. Only variables significant at the 5% level or better will be described as significant in this analysis, however, significance is denoted to the 10% level in the tables below for the reader's reference.

As shown in Table 12, little forecast error is explained by the shock variables at the one-month horizon as indicated by low R^2 values for each regression. Less than 15% of variation in forecast errors in the three time series models is explained by variation in the shock variables. Interestingly, the R^2 for futures-based forecasts is much higher than the time series models at 31%.

Very few shock variables significantly contribute to forecast error at the one-month horizon. Fed slaughter, carcass weights, cattle on feed over 150 days and income are each only significant in one of four forecast models. As mentioned above, the coefficients cannot be directly interpreted with meaningful units as they represent the expected change in forecast error given a one-unit change in the error from the univariate shock models, holding all else constant. However, the signs on the coefficients indicate that errors from carcass weight and cattle on feed over 150 days are negatively related to forecast errors. In other words, positive shocks to carcass weights and cattle on feed over 150 days are related to lower realized prices than forecasted, consistent with expectations. The positive relationship between income and fed cattle prices is also as expected. The positive sign on fed slaughter is counterintuitive since we would expect increased slaughter to result in lower prices. A possible explanation is the short-term resistance of sellers of fed cattle in reaction to lower than expected prices.

Table 12: Regression of one-month horizon forecast errors on shock variables.

	1-Month Horizon Error Regressions			
	AR2_SF	VAR5_FP	VEC5	Futures
Shock variable	Coef.	Coef.	Coef.	Coef.
(Intercept)	0.090	-0.036	-0.188	0.674*
Fed slaughter	0.005	0.006	0.011**	-0.003
Carcass weights	-0.051	-0.040	-0.046	-0.139***
Beef exports	0.005	0.003	-0.002	-0.027
Substitute cons.	-0.674	-0.937	-1.173	-0.231
COF >150 days	-0.001	0.000	-0.001	-0.007***
Disp. income	0.000	0.000	0.000	0.001**
R-squared	9.4%	8.6%	11.7%	34.0%

Note: Single, double and triple asterisk (*) denote significance at the 10%, 5% and 1% levels.

Results for error regressions at the three-month forecast horizon are given in Table 13. The R^2 values indicate that on average around 30 to 35% of forecast error is explained by market shocks. Forecast errors from futures predictions remain the most explainable, although the R^2 is not substantially larger than the others. Shocks to carcass weights are significant to errors from all forecast models, shocks to cattle on feed over 150 days are important to errors from the AR and futures forecasts and fed slaughter errors are significant to futures errors. All significant coefficients have the expected signs.

Table 13: Regression of three-month horizon forecast errors on shock variables.

	3-Month Horizon Error Regressions			
	AR2_SF	VAR5_FP	VEC4_Cn	Futures
Shock variable	Coef.	Coef.	Coef.	Coef.
(Intercept)	-0.487	-1.546**	-1.187*	-0.048
Fed slaughter	-0.013	-0.005	0.009	-0.017**
Carcass weights	-0.327***	-0.298***	-0.396***	-0.354***
Beef exports	0.011	-0.018	-0.020	-0.031
Substitute cons.	0.106	-1.321	-2.453*	0.685
COF >150 days	-0.006***	-0.004*	-0.004*	-0.009***
Disp. income	0.001	0.000	0.001	0.002*
R-squared	32.5%	23.8%	36.3%	38.3%
Note: Single, double and triple asterisk (*) denote significance at the 10%, 5% and 1% levels.				

Shock variables appear much more related to forecast errors at the six-month horizon, as shown in Table 14. Carcass weights and cattle on feed over 150 days are significant to errors in all four models, and fed slaughter and beef exports are significant in three of four models. Substitute meat consumption is significant to errors from two of four models and shocks to disposable income are important to only futures errors at the 5% level. Signs are consistent with expectations except for substitute meat consumption. We expect to see positive shocks to consumption to be negatively related to prices, as fed cattle prices are forced lower to compete with the lower prices needed to coax consumers to consume additional substitute meat supplies, but we observed the opposite. Model fits are better compared to shorter horizon forecast errors with over 50% of forecast error explained in the time series models. Less forecast error from futures market predictions is explained than from the time series forecasts at the six-month horizon, but almost half of the error is still explained.

Table 14: Regression of six-month horizon forecast errors on shock variables.

6-Month Horizon Error Regressions				
	AR3_SF	VAR4_CnFP	VEC5_P	Futures
Shock variable	Coef.	Coef.	Coef.	Coef.
(Intercept)	-1.221	-1.934**	-3.410***	-0.308
Fed slaughter	-0.036***	-0.022**	-0.014	-0.040***
Carcass weights	-0.531***	-0.511***	-0.516***	-0.591***
Beef exports	0.140***	0.124***	0.106**	0.002
Substitute cons.	3.414**	2.523	0.834	5.783***
COF >150 days	-0.008***	-0.009***	-0.013***	-0.012***
Disp. income	0.002	0.002	0.003*	0.004***
R-squared	54.5%	50.3%	51.9%	43.9%
Note: Single, double and triple asterisk (*) denote significance at the 10%, 5% and 1% levels.				

Results for the error regressions of nine-month horizon forecasts are presented in Table 15. Compared to shorter horizons, the greatest amount of variation in forecast error is explained for the time series models by shock variables at the nine-month horizon, with R^2 values of 55 to 60%. However, the R^2 for futures errors decreased to 38%. The same individual shock variables appear important to nine-month horizon forecast errors as to six-month errors. Fed slaughter and carcass weights are significant in all four error models, and beef exports and cattle on feed over 150 days are significant in three of four models. Substitute meat consumption is significant to the AR and futures errors and is again the incorrect sign.

Several major points emerge from error regressions regarding the significance of individual shock factors and general trends in forecast error explanation. Individually, shocks to carcass weights appear the most important to forecast errors, being significant at 5% or better in 13 of 16 error regressions. Fed slaughter numbers and cattle on feed over 150 days stand out as the next most important shocks with significance at the 5% level 9 and 10 times, respectively. The importance of fed slaughter in the regression analysis is surprising given the low correlations between slaughter and price forecasts in the previous section. Beef exports and substitute meat consumption are also significant in some of the extended horizon error models. However, substitute meat consumption was the opposite sign than expected when significant and an explanation is not immediately clear.

Table 15: Regression of nine-month horizon forecast errors on shock variables.

9-Month Horizon Error Regressions				
	AR3_S	VAR4_Cn	VEC4_Cn	Futures
Shock variable	Coef.	Coef.	Coef.	Coef.
(Intercept)	-1.095	-1.803	-4.411***	-0.089
Fed slaughter	-0.056***	-0.041***	-0.042***	-0.058***
Carcass weights	-0.739***	-0.792***	-0.754***	-0.642***
Beef exports	0.164***	0.148***	0.160***	0.052
Substitute cons.	4.604**	3.301	2.704	7.609***
COF >150 days	-0.007*	-0.009**	-0.010**	-0.010***
Disp. income	0.002	0.001	0.001	0.004**
R-squared	57.6%	60.5%	54.5%	38.0%

Note: Single, double and triple asterisk (*) denote significance at the 10%, 5% and 1% levels.

