Composite Quantitative and Qualitative Forecasting of Futures Prices: Using One Commodity to Help Forecast Another

by

Anzhi Li

(Under the Direction of Jeffrey H. Dorfman)

Abstract

The thesis is composed of two chapters. The first chapter examines whether commodity price forecasting model performance can be improved by the inclusion of price forecasts for other commodities within the model specification. Using Bayesian Model Averaging methodology, we estimate 1330 different models to forecast the prices of hog, cattle, corn, and soybean and find strong support for the inclusion of one or more other commodity price forecasts in the best forecasting models.

Also, sometimes the most important forecasting component is simply whether the price will move up or down. Such binary forecasts are commonly referred to as qualitative forecasts. The second chapter investigates whether qualitative forecasting of commodity prices can be improved by the inclusion within the model specification of price forecasts for other commodities. We estimate 1330 different models to forecast the price movements of hog, cattle, corn, and soybean and find strong support for the inclusion of one or more other commodity price forecasts in the best forecasting models as well. The results for both quantitative and qualitative forecasting suggest more work is called for to determine how best to use other commodity price forecasts to improve forecasting performance.

Index words: Price Forecasting, Model Specification, Bayesian Econometrics

Composite Quantitative and Qualitative Forecasting of Futures Prices: Using One Commodity to Help Forecast Another

by

Anzhi Li

B.A., Shanghai Finance University, 2012

A Thesis Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Master of Science

Athens, Georgia

2015

Composite Quantitative and Qualitative Forecasting of Futures Prices: Using One Commodity to Help Forecast Another

by

Anzhi Li

Approved:

Major Professors: Jeffrey H. Dorfman

Committee: Berna Karali
Greg Colson

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2015

Composite Quantitative and Qualitative Forecasting of Futures Prices: Using One Commodity to Help Forecast Another

Anzhi Li

November 23, 2015

Acknowledgments

Foremost, I would like to express my deepest gratitude to my advisor Dr. Jeffrey H. Dorfman for his endless help on my thesis and related research. I am greatly thankful for his patience, motivation, and immense knowledge. It was my pleasure and honor to work with him. He is a great professor and an outstanding role model. Without him, this thesis would never have been completed.

Also, I would like to thank the rest of my thesis committee, Dr. Greg Colson and Dr. Berna Karali, for their insightful comments, encouragement, and hard questions, which motivated me to widen my research from various perspectives.

In addition, I want to take this opportunity to thank all the faculty and staff at the Department of Agricultural and Applied Economics for their kind support during my graduate study at UGA.

Last but not least, I want to give my special thanks to my parents who have always been effortlessly supporting me throughout my whole life.

Table of Contents

List of Tables

Chapter 1

Composite Quantitative Forecasting of Futures Prices: Using One Commodity to Help Forecast Another

1.1 Introduction

Commodity price forecasting has a long history in both the agricultural economics literature and in the real-world application of farm and agribusiness management. People managing businesses that involve agricultural commodities need price forecasts in order to optimally plan their actions, including the use or non-use of hedging in order to manage their output or input price risk. A selective hedging strategy incorporating information attained from the forecasts of future price movements offers increased expected utility and diminished risk, compared to strictly cash marketing. Thus, the ability to generate quality forecasts of commodity prices is important.

The question this research seeks to answer is if commodity price forecasting models can be improved by the addition of forecasts of other, related commodity prices. While structural price forecasting models have commonly included variables that relate to other commodity markets (such as cattle slaughter data being included in a hog price forecasting model), the inclusion of the price forecast itself is new and untested as far as we know. Such a method is equivalent to a hybrid structural-reduced form model as the included commodity price forecasts are essentially a composite of information deemed useful to forecasting that commodity.

We test the ability of included commodity price forecasts to improve the forecasts of other commodities using data on the four most commonly forecast commodity prices: hog, cattle, corn, and soybean. For each of these four commodities, we forecast future prices both with and without other price forecasts included in the model to examine the relative forecast performance. We do all this within a Bayesian model uncertainty framework that is well-suited to the estimation and comparison of multiple models.

This paper proceeds with a literature review section, followed by an explanation of the methodology employed. Next we describe the data and present the results. The final section presents some conclusions.

1.2 Background and Literature Review

Price volatility is a fundamental feature of agricultural markets and one of the main sources of risk in commodity markets. Futures markets play a crucial role in the pricing and distribution of commodities. For farmers, processors, food manufacturers, and other participants in commodity markets to properly manage their risks and attempt to maximize profits, commodity price forecasts are often useful. Thus, these agents are continually looking for improved forecasts, as witnessed by the long history of research on this topic. In the 1970s, the increased volatility of agricultural commodity prices gained attention from scholars to create forecasting approaches in order to serve as accurate information sources to decision makers. During the past several decades, numerous forecasting methods have been developed and evaluated for agricultural commodities, including time series models such as Autoregressive Integrated Moving Average (ARIMA) models, structural econometric models, and qualitative approaches like expert judgment.

Leuthold et al. (1970) examined the economic and mathematical characteristics of the time series data of U.S. daily hog prices by using ARIMA and structural econometric models, and then compared the developed models as to their forecasting ability based on the Theil Coefficient. They found that structural econometric models did slightly better than the ARIMA models over the evaluation period.

Additional investigation revealed that each set of forecasts contains relevant and distinct information. One model would show an overall superiority while the combined forecasts of these models would possibly outperform all the individual forecasts. In addition, the optimal combined forecasts would have an error variance not greater than the smallest error variance of the individual forecasts. Brandt and Bessler (1981) confirmed the usefulness of composite forecasting by examining the empirical accuracy of several composite forecasting techniques for quarterly U.S. hog prices based on the individual structural, ARIMA, and expert opinion methods and provided empirical evidence on the usefulness of composite forecasting, using mean squared error (MSE) as the criterion for forecasting performance. Based on their findings, individual forecasts produce large errors and they are not likely to provide the most accurate information for decision making; incorporating the prior performance of the individual forecasts, either through the minimum variance or a weighting procedure, results in lower MSE than those from simple averaging of price forecasts and it is suggested that forecast users combine the forecasts from alternative forecasting techniques to reduce the risk even if the users have no prior information of the forecasting models.