As mentioned earlier in the correlation analysis, shocks to carcass weights and cattle on feed over 150 days are commonly associated with changes to currentness. The relationship between currentness and prices in the fed cattle market follows a well-known narrative. As the pace of fed cattle marketings for harvest are unexpectedly slowed, cattle spend more days on feed. As a non-storable commodity, the feedlot operator is limited on the additional time cattle can be held before they need to be marketed. Once the feedlot operator gets too behind on marketings, bargaining position is lost as the non-storable commodity needs to be processed. As a result, the meatpacker gains leverage and the short-term market power allows prices to be pushed downward. Larger carcass weights result from additional days on feed, also potentially contributing to packer bargaining position as more pounds of beef per head partially offsets the number of head needed to be harvested. Of course, the opposite of this scenario can result in higher than expected prices as the feedyard's bargaining position is improved with increased currentness.

These results add empirical evidence to the significance of the effects of currentness in the fed cattle market to prices. Understanding and anticipating the trends in currentness appears vital for sound cattle price forecasts. The concept of currentness appears to be well proxied by shocks to trends in carcass weights and cattle on feed data, providing a quantitative method to monitor this somewhat qualitative market condition.

As shown in the error regression results, R^2 values indicate that generally more of the variation in forecast error can be explained by market shocks as forecast horizon increases. One explanation is that the longer time frame between forecasting and price realization leaves more opportunity for shocks to occur, and therefore more error is explained by shocks than randomness. By comparison, less forecast error is explained by shocks at shorter forecast horizons and a greater portion of forecast errors are due to randomness. This is consistent with concepts of the price discovery process. However, another potential explanation is that the increased R^2 values are the artificial result of there being more error to explain at the more extended horizons (see forecast error by horizon in Figure 8). In other words, a greater degree of variation in forecast error makes for greater potential for shocks to explain variation in errors.

Lastly, there are significant intercepts for forecast errors in the error regressions of four time series models. The intercept reports the average forecast error, holding all the shock variables at zero. This indicates that either there is a bias in these forecast models or they represent the average effect of other factors influencing forecast error that have not been included in the regressions. Since the set of shock variables used in this analysis encompass all major fundamental forces that may contribute to forecast error, it appears that these four forecasts, even when accounting for shocks, tend to over-predict fed cattle prices at their respective horizons.

Several questions for further research emerge from our analysis of forecast error. First, research is warranted into the ability to anticipate currentness in the cattle market and the potential of improving forecasts by incorporating concepts of currentness into fed cattle price forecasting models. Second, our methods define market shocks in an ex post context. Defining shocks as the difference between purely ex ante expectations and realized outcomes may be useful. This would require more robust analysis and forecasting of each series than is performed in this analysis. Nonetheless, our approach still yields useful insights regarding the relationship between shocks and forecast errors in the fed cattle market and provides a basis for future work.

CHAPTER 6: CONCLUSIONS AND IMPLICATIONS

Cattle markets, along with many other agricultural commodities, have faced high levels of price volatility in recent years. This market environment makes establishing accurate price forecasts both more important and increasingly difficult for producers. The objective of this research has been to evaluate the performance of econometric time series forecasts in this recent period of high volatility and investigate the driving fundamental factors behind forecast errors. Forecasts and forecast errors from futures market predictions are also evaluated. This chapter first discusses conclusions regarding forecast performance, followed by the conclusions from the forecast error analysis. The final section contains a discussion of the implications of this research and suggestions for further research.

6.1 Forecast Performance

Time series methods appear valuable from a forecasting standpoint and can generally out-perform futures market predictions at nearby forecast horizons. When the forecast accuracy of competing forecast models was tested at the one, three, six and nine-month forecast horizon, the only statistically significant differences in accuracy was that all three time series models (AR, VAR and VEC) were more accurate than futures market predictions at the one-month horizon. At around three-months and beyond, futures markets appear to provide better forecasts than the time series methods. Futures market predictions are increasingly more accurate than time series models as forecast horizon increases, although the differences are never statistically significant. We expect that futures would have an even greater accuracy advantage if a basis adjustment had been made. Overall, these results are consistent with the bulk of the literature and support the generally held belief that futures markets, although not good predictors, are often the best forecasts in a comparative sense, especially at longer forecast horizons. This has positive market efficiency implications.

Incorporating futures market information appears important to forecast accuracy at shorter horizons, but not at intermediate and longer horizons as the best models at these horizons do not include futures information. This result is somewhat surprising, as we would expect including futures to become more important for accuracy at longer horizons since futures market predictions by themselves become increasingly more accurate than the time series methods as forecast horizon lengthens. Forecast encompassing results indicate that at intermediate horizons (three to six months) a composite between futures and time series forecasts may be a better method than futures alone. VAR forecasts in particular appear promising, as they contain statistically significant incremental information at both the three and six-month horizon. However, encompassing results indicate little incremental information of any time series method at the nine-month horizon. At this horizon, futures appear to be the best individual forecast and entirely encompass predictive information provided by time series methods.

When comparing the three types of econometric time series models used here, we find that simpler models can generally out-perform more robust time series models, as the AR models developed are generally more accurate than the VAR and VEC model specifications. The exception is that the AR and VAR models have virtually identical accuracy at the one-month horizon. Further, forecast encompassing tests indicate the AR and VAR models contain the same information; the optimal weight of the VAR in a composite with the AR forecasts are approximately 50% but with little or no statistical significance.

The advantage of simpler models becomes larger at longer forecast horizons, although there are no statistically significant differences between AR forecasts and VAR and VEC forecasts. Additionally, the increased complexity of the error-correcting terms of the VEC does not improve forecast accuracy compared to VAR models, adding further evidence to the value of simplicity in a forecasting context. Forecast encompassing tests largely confirm these conclusions. At the six and nine-month horizons, the more complex VAR and VEC forecast models contain no significant incremental information to the simpler AR forecasts.

We also investigate the implications of using disaggregated data on regional cattle markets to forecast fed cattle prices. Overall, differences in accuracy are very small compared to using national market data. At nearby forecast horizons, forecast accuracy within regions is generally less accurate than forecasts of the national market prices. However, at longer horizons forecasting performance is improved slightly in the Kansas and Texas-Oklahoma regions and there is mixed accuracy improvement in the Nebraska region. The commonality is that these regions are all considered “leading markets” in setting national prices. The price series of these leading markets appear to have more defined and predictable relationships through time which is consistent with their identification as leading markets.

Forecasting national prices by aggregating regional forecasts shows only minor differences in accuracy. Using regionally disaggregated data improves accuracy of most models we tested, but did not improve the accuracy of the most accurate models. Most models appear to benefit from the averaging of price forecasts by reducing the impact of a poor forecast from one regional model, consistent with the commonly cited benefits of compositing separate forecasts in the literature. This may also be the result of improved forecasting accuracy in the price-leading regional markets. Regional disaggregation of data, however, does not improve the maximum forecast accuracy achieved by the best forecasts. The best forecasts from aggregated regional forecasts show only negligible improvements in accuracy at closer horizons and negligible declines at intermediate and longer horizons compared to the best national models. Regional disaggregation may help the average forecast model, but it clearly does not improve our best efforts to foresee prices into the future.

6.2 Forecast Errors

The results of this analysis show that shocks to fundamental factors in the fed cattle market are significantly related to errors in price forecasts. This is consistent with concepts of market efficiency and price discovery. Generally, a greater percentage of forecast error is explained as forecast horizon increases. Shocks to carcass weights, cattle on feed over 150 days and fed slaughter numbers are the most frequently found to be significant to price forecast errors.

Two of the most important fundamental factors related to forecast errors are shocks to carcass weights and number of cattle on feed over 150 days. Shocks to these variables are commonly associated with currentness in the fed cattle market, or the rate of cattle being marketed compared to market readiness of those cattle. When the market is current, cattle are being marketed ahead of schedule and packers must bid more aggressively to get feedlot operators to sell cattle that could otherwise be appreciating in value in the feedlot. In an uncurrent market, cattle are behind schedule for marketing and are costing the feedlot operator to hold as inventory. Meat packers can then offer lower bids to acquire fed cattle. The short-term market power effects of the currentness in the market can therefore have a considerable influence on price levels relative to expectation. The results of this analysis add empirical evidence to the significance of currentness in the fed cattle market to prices. Understanding and anticipating the trends in currentness appears vital for sound cattle price forecasts. Currentness appears to be well proxied by shocks to trends in carcass weights and cattle on feed data.

As shown in the error regression results, R^2 values indicate that generally more of the variation in forecast error can be explained by market shocks as forecast horizon increases. One explanation is that the longer time frame between forecasting and price realization leaves more opportunity for shocks to occur, and therefore more error is explained by shocks as compared to the randomness of the stochastic process. By contrast, less forecast error is explained by shocks at shorter forecast horizons and a greater portion of forecast errors are due to randomness. This conclusion has an intuitive appeal and is consistent with concepts of the price discovery process. However, another possibility for the increasing R^2 values is the increasing levels of error at more extended forecast horizons. In other words, is there a greater potential for shocks to explain variation in forecast errors because there is more variation to be explained? The answer may be a combination of both explanations and therefore caution should be taken when comparing R^2 of the error regressions across forecast horizons or between models due to differences in accuracy.

Some forecast models appear to be biased as shown by significant intercepts in the error regressions of four time series models. The intercept reports the average forecast error, holding all the shock variables at zero. This indicates that either there is a bias in these forecast models or they represent the average effect of other factors influencing forecast error that have not been included in the regressions. The set of shock variables used in this analysis was selected to largely encompasses the major fundamental forces that may contribute to forecast error, but important shocks could have been excluded. If no other significant shocks are in play, it appears that these four forecasts tend to over-predict fed cattle prices at these specific horizons, even when accounting for market shocks.

Forecast errors are also related to the volatility and swiftness of price movements in the market. Correlations indicate that forecast errors from the time series models at the six- and nine-month forecast horizons are moderately related to the swiftness of price movements, with smaller correlations at the three-month horizon. In this way, forecast errors from the time series models tend to be larger during periods of swift price movements as it appears forecasts don't anticipate the large movements. However, errors from futures market predictions do not appear to be as strongly related to the momentum of price movements.

In terms of the volatility of prices, both time series and future market predictions tend to have larger errors during periods of greater volatility. At longer and intermediate horizons, forecast errors from the more complex VAR and VEC model specifications are less related to volatility than the AR models. If forecasts from multiple-equation models are less prone to errors due to volatility, this would suggest these models may have an advantage during periods of higher volatility, despite our findings that the simpler AR models have an accuracy advantage overall.

6.3 Implications and Further Research

The present research finds that previous conclusions on cattle price forecast performance largely continue to apply in this recent volatile environment. Namely, that futures can be improved upon at nearby horizons but are superior at more extended horizons. This has positive market efficiency implications. While it is reasonable for forecast users to be concerned whether their forecasts are accurate, they can be reassured that simply using futures-based forecasts are providing accurate forecasts relative to other available forecasts. Our use of futures information for forecasting is very simple and it is anticipated that more robust methods, especially using basis adjustments, would result in futures-based forecasts that have an even greater accuracy advantage compared to time series methods.

Nonetheless, time series methods appear valuable from a forecasting standpoint in several applications. A well-specified AR model is independently the most accurate of the time series methods we evaluated, especially at longer horizons. In a compositing sense, VAR forecasts are particularly useful complements to other forecasts. Although AR forecasts are more accurate alone, VAR forecast models contain incrementally useful information when composited with futures-based forecasts and AR forecasts, especially at horizons of around six months and shorter. These findings are somewhat contradictory, but indicate that practitioners may be best suited to use AR models when a separate forecast is needed but a VAR model when the forecast is to be combined with other forecasts. However, at longer forecast horizons futures market predictions appear to be the most useful information and time series methods offer very little added valuable information.

The decline in forecast accuracy for many forecast models when using a rolling estimation framework versus a single estimation was surprising. We hypothesize that the decline is related to a changing drift term in the models as the mean of the price series changes over time or related to changing basis information over time. This may suggest a change in the nature of the stochastic process that makes it more difficult to forecast. This question warrants further research, including whether manipulating the drift term in forecasts can improve forecast accuracy. Can a forecaster improve accuracy by selecting different drift terms in upward price cycles than downward cycles? Additionally, further research is warranted on basis predictability through the last decade. Basis observations in recent years may not be reflective of the underlying data generating process, possibly due to the highly volatile market environment, and need to be excluded from samples used for forecasting future basis levels.

Our analysis of forecasting with regional data shows that forecasters should stick with national models unless one of the leading markets (namely Kansas or Texas-Oklahoma) is of specific interest since the national market prices can generally be forecasted more accurately. Using disaggregated regional data to forecast national prices improves forecasts only for mediocre models, and this appears to be due to the averaging of forecasts rather than improved modeling of the underlying data generating process. Forecasting efforts are more wisely spent developing accurate national models than building national forecasts from regional models.

This research has shown that shocks to some fundamental factors in the fed cattle market significantly contribute to price forecasting errors, the most important factors being carcass weights, cattle on feed over 150 days and fed slaughter numbers. The importance of shocks to carcass weights and cattle on feed over 150 days implies that modelling and incorporating predictions of market currentness appears the most important fundamental factor to investigate. Based on the results of this research, understanding and accounting for an expectation of market currentness shows the most potential for improving efforts to forecast fed cattle prices.