Brandt and Bessler (1983) later used seven methods, including exponential smoothing, ARIMA, a structural econometric model, expert judgement, and a composite forecasting approach, to explore forecasting performance improvement of U.S. hog prices and evaluated their forecasting performances based on MSE and mean absolute percentage error (MAPE) criteria. They found that combining forecasts from individual methods into a composite reduced the forecast error below that of any individual approach. These results are generally consistent with previous findings from other scholars (Bates and Granger, 1969; Falconer and Sivesind, 1977). Further, they found that the use of price forecasts in developing a market strategy can improve the average price received for the product. In addition, Brandt (1985) developed alternative forecasting approaches generating commodity price forecasts and noted how decision makers could reduce price variability by combining price forecasts with hedging, using an empirical example of the live hog market. These results suggest that decision makers should consider composite forecasting when planning marketing strategies.

Cromarty and Myers (1975) noted that parsimony is desirable in forecasting model selection, providing better forecasts and policy prescriptions, and good forecasting models are designed to deal explicitly with decisions of major price consequences by incorporating major policy changes, currency alignment, shifts in world demand, weather and other new information as it becomes available. This makes the Bayesian framework ideal. Brandt and Bessler (1983) also agreed with the idea of obtaining a parsimonious model that predicts out-of-sample data well, arguing that profligate models perform poorly at out-of-sample forecasting.

Dorfman (1998) later created a new Bayesian method to form composite qualitative forecasts and showed that forming composite forecasts from a set of forecasts in the Bayesian framework improved performance in an application to the hog prices. Dorfman and Sanders (2006) also introduced a systematic Bayesian approach to handle model specification uncertainty in hedging models, which can be applied to data on the hedging of corn and soybeans and on cross-hedging of corn oil using soybean oil futures.

In this paper, we are interested in investigating whether the forecasts of one commodity can help improve the forecasts of a second commodity. Hog, cattle, corn, and soybean are chosen in this paper because they are the four most common commodities that have been looked at the agricultural economics literature on forecasting. Essentially, this is a new form of composite forecasting where model specification uncertainty is taken to include the possible inclusion of the forecasts from models of other, related commodities. We demonstrate this by constructing price forecasts for each commodity (hog, cattle, corn, and soybean), with a set of models some of which include price forecasts of other commodities.

1.3 Methodology

The Basics

In this paper, we use the Bayesian approach to deal with model specification uncertainty for each commodity price forecasting model. For each commodity price to be forecast, we start with a set of possible forecasting models, estimate them all, and see which have the most posterior support from the data. This is done in two parts: the estimation of each model and the computation of each model's support.

Given a model j, for one commodity price, assume a linear regression model:

$$y = X_j \beta_j + \epsilon_j, \ j = 1, \ldots, M, \tag{1.1}$$

where y is the vector of observations on the dependent variable assumed identical in all models, X_j is the matrix of the independent variables for the j^{th} model considered, ϵ_j is the vector of random errors for the j^{th} model, and j denotes the model in the set of M models considered. The dependent variable here is assumed to be identical in all models, and therefore the differences between the models are restricted to the matrix X of independent variables.

The prior distributions on the regression parameters β_j can be specified as

$$p(\beta_j) \sim N(b_{0j}, \sigma_j^2 V_{0j}), \ j = 1, \ldots, M, \tag{1.2}$$

where N represents the multivariate normal distribution, b_{0j} is the prior mean of the regression parameters for the j^{th} model, and $\sigma_j^2 V_{0j}$ is the prior covariance matrix. The prior distribution on σ_j^2 is specified as an inverse-gamma, which is equivalent to a gamma distribution on σ_j^{-2},

$$p(\sigma_j^{-2}) \sim G(s_{0j}^{-2}, d_{0j}),\ j = 1, \ldots, M, \tag{1.3}$$

where G stands for the gamma distribution, s_{0j}^{-2} is the prior mean for the inverse error variance, and d_{0j} is the prior degrees of freedom. A higher value of d_{0j} indicates a more informative prior (Koop, 2003).

The likelihood function for each model can be specified as

$$L_j(y|\beta_j, \sigma_j^2, X_j) = (2\pi\sigma_j^2)^{-n/2} exp\{-0.5(y - X_j\beta_j)' \sigma_j^{-2}(y - X_j\beta_j)\},\ j = 1, \ldots, M, \tag{1.4}$$

which is assumed to follow a standard form based on identically and normally distributed random error terms ϵ_j.

Given these priors and the likelihood function above, the joint posterior distribution of β_j and σ_j^2 can be derived by Bayes' Theorem that the posterior distribution is proportional to the prior distribution times the likelihood function. The joint posterior can be written as

$$p(\beta_j, \sigma_j^2|y, X_j) \sim NG(b_{pj}, V_{pj}, s_{pj}^2, d_{pj}),\ j = 1, \ldots, M, \tag{1.5}$$

where

$$V_{pj} = (V_{0j}^{-1} + X_j' X_j)^{-1}, \tag{1.6}$$

$$b_{pj} = V_{pj}(V_{0j}^{-1} b_{0j} + (X_j' X_j)\hat{\beta}_j), \tag{1.7}$$

$$d_{pj} = d_{0j} + n_j, \tag{1.8}$$

and

$$s_{pj}^2 = d_{pj}^{-1}[d_{0j} s_{0j}^2 + (n_j - k_j)s_j^2 + (\hat{\beta}_j - b_{0j})' (V_{0j} + (X_j' X_j)^{-1})^{-1} (\hat{\beta}_j - b_{0j})], \tag{1.9}$$

where NG represents the joint normal-gamma distribution, $\hat{\beta}_j$ and s_j^2 are the standard OLS quantities and n_j and k_j are the rows and columns of X_j, respectively. Equations (6) to

(9) together define the parameters in the distribution. $s_{pj}^2 V_{pj}$ is the posterior mean of the variance, b_{pj} is the posterior mean of the coefficients, which are the weighted averages of the means of the prior distribution and the parameters that are derived from the maximum likelihood estimator based on the data, and d_{pj} is the posterior degrees of freedom.