Future efforts to improve forecasting may include incorporating these variables directly into forecast models or by forecast adjustment. In the case of forecast adjustment, the error regression framework used in this analysis presents a potential mechanism to make such adjustments by incorporating expected deviations of other variables from trendline and autoregressive-based patterns. Further research is warranted on the efficiency of forecasting these other variables and adjusting fed cattle forecasts. Of course, forecasts can only be improved to the extent that the future unfolds as expected. Accurate forecasting provides critical information to producers but these results indicate that forecast errors will always be impacted by random errors and unforeseen market shocks.

REFERENCES

Allen, P.G. 1994. "Economic Forecasting in Agriculture." *International Journal of Forecasting*. 10:81-125.

Bailey, D.V., and B.W. Brorsen. 1985. "Dynamics of Regional Fed Cattle Prices." *Western Journal of Agricultural Economics*. 10(1):126-133.

Bechtel, W. 2017. "Rabobank: Cow Herd to Continue Increasing Through 2020." *Drovers Cattle Network,* in press. http://www.cattlenetwork.com/news/industry/rabobank-cow-herd-continue-increasing-through-2020.

Clark, T.E., and M.W. McCracken. 2009. "Improving Forecast Accuracy by Combining Recursive and Rolling Forecasts." *International Economic Review*. 50(2):363-395.

Darko, F.A., and J.S. Eales. 2013. "Meat Demand in the US During and After the Great Recession." *Selected paper for the Agricultural & Applied Economics Association 2013 Annual Meeting*.

Etienne, X. L., S. H. Irwin, and P. Garcia. 2014. "A Structural Approach to Disentangling Speculative and Fundamental Influences on the Price of Corn." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. http://www.farmdoc.illinois.edu/nccc134.

Garcia, P., R.M. Leuthold, R.T. Fortenbery and F. Sarassoro. 1988. "Pricing Efficiency in the Live Cattle Futures Market: Measurement and Interpretation." *American Journal of Agricultural Economics*. 70: 162-169.

Gee, Kelsey. 2016. "Welcome to the 'Meat Casino'! The Cattle Futures Market Descends Into Chaos." *The Wall Street Journal,* in press. https://www.wsj.com/articles/welcome-to-the-meat-casino-the-cattle-futures-market-descends-into-chaos-1471475438.

Granger, C.W.J., and P. Newbold. 1973. "Some Comments on the Evaluation of Economic Forecasts." *Applied Economics*. 5:35-47.

Gujarati, D.N. and D.C. Porter. 2009. *Basic Econometrics*. New York: McGraw Hill Book Co.

Harvey, D.I., S.J. Leybourne and P. Newbold. 1997. "Testing the Equality of Prediction Mean Squared Errors." *International Journal of Forecasting*. 13(2):281-291.

Harvey, D.I., S.J. Leybourne and P. Newbold. 1998. "Tests for Forecast Encompassing." *Journal of Business and Economic Statistics*. 16(2):254-259.

Johansen, S. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. New York: Oxford University Press.

Kastens, T.L., T.C. Schroeder, and R. Plain. 1998. "Evaluation of Extension and USDA Price and Production Forecasts." *Journal of Agricultural and Resource Economics*. 23: 244-261.

Koontz, S.R. 2016. "Objective Measures of Price Discovery in Thinning Fed Cattle Markets" Unpublished research for the National Cattlemen's Beef Association and Livestock Marketing Information Center. Colorado State University.

Livestock Marketing Information Center (LMIC). *Analysis and Comments: U.S. Cattle Market Drivers*. State Extension Services in Cooperation with U.S. Department of Agriculture: Letter #50. 15 December 2016.

Mulvany, Lydia. 2016. "CME Says Cattle Market May Be 'Broken' as Volatility Soars." *Bloomberg,* in press. https://www.bloomberg.com/amp/news/articles/2016-02-01/cme-says-cattle-market-may-be-broken-as-price-volatility-soars.

Meyer, George. 2016. "Cattlemen Lock Horns with Futures Exchange over Market Volatility." *Financial Times,* in press. https://www.ft.com/content/6eed1268-c130-11e5-846f-79b0e3d20eaf.

Nordhaus, W.D. 1987. "Forecasting Efficiency: Concepts and Applications." *Review of Economics and Statistics*. 69(4):667-674.

Oliveira, R.A., C.W. O'Connor, and G.W. Smith. 1979. "Short-Run Forecasting Models of Beef Prices." *Western Journal of Agricultural Economics* 4:45-55.

Pritchett, J., K. Johnson, D. Thilmany and W. Hahn. 2007. "Consumer Responses to Recent BSE Events." *Journal of Food Distribution Research* 38(2):57-68.

Purcell, W.D., and S.R. Koontz. 1999. *Agricultural Futures and Options*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall.

Sanders, D.R. and M.R. Manfredo. 2005. "Forecast Encompassing as the Necessary Condition to Reject Futures Market Efficiency: Fluid Milk Futures." *American Journal of Agricultural Economics* 87(3):610-620.

Sanders, D.R. and M.R. Manfredo. 2003. "USDA Livestock Price Forecasts: A Comprehensive Evaluation." *Journal of Agricultural and Resource Economics* 28(2):316-334.

Stein, J. L. 1981. "Speculative Price: Economic Welfare and the Idiot of Chance." *The Review of Economics and Statistics* 63:223-232.

Zapata, H.O. and P. Gracia. 1990. “Price Forecasting with Time-Series Methods and Nonstationary Data: An Application to Monthly U.S. Cattle Prices.” *Western Journal of Agricultural Economics* 15:123-132.

Yang, J., R.B. Balyeat, and D.J. Leatham. 2005. “Futures Trading Activity and Commodity Cash Price Volatility.” *Journal of Business Finance and Accounting*. 32(1&2):297-323.

APPENDIX A

Auxiliary regressions: regression demonstrating structural change with USDA Cattle on Feed report change and regression model used to forecast feedlot placements.

Table A1: Regression of cattle on feed data on a trend, monthly dummy variables and mean shift for report change starting December 1991, Jan. 1990-Dec. 2016.

Residuals:

Min	1Q	Median	3Q	Max
-1971.71	-357.57	12.96	398.12	1211.62

Coefficients	Estimate	Std. Error	t-value	p-value	Signif.
(Intercept)	10685.46	137.31	77.82	0.0000	***
old_cof	-1924.91	103.76	-18.55	0.0000	***
trend	2.69	0.46	5.87	0.0000	***
season2	-34.88	146.59	-0.24	0.8121	
season3	-167.46	146.59	-1.14	0.2542	
season4	-207.01	146.59	-1.41	0.1589	
season5	-521.78	146.60	-3.56	0.0004	***
season6	-625.36	146.61	-4.27	0.0000	***
season7	-1127.24	146.61	-7.69	0.0000	***
season8	-1383.31	146.62	-9.43	0.0000	***
season9	-1322.56	146.63	-9.02	0.0000	***
season10	-842.99	146.65	-5.75	0.0000	***
season11	-83.02	146.66	-0.57	0.5718	
season12	106.47	146.63	0.73	0.4683	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 538.6 on 310 degrees of freedom
Multiple R-squared: 0.8189,   Adjusted R-squared: 0.8113
F-statistic: 107.8 on 13 and 310 DF, p-value: < 2.2e-16

Table A2: Placement forecasting model, January 1990-December 2006.

Residuals:

Min	1Q	Median	3Q	Max
-388.32	-89.54	0.67	89.01	352.71

Coefficients:

	Estimate	Std. Error	t-value	p-value	Signif.
(Intercept)	-62.62	440.57	-0.14	0.8871	
plmt.l1	0.34	0.05	6.57	0.0000	***
plmt.l11	0.14	0.05	2.54	0.0115	*
plmt.l12	0.20	0.06	3.64	0.0003	***
cows.l6	0.02	0.01	1.99	0.0471	*
oldcof	-81.82	45.52	-1.80	0.0733	.
season2	-289.58	47.18	-6.14	0.0000	***
season3	-6.16	39.09	-0.16	0.8749	
season4	-337.00	51.05	-6.60	0.0000	***
season5	115.15	39.93	2.88	0.0042	**
season6	-366.48	49.18	-7.45	0.0000	***
season7	-123.44	47.70	-2.59	0.0101	*
season8	23.16	51.96	0.45	0.6561	
season9	68.96	62.93	1.10	0.2741	
season10	298.93	57.73	5.18	0.0000	***
season11	-253.24	62.25	-4.07	0.0001	***
season12	-305.34	48.65	-6.28	0.0000	***
trend	0.09	0.29	0.31	0.7590	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 140.7 on 294 degrees of freedom
Multiple R-squared: 0.8513, Adjusted R-squared: 0.8427
F-statistic: 99.02 on 17 and 294 DF, p-value: < 2.2e-16

APPENDIX B

Forecast Model Specifications

Forecasts from all candidate forecast models (48 total) are generated and evaluated in a single estimation framework. Forecast models with **bolded equation names** indicate equations re-evaluated in a rolling estimation framework (20 models). Lastly, ***bolded and italicized equation names*** indicate forecast models estimated and evaluated on regional data, also in a rolling estimation framework (10 models). Autoregressive models are grouped by similar exogenous independent variables and VAR and VEC models are grouped by similar endogenous variables. Coefficients represented by lower case letters are single parameters and capital letters denote vectors or matrices of parameters. General forms of models are described in Chapter 3. The full names and descriptions for the variable symbols can be found in Table 1.

AR Models

AR: Simple and seasonal

AR(1-3, 6)

(1) $$\Delta pfed_t = \alpha_0 + A_i \begin{bmatrix} \Delta pfed_{t-1} \\ \Delta pfed_{t-2} \\ \Delta pfed_{t-3} \\ \Delta pfed_{t-6} \end{bmatrix} + u_t$$

AR(3)_S6

(2) $$\Delta pfed_t = \alpha_0 + A_i \begin{bmatrix} \Delta pfed_{t-1} \\ \Delta pfed_{t-2} \\ \Delta pfed_{t-3} \end{bmatrix} + B \begin{bmatrix} Mar_t \\ Apr_t \\ June_t \\ July_t \\ Aug_t \\ Nov_t \end{bmatrix} + u_t$$

AR(3)_S12

$$(3)\qquad \Delta pfed_t = \alpha_0 + A_i \begin{bmatrix} \Delta pfed_{t-1} \\ \Delta pfed_{t-2} \\ \Delta pfed_{t-3} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t$$

AR: Futures only

AR(2)-F

$$(4)\qquad \Delta pfed_t = \alpha_0 + A_i \begin{bmatrix} \Delta pfed_{t-1} \\ \Delta pfed_{t-2} \end{bmatrix} + \beta fut_{t-1}^{t} + u_t$$

AR(3)-F

$$(5)\qquad \Delta pfed_t = \alpha_0 + A_i \begin{bmatrix} \Delta pfed_{t-1} \\ \Delta pfed_{t-2} \\ \Delta pfed_{t-3} \end{bmatrix} + \beta fut_{t-1}^{t} + u_t$$

AR: Futures and seasonalities

AR(2)_S6-F

$$(6)\qquad \Delta pfed_t = \alpha_0 + A_i \begin{bmatrix} \Delta pfed_{t-1} \\ \Delta pfed_{t-2} \end{bmatrix} + B \begin{bmatrix} Mar_t \\ Apr_t \\ June_t \\ July_t \\ Aug_t \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t$$

AR(3)_S6-F

$$(7) \qquad \Delta pfed_t = \alpha_0 + A_i \begin{bmatrix} \Delta pfed_{t-1} \\ \Delta pfed_{t-2} \\ \Delta pfed_{t-3} \end{bmatrix} + B \begin{bmatrix} Mar_t \\ Apr_t \\ June_t \\ July_t \\ Aug_t \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t$$

AR(2)_S12-F

$$(8) \qquad \Delta pfed_t = \alpha_0 + A_i \begin{bmatrix} \Delta pfed_{t-1} \\ \Delta pfed_{t-2} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t$$

AR(3)_S12-F

$$(9) \qquad \Delta pfed_t = \alpha_0 + A_i \begin{bmatrix} \Delta pfed_{t-1} \\ \Delta pfed_{t-2} \\ \Delta pfed_{t-3} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t$$

AR: Futures, seasonalities and placements

AR(2)_S12-FP

$$(10) \qquad \Delta pfed_t = \alpha_0 + A_i \begin{bmatrix} \Delta pfed_{t-1} \\ \Delta pfed_{t-2} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t$$

VAR Models

VAR3: pfed, beefcon, pfeeder

VAR3

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \end{bmatrix} = A_0 + \textstyle\sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t \tag{1}$$

VAR3-FP

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \end{bmatrix} = A_0 + \textstyle\sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t \tag{2}$$

VAR3: pfed, beefcon, pcorn

VAR3_Corn

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pcorn_t \end{bmatrix} = A_0 + \textstyle\sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pcorn_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t \tag{3}$$

VAR3_Corn-FP

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pcorn_t \end{bmatrix} = A_0 + \textstyle\sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pcorn_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t \tag{4}$$

VAR3: pfed, beefcon, cof

VAR3_COF-FP

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ cof_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t \tag{5}$$

VAR4_Corn: pfed, beefcon, pfeeder, pcorn

VAR4_Corn

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t \tag{6}$$

VAR4_Corn-F

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t \tag{7}$$

VAR4_Corn-FP

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t \tag{8}$$

VAR4_Corn-P

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ plmts_{t-5} \end{bmatrix} + u_t \tag{9}$$