Model Specification Uncertainty

Now we describe the process for handling model specification uncertainty. First, a discrete prior weight is assigned to each model

$$p(M_j) = \mu_j, \ \sum_{j=1}^{M} \mu_j = 1. \tag{1.10}$$

Here we choose to use uninformative priors across the model specification, so all models are treated equally. In this case, $\mu_j = 1/M, \forall j$. Then, using the above results for the posterior distributions shown in (5), we derive the marginal likelihood functions by integrating out the parameter uncertainty to leave

$$p(y|M_j) = c_j[|V_{pj}|/|V_{0j}|]^{1/2}(d_{pj}s_{pj}^2)^{-d_{pj}/2}, \tag{1.11}$$

where

$$c_j = \frac{\Gamma(d_{pj}/2)(d_{0j}s_{0j}^2)^{d_{0j}/2}}{\Gamma(d_{0j}/2)\pi^{n/2}}, \tag{1.12}$$

and $\Gamma(\cdot)$ is the Gamma function. The marginal likelihood tells how well the model fits on average, where the averaging is over all possible parameter values. As shown in equation (11), the smaller the posterior mean of the variance is, the larger the marginal likelihood will be, indicating that the better the model fits, the larger the marginal likelihood will be. Combining (11) and (12) by Bayes' Theorem, the posterior probability of each model can be derived as follows:

$$p(M_j|y) \propto \mu_j[|V_{pj}|/|V_{0j}|]^{1/2}(d_{pj}s^2_{pj})^{-d_{pj}/2} = \mu_j p(y|M_j),\ j = 1, \ldots, M. \tag{1.13}$$

Normalizing the values in (13) by dividing each value by the sum of the unnormalized posterior probabilities across all M models will make sure that these posterior model probabilities sum to unity. Denote these normalized posterior probabilities by

$$\omega_j = \frac{\mu_j p(y|M_j)}{\sum_{j=1}^{M} \mu_j p(y|M_j)},\ j = 1, \ldots, M. \tag{1.14}$$

These posterior probabilities ω_j are the key to evaluating both general model specification uncertainty and the advantage of including forecasts of other commodity prices in the forecasting model. Models which receive higher posterior probabilities are better supported by the data, indicating that those models are preferred choices and can be expected to yield better forecasting performance. We further obtain the posterior support for model traits by summing the posterior probabilities of each variable across models, thereby determining ideal model specification. The higher the summed posterior probability of one variable, the more support that variable has for being included in the model specification.

1.4 Data

Data on the four commodity prices are collected from the Chicago Mercantile Exchange (CME) Group, using monthly futures prices for lean hog futures ($/lb), live cattle futures ($/lb), corn futures ($/bushel), and soybean futures ($/bushel).

Possible independent variables, including autoregressive (AR) processes and exogenous variables, are selected based on analyses of previous studies in the literature. For the hog price forecasting models, the AR terms to be considered range from AR(3) to AR(12) and the exogenous variables include monthly disposable personal income (logged), monthly commercial cattle slaughter (thousand heads), monthly broiler-type poultry eggs hatched (million eggs), monthly number of sows farrowing (thousand heads), and monthly pork cold storage (million pounds). For the cattle price forecasting models, the independent variables considered are the same as the hog price forecasting model except pork storage is not included. In the corn price forecasting models, the AR terms to be considered range from AR(3) to AR(6) and the exogenous variables to be included are monthly corn export (thousand units), monthly corn inventory (million bushels), monthly lagged acres planted to corn (thousand acres), and monthly fuel ethanol production (million gallons). For the soybean price forecasting models, the independent variables considered are the same as in the corn model except the ethanol variable is not included. All the data of the exogenous variables for the four commodity price forecasting models come from the National Agricultural Statistics Service (NASS). Additionally, the monthly lagged acres planted to corn/soybean is a weighted average of lagged acres of corn/soybean based on the Palmer Drought Severity Index (PDSI) for each state. The data on the PDSI are provided by the National Oceanic and Atmospheric Administration (NOAA).

All data are monthly extending from January 1981 to December 2013. We use the first twenty-six years (January 1981-December 2006) for in-sample estimation, and then evaluate out-of-sample forecasting performance over the period from January 2010 to December 2012, 36 observations. Due to the high volatility of these four commodity prices over the period from 2007 to 2009, the out-of-sample estimation of that period has not been considered here. Additionally, in order to check the consistency of the forecasting performance of the models, we further examine the forecasting performance over the period from January 2013 to December 2013, 12 observations.