VAR4_COF: pfed, beefcon, pfeeder, cof

VAR4_COF

$$(10)\quad \begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ cof_t \end{bmatrix} = A_0 + \textstyle\sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ cof_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t$$

VAR4_COF-F

$$(11)\quad \begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ cof_t \end{bmatrix} = A_0 + \textstyle\sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ cof_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t$$

VAR4_COF-FP

$$(12)\quad \begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ cof_t \end{bmatrix} = A_0 + \textstyle\sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ cof_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t$$

VAR4_COF-P

$$(13)\quad \begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ cof_t \end{bmatrix} = A_0 + \textstyle\sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ cof_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ plmts_{t-5} \end{bmatrix} + u_t$$

VAR5: pfed, beefcon, pfeeder, pcorn, cof

VAR5

$$(14)\quad \begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ cof_t \end{bmatrix} = A_0 + \textstyle\sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ cof_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t$$

VAR5-F

$$(15)\quad \begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ cof_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t$$

VAR5-FP

$$(16)\quad \begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ cof_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t$$

VAR5-P

$$(17)\quad \begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ cof_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ plmts_{t-5} \end{bmatrix} + u_t$$

VAR5_I: pfed, beefcon, pfeeder, pcorn, income

VAR5_I

$$(18)\quad \begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ \Delta income_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ \Delta income_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t$$

VAR5_I-F

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ \Delta income_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ \Delta income_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t \tag{19}$$

VAR5_I-FP

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ \Delta income_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ \Delta income_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t \tag{20}$$

VAR5_I-P

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ \Delta income_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ \Delta income_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ plmts_{t-5} \end{bmatrix} + u_t \tag{21}$$

VAR6: pfed, beefcon, pfeeder, pcorn, cof, income

VAR6

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ cof_t \\ \Delta income_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ cof_{t-i} \\ \Delta income_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t \tag{22}$$

VAR6-F

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ cof_t \\ \Delta income_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ cof_{t-i} \\ \Delta income_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t \tag{23}$$

VAR6-FP

$$\begin{bmatrix} \Delta pfed_t \\ beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ cof_t \\ \Delta income_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ cof_{t-i} \\ \Delta income_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t \tag{24}$$

VAR6D: pfed, beefcon, pfeeder, pcorn, cof, income

VAR6D

$$\begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ \Delta cof_t \\ \Delta income_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ \Delta cof_{t-i} \\ \Delta income_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t \tag{25}$$

VAR6D-FP

$$\begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ \Delta cof_t \\ \Delta income_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ \Delta cof_{t-i} \\ \Delta income_{t-i} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t \tag{26}$$

VEC Models

VEC4_Corn: pfed, beefcon, pfeeder, pcorn

VEC4_Corn

$$\begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ pcorn_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t \tag{1}$$

VEC4_Corn-F

$$\begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ pcorn_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t \tag{2}$$

VEC4_Corn-FP

$$\begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ pcorn_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t \tag{3}$$

VEC4_Corn-P

$$\begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ pcorn_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ plmts_{t-5} \end{bmatrix} + u_t \tag{4}$$

VEC4_COF: pfed, beefcon, pfeeder, cof

VEC4_COF

$$(5)\quad \begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta cof_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ cof_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t$$

VEC4_COF-F

$$(6)\quad \begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta cof_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ cof_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t$$

VEC4_COF-FP

$$(7)\quad \begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta cof_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ cof_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t$$

VEC4_COF-P

$$(8)\quad \begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta cof_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ cof_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ plmts_{t-5} \end{bmatrix} + u_t$$

VEC5: pfed, beefcon, pfeeder, pcorn, cof

VEC5

$$(9)\quad \begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ \Delta cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ \Delta cof_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ pcorn_{t-1} \\ cof_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \end{bmatrix} + u_t$$

VEC5-F

$$(10)\quad \begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ \Delta cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ \Delta cof_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ pcorn_{t-1} \\ cof_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \end{bmatrix} + u_t$$

VEC5-FP

$$(11)\quad \begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ \Delta cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ \Delta cof_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ pcorn_{t-1} \\ cof_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ fut_{t-1}^{t} \\ plmts_{t-5} \end{bmatrix} + u_t$$

VEC5-P

$$(12)\quad \begin{bmatrix} \Delta pfed_t \\ \Delta beefcon_t \\ \Delta pfeeder_t \\ \Delta pcorn_t \\ \Delta cof_t \end{bmatrix} = A_0 + \sum_{i=1}^{p} A_i \begin{bmatrix} \Delta pfed_{t-i} \\ \Delta beefcon_{t-i} \\ \Delta pfeeder_{t-i} \\ \Delta pcorn_{t-i} \\ \Delta cof_{t-i} \end{bmatrix} + \Pi \begin{bmatrix} pfed_{t-1} \\ beefcon_{t-1} \\ pfeeder_{t-1} \\ pcorn_{t-1} \\ cof_{t-1} \end{bmatrix} + B \begin{bmatrix} Jan_t \\ \vdots \\ Nov_t \\ plmts_{t-5} \end{bmatrix} + \Gamma + u_t$$

APPENDIX C

Forecast accuracy of single-estimation and rolling-estimation forecast models based on root mean squared error (RMSE) and mean absolute percent error (MAPE).

Table A3: Forecast accuracy based on RMSE and MAPE of single estimation models by forecast horizon, Jan. 2007-Dec. 2016.