Table 1.1 shows the set of variables considered in the model specification and the total number of forecasting models estimated for each of the four commodity prices. In the hog price forecasting model, the hog price (PH) to be forecast is the monthly lean hog futures price ($/lb) as given by CME group. Among the exogenous variables considered for the hog price forecasting model, DSPI denotes the natural logarithm of monthly disposable personal income (billion dollars); CTSL denotes the monthly commercial cattle slaughter (thousand heads); HATCH denotes the monthly broiler-type poultry eggs hatched (million eggs); SF denotes the monthly number of sows farrowing (thousand heads); PKST denotes the monthly pork cold storage (million pounds). In the cattle price forecasting model, the cattle price (PCA) to be forecast is the monthly live cattle futures price ($/lb) as given by CME group. The independent variables considered are basically the same as in the hog model except the PKST variable. In the corn price forecasting model, the corn price (PC) to be forecast is the monthly corn futures price ($/bushel) as given by CME group. Among the exogenous variables considered for the corn price forecasting model, EXPORT^c denotes the monthly corn export (thousand units); INVENTORY^c denotes the monthly corn inventory (million bushels); ACRES^c denotes the monthly lagged acreages planted for corn (thousand acres); ETHANOL denotes the monthly fuel ethanol production (million gallons). In the soybean price forecasting model, the soybean price (PS) to be forecast is the monthly soybean futures price ($/bushel) as given by CME group. The independent variables considered are the same as in the corn model except the ETHANOL variable. The data of the exogenous variables for the four commodity price forecasting models are provided by NASS. In addition, $\text{ACRES}^c/\text{ACRES}^s$ is a weighted average lagged acres of corn/soybean based on the Palmer Drought Severity Index (PDSI) given by NOAA for each of the 48 states.

1.5 Empirical Results

Beginning with the hog price forecasting models, Table 1.2 presents the posterior probabilities for the model specification. The probabilities shown in Table 1.2 are the probability that each of the variables listed belongs in the true model. These probabilities show that there is clear and overwhelming support for the inclusion of AR(3) (0.992), disposable personal income (1.000), egg hatching (0.977), sows farrowing (1.000), and pork storage (0.999) in the hog price forecasting model. Also, cattle forecasts have a 0.878 posterior probability of inclusion. Other variables have little to no posterior support for inclusion in the hog price forecasting model.

In terms of forecasting performance, Table 1.3 presents the out-of-sample mean squared error (MSE) over the two periods of time for hog price for the five best and five worst performing forecasting models. Note that the five best and five worst performing forecasting models are decided based on MSE over the 2010-2012 period. Table 1.4 displays the MSEs of the five most probable and five least probable models; these are the models with the highest and lowest posterior model probabilities. The five most probable models are those that one would be most likely to choose ex ante before seeing out-of-sample forecasting performance. As shown in Tables 1.3 and 1.4, the five most probable and five best performing models all have disposable personal income, sows farrowing and pork storage as the exogenous variables and include either one or more commodity forecasts. In Table 1.4, for the 2010-2012 period, two of the five most probable models have excellent forecasting performance, as measured by MSE, that is close to the best hog price forecasting performance models in Table 1.3, and they all have smaller MSEs (better forecasting performance) than the mean and median level of the total 420 hog price forecasting models; for the 2013 period, the models with the best forecasting performance over the 2010-2012 period no longer have top performance and the most along with fourth most probable model actually beats all of the five best performance models on MSE. In addition, the composite forecasts computed based on the 420 hog models over the 2010-2012 period perform better than four of the five most probable models as well as the mean and median level, indicating that our Bayesian methodology works for the hog price forecasting model.

Moving to the cattle price forecasting models, Table 1.5 presents the posterior probabilities in favor of variable inclusion in the cattle forecasting model. These results show that disposable personal income (0.998), cattle slaughter (0.965), sows farrowing (0.990), and hog price forecasts (0.998) have enormous support for inclusion in the cattle price forecasting model. Other variables have little to no posterior support for inclusion in the cattle price forecasting model.

Regarding forecasting performance, Tables 1.6 and 1.7 hold the MSEs of the best/worst performing models and the most/least probable models, respectively. Table 1.7 shows that none of the most probable cattle price forecasting models perform close to the best forecasting performance models, based on their MSEs over the 2010-2012 period. Also, none of these models outperform the mean and median level of the total 350 cattle price forecasting models and even the composite forecasts. During the 2013 period, the models that have the best forecasting performance over the 2010-2012 period are no longer the best performing models. Beyond that, we find that the MSE of the composite forecasts computed from the 350 cattle models, unfortunately, happens to be larger than the most probable model for both out-of sample forecast periods (2010-2012 period: $28.455 > 27.415$; 2013 period: $18.085 > 14.315$), indicating that the composite forecasts in the cattle price forecasting case actually perform worse than the most probable model.

Next for the corn price forecasting models, Table 1.8 presents the posterior probabilities in favor of variable inclusion in the corn price forecasting model. AR(3) has a 0.999 posterior probability of inclusion, ethanol production has a 0.864 probability. No other variables have posterior support that reaches 0.20, so the model specification is also quite clear.

Table 1.9 shows the MSEs of the five best and five worst performing corn price forecasting models. Table 1.10 displays the MSEs of the five most and five least probable corn price forecasting models. Based on the MSEs over the 2010-2012 period, though four of the five most probable models have better forecasting performance than the mean and median level of the total 308 corn price forecasting models, they are still noticeably worse than the best corn price forecasting performance models in Table 1.9. As in the hog and cattle price forecasting cases, the models with best corn price forecasting performance over the 2010-2012 period no longer have top performance over the 2013 period while four of the worst five performing models over the 2010-2012 period still have worst forecasting performance over the 2013 period. Moreover, all the five most probable models actually outperform all the best performing models on MSE over the 2013 period. Additionally, the composite forecasts computed based on the total 308 corn models outperform both the mean and median level for both out-of-sample periods, and more importantly, the composite forecasts beat the two most probable models for the 2013 out-of-sample period, indicating that our Bayesian methodology also has promise for the corn price forecasting model.

Finally, the soybean price forecasting model specification results are in Table 1.11. The posterior probabilities show strong support for including AR(3) (0.999) and soybean export (0.970) in the soybean price forecasting model. Also, hog forecasts have a 0.607 posterior probability of inclusion. Table 1.12 presents the MSEs for the five best and five worst performing forecasting models, while Table 1.13 displays the MSEs for the five most and five least probable models. For the 2010-2012 period, based on the MSE criterion, while two of the five most probable soybean price forecasting models have better forecasting performance than the mean and median level of the total 252 soybean models, none is as good as the best performing models in Table 1.12. For the 2013 period, the models with best soybean price forecasting performance over the 2010-2012 period no longer have top performance over the 2013 period while the fifth most probable model has the tenth best forecasting performance. Similarly, for the 2013 out-of-sample period, the composite forecasts computed based on the total 252 soybean models outperform both the mean and median level, and more importantly, the composite forecasts also beat three of the most probable models including the most probable one. This suggests that our Bayesian methodology also works for the soybean price forecasting model.

Overall, of those twenty top probable models for the four commodities, nine have above average forecasting performance. Also we find that within the lists of the five best forecasting models for each of the four commodity prices, models that include commodity price forecasts are heavily represented. Of those twenty best performing models, seventeen include one or more commodity price forecasts. This suggests that it is worth pursuing how commodity price forecasts can be improved by the inclusion of other commodity price forecasts in the forecasting models.

Furthermore, in terms of out-of-sample forecasts over the 2013 period, the best performing models over the 2010-2012 period still perform relatively well, though falling short of the best performing models. The worst performing models over the 2010-2012 period generally still stay at the bottom over the 2013 period.

1.6 Conclusions

The Bayesian Model Averaging methodology applied here for model specification to the forecasting of four important commodity prices provides clear signals for variable inclusion in the forecasting models, although the results of the Bayesian Model Averaging are somewhat mixed with regard to signaling which models are likely to have the best out-of-sample forecasting performance. Based on our findings, in general, the models with the highest model probabilities based on the in-sample data deliver around average out-of-sample forecasting performance. For price forecasting of hog, corn, and soybean, the composite forecasts computed under the Bayesian framework outperform the most probable model among the entire set of models estimated but that is not the case for cattle price forecasting. Also, the fact that seventeen of the twenty best performing forecasting model, as measured by out-of-sample MSE, contain price forecasts for one or more different commodities suggests that the idea of improving commodity price forecasting by including the composite forecasts of other commodities in the model is a good one. Still, additional work is needed to evaluate multiple models based on the out-of-sample forecasting performance so that users of such forecasts can have some scientific basis for choosing a model specification (including possibly using a composite forecast). We believe the results here show that we are on the right track, but have not yet arrived at out desired destination.

Table 1.1: Variables Used to Predict Commodity Prices

Dependent Variable	Lags	Exogenous Variables
PH (cents per pound) (420 models)	AR(3) $\sim$ AR(12)	DSPI; CTSL; HATCH; SF ;PKST
PCA (cents per pound) (350 models)	AR(3) $\sim$ AR(12)	DSPI; CTSL; HATCH; SF
PC (cents per bushel) (308 models)	AR(3) $\sim$ AR(6)	EXPORT^c; INVENTORY^c; ACRES^c; ETHANOL
PS (10 cents per bushel) (252 models)	AR(3) $\sim$ AR(6)	EXPORT^s; INVENTORY^s; ACRES^s

Table 1.2: Hog Price Forecasting Model Specification (420 Models)

Model Traits	Post Probability
Include AR(3)	0.992
Include DSPI	1.000
Include CTSL	0.024
Include HATCH	0.978
Include SF	1.000
Include PKST	0.999
Include Cattle Forecasts	0.878
Include Corn Forecasts	0.070
Include Soybean Forecasts	0.054
No Forecasts	<0.001

Table 1.3: Top 5 and Bottom 5 Hog Price Forecasting Models by MSE

Top 5 Models by 2010-2012 MSE	2010-2012 MSE	2013 MSE	Post Probability
1) AR(6)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Cattle Forecasts$_t$ +Soybean Forecasts$_t$	39.792	25.527	<0.001
2) AR(6)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Soybean Forecasts$_t$	39.823	27.066	<0.001
3) AR(4)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Soybean Forecasts$_t$	39.835	27.940	<0.001
4) AR(4)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Cattle Forecasts$_t$ +Soybean Forecasts$_t$	39.854	27.290	<0.001
5) AR(7)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Cattle Forecasts$_t$ +Soybean Forecasts$_t$	39.876	26.744	<0.001
Bottom 5 Models by 2010-2012 MSE			
1) AR(10)+DSPI$_t$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+PKST$_{t-1}$+Cattle Forecasts$_t$ +Corn Forecasts$_t$	62.277	45.697	<0.001
2) AR(11)+DSPI$_t$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+PKST$_{t-1}$+Cattle Forecasts$_t$ +Corn Forecasts$_t$	61.902	44.800	<0.001
3) AR(10)+DSPI$_t$+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +PKST$_{t-1}$+Corn Forecasts$_t$	61.540	44.039	<0.001
4) AR(9)+DSPI$_t$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+PKST$_{t-1}$+Cattle Forecasts$_t$ +Corn Forecasts$_t$	61.517	46.756	<0.001
5) AR(10)+DSPI$_t$+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +PKST$_{t-1}$+Corn Forecasts$_t$ +Soybean Forecasts$_t$	61.479	45.140	<0.001
Mean MSE	47.277	30.099	
Median MSE	44.912	28.178	
Composite Forecasts MSE	44.213	25.618	

Table 1.4: Top 5 and Bottom 5 Hog Price Forecasting Models by Posterior Probability