	RMSE by Forecast Horizon (Months)									**MAPE (%) by Forecast Horizon (Months)**								
Forecast Model	H1	H2	H3	H4	H5	H6	H7	H8	H9	H1	H2	H3	H4	H5	H6	H7	H8	H9
Futures	**4.68**	6.32	**7.75**	8.99	10.06	**10.99**	11.76	12.57	**13.34**	3.03	4.20	5.05	5.88	6.58	7.13	7.72	8.51	9.22
AR(3)-S(6)	4.50	6.94	8.59	9.89	10.95	11.96	13.08	14.29	15.45	2.93	4.60	5.72	6.45	7.14	7.78	8.62	9.45	10.20
AR(3)_S12	4.23	6.75	8.45	9.88	11.03	12.05	13.11	14.22	15.29	2.74	4.44	5.60	6.34	7.20	7.87	8.59	9.32	10.05
AR(2)-F	4.23	6.29	7.85	9.44	11.21	12.94	14.45	16.21	18.14	2.71	4.20	5.38	6.41	7.25	8.15	8.83	9.84	10.75
AR(3)-F	4.18	6.48	8.13	9.65	11.08	12.40	13.58	15.05	16.69	2.71	4.20	5.38	6.41	7.25	8.15	8.83	9.84	10.75
AR(2)_S6-F	4.23	6.35	7.81	9.28	10.84	12.25	13.48	14.96	16.52	2.73	4.25	5.16	6.18	7.06	8.00	8.64	9.72	10.61
AR(3)-S(6)F	4.20	6.54	8.11	9.56	10.88	12.01	13.05	14.36	15.77	2.70	4.29	5.36	6.28	7.10	7.85	8.45	9.45	10.32
AR(2)_S12-F	4.17	6.35	**7.82**	9.15	10.45	11.59	12.58	13.74	14.96	2.72	4.25	5.17	6.02	6.77	7.55	8.18	9.11	9.91
AR(3)_S12-F	**4.15**	6.54	8.14	9.49	10.62	**11.57**	12.47	13.54	**14.67**	2.71	4.30	5.39	6.22	6.98	7.64	8.22	8.98	9.91
AR(2)-FPS(12)	4.51	6.48	7.92	9.23	10.53	11.71	12.74	13.94	15.27	2.87	4.39	5.33	6.21	6.89	7.64	8.23	9.17	10.11
VAR5_I	4.35	6.63	7.94	9.24	10.59	11.70	12.84	13.91	14.74	2.87	4.56	5.45	6.20	7.17	7.87	8.51	9.22	9.68
VAR5_I-F	4.22	6.38	7.62	8.86	10.20	11.32	12.36	13.55	14.67	2.75	4.33	5.23	6.00	6.79	7.41	8.09	9.13	10.02
VAR5_I-FP	4.24	6.48	7.77	9.01	10.35	11.43	12.43	13.58	14.66	2.77	4.36	5.25	6.09	6.81	7.40	7.82	8.79	9.59
VAR5_I-P	4.39	6.78	8.17	9.50	10.86	11.97	13.11	14.18	15.03	2.90	4.64	5.58	6.44	7.41	8.15	8.92	9.61	9.98
VAR6D	4.41	7.16	9.43	11.56	13.73	15.46	17.25	19.35	21.00	2.88	4.85	6.51	8.01	9.36	11.01	12.24	13.50	14.42
VAR6D-FP	4.51	7.54	9.91	11.89	13.62	15.12	16.51	18.14	19.95	3.00	5.13	6.73	8.18	9.46	10.45	11.68	13.16	14.81
VAR6	4.01	6.52	8.79	11.45	14.15	16.24	18.78	21.03	23.57	2.67	4.58	6.43	8.12	9.74	11.45	13.33	15.00	16.57
VAR6-F	3.89	6.05	7.67	9.48	11.36	13.05	14.63	16.37	18.13	2.57	4.23	5.45	6.59	7.79	8.98	10.22	11.60	13.15
VAR6-FP	3.88	6.03	7.67	9.55	11.56	13.40	15.14	17.15	19.39	2.57	4.23	5.41	6.65	8.04	9.35	10.81	12.56	14.62
VAR5	4.01	6.51	8.78	11.42	14.10	16.17	18.70	20.94	23.46	2.67	4.57	6.42	8.09	9.69	11.39	13.27	14.92	16.47
VAR5-F	3.89	6.04	7.67	9.48	11.36	13.05	14.63	16.36	18.12	2.57	4.23	5.45	6.59	7.79	8.97	10.21	11.59	13.15
VAR5-FP	**3.88**	6.03	7.68	9.55	11.56	13.40	15.14	17.15	19.39	2.57	4.23	5.41	6.65	8.04	9.35	10.81	12.56	14.62
VAR5-P	4.02	6.23	7.78	9.45	11.12	12.50	13.82	15.04	16.12	2.66	4.35	5.38	6.26	7.09	7.84	8.54	9.25	9.74
VAR4_COF	4.23	6.95	9.67	12.68	16.12	18.64	20.87	22.98	24.59	2.85	4.84	6.76	8.77	10.96	13.12	14.71	16.53	17.59
VAR4_COF-F	3.90	6.05	7.69	9.49	11.47	13.35	15.11	17.00	18.96	2.59	4.25	5.41	6.68	7.96	9.25	10.66	12.14	14.01
VAR4_COF-FP	3.89	6.04	7.70	9.57	11.66	13.67	15.60	17.78	20.28	2.58	4.26	5.42	6.71	8.15	9.63	11.24	13.21	15.45
VAR4_COF-P	4.25	6.84	8.93	10.60	11.97	13.02	13.97	14.91	15.84	2.86	4.70	6.18	7.23	8.13	8.81	9.47	10.20	10.85

VAR4_Corn	4.35	6.88	8.83	11.03	13.21	14.79	17.09	19.20	21.67	2.87	4.81	6.21	7.78	9.10	10.64	12.21	13.69	15.52
VAR4_Corn-F	4.22	6.38	**7.62**	8.86	10.21	**11.32**	12.36	13.56	14.67	2.75	4.33	5.23	6.00	6.79	7.41	8.08	9.13	10.02
VAR4_Corn-FP	4.25	6.48	7.77	9.02	10.35	11.43	12.43	13.58	**14.66**	2.77	4.36	5.25	6.09	6.81	7.40	7.82	8.79	9.59
VAR4_Corn-P	4.39	6.78	8.17	9.50	10.86	11.97	13.11	14.18	15.03	2.90	4.63	5.58	6.44	7.41	8.15	8.92	9.61	9.98
VAR3	4.64	8.05	10.99	13.10	14.97	16.13	17.59	19.10	20.28	3.11	5.63	7.55	9.09	10.07	11.08	12.23	13.40	14.40
VAR3-FP	4.44	7.44	9.79	11.40	12.55	13.47	14.32	15.22	16.14	2.97	5.12	6.69	7.78	8.50	9.04	9.68	10.38	11.07
VAR3_FBCn	4.62	8.10	11.13	13.42	15.33	16.51	17.94	19.56	20.91	3.11	5.57	7.59	9.28	10.31	11.34	12.52	13.62	14.82
VAR3_FBCn-FP	4.49	7.59	10.06	11.67	12.68	13.42	14.12	14.97	15.90	3.01	5.12	6.80	7.74	8.44	8.95	9.41	10.03	10.75
VAR3_FBCof-FP	4.22	6.86	8.84	10.34	11.50	12.38	13.20	14.15	15.20	2.82	4.62	6.02	7.01	7.87	8.52	9.05	9.70	10.49
VEC5	3.96	5.17	8.03	10.03	12.67	14.27	16.32	17.56	18.85	2.57	3.34	5.56	7.01	8.71	9.78	11.36	12.33	13.17
VEC5-F	4.08	5.31	8.10	10.50	13.14	15.47	17.94	20.06	21.67	2.75	3.64	5.91	7.50	9.00	10.36	11.72	13.50	14.60
VEC5-FP	4.05	5.14	7.70	9.73	12.22	14.33	16.66	18.52	19.99	2.70	3.51	5.58	6.87	8.23	9.49	10.89	12.41	13.30
VEC5-P	**3.94**	5.06	**7.53**	9.31	11.66	**13.37**	15.23	16.59	**17.79**	2.56	3.25	5.24	6.38	7.93	9.01	10.50	11.67	12.43
VEC4_Corn	4.22	5.15	8.12	9.68	12.26	13.44	15.60	17.18	18.46	2.79	3.47	5.83	6.53	8.23	9.19	10.78	11.83	12.63
VEC4_Corn-FP	4.17	5.41	8.05	10.11	11.95	13.64	15.76	18.46	20.65	2.65	3.42	5.39	6.74	7.72	8.94	10.26	12.15	13.52
VEC4_Corn-F	4.21	5.51	8.21	10.40	12.33	14.21	16.39	19.19	21.41	2.68	3.49	5.47	6.90	7.96	9.30	10.72	12.72	14.10
VEC4_Corn-P	4.22	5.15	8.12	9.67	12.20	13.36	15.49	17.03	18.29	2.78	3.46	5.83	6.53	8.19	9.14	10.71	11.74	12.48
VEC4_COF	4.33	5.59	8.94	11.30	14.73	17.11	19.68	21.48	22.84	2.92	3.82	6.22	7.57	9.81	11.49	13.41	15.06	16.25
VEC4_COF-FP	4.29	5.94	9.00	11.75	15.19	18.59	21.69	24.54	26.95	2.91	4.10	6.42	8.16	10.24	12.49	14.75	17.13	18.70
VEC4_COF-F	4.20	5.37	8.63	10.92	14.41	17.28	20.26	22.70	24.42	2.78	3.64	6.11	7.45	9.68	11.94	14.34	16.24	17.31
VEC4_COF-P	4.35	5.63	8.98	11.31	14.78	17.36	20.10	22.30	24.01	2.92	3.86	6.22	7.48	9.66	11.57	13.78	15.84	17.15