5 Most Probable Models	Post Probability	2010-2012 MSE	2013 MSE
1) AR(3)+DSPI$_t$+HATCH$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Cattle Forecasts$_t$	0.856	44.217	24.814
2) AR(3)+DSPI$_t$+HATCH$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Corn Forecasts$_t$	0.060	56.500	39.099
3) AR(3)+DSPI$_t$+HATCH$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Soybean Forecasts$_t$	0.049	49.757	36.469
4) AR(3)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Cattle Forecasts$_t$	0.011	46.420	24.908
5) AR(3)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Corn Forecasts$_t$	0.007	42.576	28.985
5 Least Probable Models			
1) AR(12)+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+PKST$_{t-1}$+Corn Forecasts$_t$ +Soybean Forecasts$_t$	$<$0.001	46.236	25.965
2) AR(12)+DSPI$_t$+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+PKST$_{t-1}$+Corn Forecasts$_t$ +Soybean Forecasts$_t$	$<$0.001	53.112	37.736
3) AR(12)+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+PKST$_{t-1}$+Cattle Forecasts$_t$ +Soybean Forecasts$_t$	$<$0.001	42.535	21.400
4) AR(11)+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+PKST$_{t-1}$+Corn Forecasts$_t$ +Soybean Forecasts$_t$	$<$0.001	46.343	28.700
5) AR(12)+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+PKST$_{t-1}$+Cattle Forecasts$_t$ +Corn Forecasts$_t$	$<$0.001	46.769	24.538
Mean MSE		47.277	30.099
Median MSE		44.912	28.178
Composite Forecasts MSE		44.213	25.618

Table 1.5: Cattle Price Forecasting Model Specification (350 Models)

Model Traits	Post Probability
Include AR(3)	0.321
Include AR(6)	0.451
Include DSPI	0.998
Include CTSL	0.965
Include HATCH	0.047
Include SF	0.990
Include Hog Forecasts	0.998
Include Corn Forecasts	0.002
Include Soybean Forecasts	0.001
No Forecasts	<0.001

Table 1.6: Top 5 and Bottom 5 Cattle Price Forecasting Models by MSE

Top 5 Models by 2010-2012 MSE	2010-2012 MSE	2013 MSE	Post Probability
1) AR(5)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$ +Hog Forecasts$_t$+Soybean Forecasts$_t$	19.907	12.447	<0.001
2) AR(6)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$ +Soybean Forecasts$_t$	19.962	11.385	<0.001
3) AR(3)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$+Soybean Forecasts$_t$	19.982	14.544	<0.001
4) AR(6)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$ +Hog Forecasts$_t$+Soybean Forecasts$_t$	19.983	11.415	<0.001
5) AR(5)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$ +Soybean Forecasts$_t$	19.995	12.255	<0.001
Bottom 5 Models by 2010-2012 MSE			
1) AR(3)+DSPI$_t$+HATCH$_{t-1}$+SF$_{t-1,t-2}$	36.488	38.724	<0.001
2) AR(3)+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$+Corn Forecasts$_t$	35.468	28.446	<0.001
3) AR(3)+DSPI$_t$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	35.254	37.029	<0.001
4) AR(4)+DSPI$_t$+HATCH$_{t-1}$SF$_{t-1,t-2}$ +Hog Forecasts$_t$	34.247	31.406	<0.001
5) AR(3)+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Corn Forecasts$_t$+Soybean Forecasts$_t$	33.766	23.440	<0.001
Mean	24.872	13.376	
Median	24.732	11.760	
Composite Forecasts	28.455	18.085	

Table 1.7: Top 5 and Bottom 5 Cattle Price Forecasting Models by Posterior Probability

5 Most Probable Models	Post Probability	2010-2012 MSE	2013 MSE
1) AR(6)+DSPI$_t$+CTSL$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	0.424	27.415	14.315
2) AR(3)+DSPI$_t$+CTSL$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	0.305	32.309	27.646
3) AR(5)+DSPI$_t$+CTSL$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	0.161	27.702	16.912
4) AR(4)+DSPI$_t$+CTSL$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	0.058	30.360	21.634
5) AR(6)+DSPI$_t$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	0.025	28.226	16.625
5 Least Probable Models			
1) AR(12)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Corn Forecasts$_t$+Soybean Forecasts$_t$	<0.001	21.768	9.890
2) AR(12)+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Corn Forecasts$_t$+Soybean Forecasts$_t$	<0.001	28.815	9.057
3) AR(11)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Corn Forecasts$_t$+Soybean Forecasts$_t$	<0.001	21.508	9.937
4) AR(12)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$+Soybean Forecasts$_t$	<0.001	21.205	11.170
5) AR(11)+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Corn Forecasts$_t$+Soybean Forecasts$_t$	<0.001	28.465	9.385
Mean MSE		24.872	13.376
Median MSE		24.732	11.760
Composite Forecasts MSE		28.455	18.085

Table 1.8: Corn Price Forecasting Model Specification (308 Models)

Model Traits	Post Probability
Include EXPORT$_t^c$	<0.001
Include INVENTORY$_t^c$	0.125
Include ETHANOL$_t$	0.864
Include ACRES$_t^c$	0.011
Include AR(3)	0.999
Include Hog Forecasts	0.009
Include Cattle Forecasts	0.098
Include Soybean Forecasts	<0.001
No Forecasts	0.894

Table 1.9: Top 5 and Bottom 5 Corn Price Forecasting Models by MSE

Top 5 Models by 2010-2012 MSE	2010-2012 MSE	2013 MSE	Post Probability
1) AR(3)+ETHANOL$_t$+INVENTORY$_t^c$ +Hog Forecasts$_t$+Cattle Forecasts$_t$	4145.528	2646.448	<0.001
2) AR(3)+ETHANOL$_t$+INVENTORY$_t^c$ +Hog Forecasts$_t$+Soybean Forecasts$_t$	4148.722	2646.386	<0.001
3) AR(4)+ETHANOL$_t$+INVENTORY$_t^c$ +Hog Forecasts$_t$+Cattle Forecasts$_t$	4148.981	2665.617	<0.001
4) AR(3)+ETHANOL$_t$+INVENTORY$_t^c$ +Hog Forecasts$_t$	4150.145	2637.776	<0.001
5) AR(4)+ETHANOL$_t$+INVENTORY$_t^c$ +Hog Forecasts$_t$+Soybean Forecasts$_t$	4151.836	2664.314	<0.001
Bottom 5 Models by 2010-2012 MSE			
1) AR(5)+ACRES$_{t-1}^c$+Hog Forecasts$_t$	5818.718	2827.018	<0.001
2) AR(5)+ACRES$_t^c$+Hog Forecasts$_t$	5816.814	2823.805	<0.001
3) AR(4)+ACRES$_{t-1}^c$+Hog Forecasts$_t$	5770.938	2988.628	<0.001
4) AR(4)+ACRES$_t^c$+Hog Forecasts$_t$	5768.744	2985.898	<0.001
5) AR(3)+ACRES$_t^c$+Hog Forecasts$_t$	5764.925	2962.363	<0.001
Mean	4864.161	2604.911	
Median	4850.833	2595.945	
Composite Forecasts	4623.618	2573.111	

Table 1.10: Top 5 and Bottom 5 Corn Price Forecasting Models by Posterior Probability

5 Most Probable Models	Post Probability	2010-2012 MSE	2013 MSE
1) AR(3)+ETHANOL$_t$	0.396	4572.955	2596.415
2) AR(3)+ETHANOL$_{t-1}$	0.373	4602.819	2583.324
3) AR(3)+INVENTORY$_t^c$	0.089	4894.200	2461.814
4) AR(3)+ETHANOL$_t$+Cattle Forecasts$_t$	0.044	4530.603	2619.124
5) AR(3)+ETHANOL$_{t-1}$+Cattle Forecasts$_t$	0.041	4574.050	2599.164
5 Least Probable Models			
1) AR(6)+ACRES$_t^c$+INVENTORY$_t^c$ +Cattle Forecasts$_t$+Soybean Forecasts$_t$	<0.001	4695.957	2440.241
2) AR(6)+ACRES$_t^c$+INVENTORY$_t^c$ +Hog Forecasts$_t$+Soybean Forecasts$_t$	<0.001	5167.274	2458.448
3) AR(6)+ACRES$_t^c$+ETHANOL$_t$ +Cattle Forecasts$_t$+Soybean Forecasts$_t$	<0.001	4549.190	2632.452
4)AR(6)+ACRES$_t^c$+INVENTORY$_t^c$ +Soybean Forecasts$_t$	<0.001	4913.645	2425.709
5) AR(6)+ACRES$_t^c$+INVENTORY$_t^c$ +Hog Forecasts$_t$+Cattle Forecasts$_t$	<0.001	4826.405	2396.334
Mean MSE		4864.161	2604.911
Median MSE		4850.833	2595.945
Composite Forecasts MSE		4623.618	2573.111

Table 1.11: Soybean Price Forecasting Model Specification (252 Models)

Model Traits	Post Probability
Include AR(3)	0.999
Include EXPORT^s	0.970
Include INVENTORY^s	0.028
Include ACRES^s	0.002
Include Hog Forecasts	0.607
Include Cattle Forecasts	0.367
Include Corn Forecasts	0.029
No Forecasts	<0.001

Table 1.12: Top 5 and Bottom 5 Soybean Price Forecasting Models by MSE

Top 5 Models by 2010-2012 MSE	2010-2012 MSE	2013 MSE	Post Probability
1) AR(5)+EXPORT_t^s+INVENTORY_{t-1}^s +Corn Forecasts_t	116.602	71.624	<0.001
2) AR(6)+EXPORT_t^s+INVENTORY_{t-1}^s +Corn Forecasts_t	117.063	74.198	<0.001
3) AR(5)+EXPORT_t^s+ACRES_t^s +Corn Forecasts_t	117.283	73.125	<0.001
4) AR(5)+EXPORT_t^s+ACRES_t^s +Cattle Forecasts_t+Corn Forecasts_t	117.322	72.953	<0.001
5) AR(5)+EXPORT_t^s+INVENTORY_{t-1}^s +Cattle Forecasts_t+Corn Forecasts_t	117.338	72.512	<0.001
Bottom 5 Models by 2010-2012 MSE			
1) AR(3)+ACRES_{t-1}^s+Hog Forecasts_t	166.272	131.824	<0.001
2) AR(5)+ACRES_{t-1}^s+Hog Forecasts_t	163.178	124.714	<0.001
3) AR(4)+ACRES_{t-1}^s+Hog Forecasts_t	163.159	123.488	<0.001
4) AR(6)+ACRES_{t-1}^s+Hog Forecasts_t	162.600	122.835	<0.001
5) AR(3)+ACRES_t^s+Hog Forecasts_t	162.170	125.890	<0.001
Mean MSE	134.360	91.467	
Median MSE	133.350	88.695	
Composite Forecasts MSE	134.800	88.241	

Table 1.13: Top 5 and Bottom 5 Soybean Price Forecasting Models by Posterior Probability