Note: All models with bolded names were selected for re-estimation in rolling estimation framework. Bolded RMSE values indicate the forecast model was considered the best forecast of that specification for the respective horizon and are denoted by an italicized model name.

Table A4: Forecast accuracy based on RMSE and MAPE of rolling estimation models by forecast horizon, Jan. 2007-Dec. 2016.

	RMSE by Forecast Horizon (Months)									MAPE (%) by Forecast Horizon (Months)								
Forecast Model	H1	H2	H3	H4	H5	H6	H7	H8	H9	H1	H2	H3	H4	H5	H6	H7	H8	H9
Futures	**4.68**	6.32	**7.75**	8.99	10.06	**10.99**	11.76	12.57	**13.34**	3.03	4.20	5.05	5.88	6.58	7.13	7.72	8.51	9.22
AR(2)_S12-F	**3.77**	6.17	**8.00**	9.78	11.66	13.45	15.13	17.05	19.05	2.46	4.21	5.38	6.59	7.80	8.98	10.04	11.41	12.73
AR(2)_S6-F	3.80	6.11	7.97	9.87	11.93	13.88	15.70	17.76	19.91	2.45	4.16	5.33	6.65	7.96	9.19	10.45	11.66	13.08
AR(2)-F	3.82	6.06	8.03	10.11	12.39	14.65	16.76	19.07	21.51	2.44	4.08	5.35	6.77	8.19	9.62	11.16	12.45	14.11
AR(3)_S12	3.88	6.44	8.22	9.69	11.02	**12.20**	13.34	14.43	**15.47**	2.53	4.39	5.45	6.23	7.11	7.89	8.70	9.38	10.10
AR(3)_S12-F	3.81	6.27	8.10	9.80	11.51	13.13	14.65	16.39	18.19	2.51	4.27	5.49	6.60	7.75	8.78	9.70	11.00	12.15
AR(3)-F	3.87	6.24	8.20	10.12	12.08	13.99	15.77	17.74	19.84	2.51	4.15	5.48	6.78	8.01	9.23	10.47	11.61	12.95
VAR4_Corn	3.89	6.45	8.15	9.68	11.16	**12.50**	13.75	14.96	16.05	2.49	4.43	5.63	6.54	7.49	8.34	9.14	9.97	10.64
VAR4_Corn-F	3.79	6.27	8.10	9.88	11.71	13.46	15.09	16.87	18.62	2.43	4.29	5.55	6.70	7.90	8.89	10.02	11.43	12.68
VAR4_Corn-FP	3.84	6.32	8.14	9.86	11.64	13.32	14.91	16.70	18.52	2.46	4.31	5.55	6.66	7.82	8.80	9.76	10.99	12.26
VAR4_Corn-P	3.92	6.54	8.28	9.80	11.28	12.59	13.78	14.93	**15.95**	2.50	4.51	5.73	6.57	7.58	8.50	9.27	10.02	10.62
VAR5	3.84	6.39	8.09	9.66	11.26	12.71	14.05	15.35	16.45	2.49	4.48	5.66	6.49	7.50	8.32	9.05	9.92	10.56
VAR5-F	3.74	6.23	8.11	10.05	12.23	14.38	16.44	18.72	21.05	2.45	4.35	5.67	6.81	8.33	9.85	11.47	13.16	14.86
VAR5-FP	**3.77**	6.25	**8.03**	9.89	12.11	14.39	16.64	19.19	21.93	2.47	4.37	5.54	6.63	8.20	9.80	11.58	13.52	15.54
VAR5-P	3.88	6.48	8.15	9.67	11.23	12.62	13.87	15.06	16.03	2.51	4.50	5.67	6.46	7.42	8.22	8.73	9.38	9.85
VEC4_Corn	3.90	6.55	8.48	10.21	11.86	13.33	14.69	16.09	17.33	2.50	4.46	5.75	6.79	7.89	8.94	9.94	11.07	11.98
VEC4_Corn-FP	3.88	6.41	8.30	10.15	12.20	14.25	16.23	18.35	20.49	2.51	4.43	5.68	6.93	8.30	9.70	11.13	12.85	14.48
VEC4_Corn-P	3.94	6.64	8.58	10.33	12.01	13.51	14.90	16.31	17.54	2.56	4.52	5.86	6.89	8.07	9.13	10.15	11.21	12.08
VEC5	**3.84**	6.44	**8.31**	10.00	11.67	13.27	14.83	16.36	17.76	2.47	4.29	5.51	6.42	7.53	8.58	9.62	10.84	11.73
VEC5-FP	3.88	6.53	8.51	10.33	12.29	14.28	16.25	18.51	20.71	2.52	4.43	5.69	6.82	8.08	9.48	11.19	12.91	14.62
VEC5-P	3.86	6.50	8.32	9.92	11.52	**13.04**	14.54	16.05	**17.40**	2.47	4.29	5.52	6.33	7.45	8.50	9.45	10.66	11.52

Note: Bolded RMSE values indicate the forecast model was considered the best forecast of that specification for the respective horizon and selected for forecasting performance comparisons and forecast error analysis. The names of these selected forecasts are also in bold.