5 Most Probable Models	Post Probability	2010-2012 MSE	2013 MSE
1) AR(3)+EXPORT_t^s+Hog Forecasts_t	0.295	135.410	88.698
2) AR(4)+EXPORT_{t-1}^s+Hog Forecasts_t	0.292	140.137	91.897
3) AR(3)+EXPORT_t^s+Cattle Forecasts_t	0.204	131.148	86.851
4) AR(3)+EXPORT_{t-1}^s+Cattle Forecasts_t	0.149	133.718	87.734
5) AR(3)+EXPORT_t^s+Corn Forecasts_t	0.015	119.043	70.727
5 Least Probable Models			
1) AR(6)+INVENTORY_t^s+ACRES_{t-1}^s	<0.001	137.605	97.721
2) AR(6)+EXPORT_t^s+ACRES_t^s	<0.001	125.760	86.537
3) AR(6)+INVENTORY_t^s+ACRES_{t-1}^s +Cattle Forecasts_t+Corn Forecasts_t	<0.001	122.478	82.189
4) AR(6)+INVENTORY_t^s+ACRES_{t-1}^s +Hog Forecasts_t+Corn Forecasts_t	<0.001	134.750	95.451
5) AR(5)+INVENTORY_t^s+ACRES_{t-1}^s	<0.001	139.836	99.767
Mean		134.360	91.467
Median		133.350	88.695
Composite Forecasts		134.800	88.241

Chapter 2

Composite Qualitative Forecasting of Futures Prices: Using One Commodity to Help Forecast Another

2.1 Introduction

Commodity price forecasting has a long history in both the agricultural economics literature and in the real-world application of farm and agribusiness management. People managing businesses that involve agricultural commodities need price forecasts in order to optimally plan their actions, including the use or non-use of hedging in order to manage their output or input price risk. A selective hedging strategy incorporating information attained from the forecasts of future price movements offers increased expected utility and diminished risk, compared to strictly cash marketing. Thus, the ability to generate quality forecasts of commodity prices is important.

The question this research seeks to answer is if commodity price forecasting models can be improved by the addition of forecasts of other, related commodity prices. While structural price forecasting models have commonly included variables that relate to other commodity markets (such as cattle slaughter data being included in a hog price forecasting model), the inclusion of the price forecast itself is new and untested as far as we know. Such a method is equivalent to a hybrid structural-reduced form model as the included commodity price forecasts are essentially a composite of information deemed useful to forecasting that commodity.

Because in many situations, the key part of a price forecast is whether the price will move up or down in the future, we focus here on qualitative forecasts of the direction of price changes. We test the ability of included commodity price forecasts to improve the qualitative forecasts of other commodities using data on the four most commonly forecast commodity prices: hog, cattle, corn, and soybean. For each of these four commodities, we forecast future prices both with and without other price forecasts included in the model to examine the relative forecast performance. We do all this within a Bayesian model uncertainty framework that is well-suited to the estimation and comparison of multiple models.

The paper proceeds with a literature review section, followed by an explanation of the methodology employed. Next we describe the data and present the results. The final section presents some conclusions.

2.2 Background and Literature Review

Price volatility is a fundamental feature of agricultural markets and one of the main sources of risk in commodity markets. Futures markets play a crucial role in the pricing and distribution of commodities. For farmers, processors, food manufacturers, and other participants in commodity markets to properly manage their risks and attempt to maximize profits, commodity price forecasts are often useful. Thus, these agents are continually looking for improved forecasts, as witnessed by the long history of research on this topic. In the 1970s, the increased volatility of agricultural commodity prices focused the attention of scholars on creating forecasting approaches in order to serve as accurate information sources for decision makers. During the past several decades, numerous forecasting methods have been developed and evaluated for agricultural commodities, including time series models such as Autoregressive Integrated Moving Average (ARIMA) models, structural econometric models, and qualitative approaches like expert judgment. Leuthold et al. (1970) examined the economic and mathematical characteristics of the time series data of U.S. daily hog prices by using ARIMA and structural econometric models, and then compared the developed models as to their forecasting ability based on the Theil Coefficient. They found that structural econometric models did slightly better than the ARIMA models over the evaluation period.

Additional investigation revealed that each set of forecasts contains relevant and distinct information. One model would show an overall superiority while the combined forecasts of these models would possibly outperform all the individual forecasts. In addition, the optimal combined forecasts would have an error variance not greater than the smallest error variance of the individual forecasts. Brandt and Bessler (1981) confirmed the usefulness of composite forecasting by examining the empirical accuracy of several composite forecasting techniques for quarterly U.S. hog prices based on the individual structural, ARIMA, and expert opinion methods and provided empirical evidence on the usefulness of composite forecasting, using mean squared error (MSE) as the criterion for forecasting performance. Based on their findings, individual forecasts produce large errors and they are not likely to provide the most accurate information for decision making; incorporating the prior performance of the individual forecasts, either through the minimum variance or a weighting procedure, results in lower MSE than those from simple averaging of price forecasts and it is suggested that forecast users combine the forecasts from alternative forecasting techniques to reduce the risk even if the users have no prior information of the forecasting models.

Brandt and Bessler (1983) later used seven methods, including exponential smoothing, ARIMA, a structural econometric model, expert judgement, and a composite forecasting approach, to explore forecasting performance improvement of U.S. hog prices and evaluated their forecasting performances based on MSE and mean absolute percentage error (MAPE) criteria. They found that combining forecasts from individual methods into a composite reduced the forecast error below that of any individual approach. These results are generally consistent with previous findings from other scholars (Bates and Granger, 1969; Ealconer and Sivesind, 1977). Further, they found that the use of price forecasts in developing a market strategy can improve the average price received for the product. In addition, Brandt (1985) developed alternative forecasting approaches generating commodity price forecasts and noted how decision makers could reduce price variability by combining price forecasts with hedging, using an empirical example of the live hog market. These results suggest that decision makers should consider composite forecasting when planning marketing strategies. Feather and Kaylen (1989) suggested a procedure for the formation of a conditional "composite" qualitative forecast, the theoretical development of which was followed by an empirical application using quarterly hog prices. The results showed the composite allows the possibility of avoiding reliance on an inferior forecasting method.

Cromarty and Myers (1975) noted that parsimony is desirable in forecasting model selection, providing better forecasts and policy prescriptions, and good forecasting models are designed to deal explicitly with decisions of major price consequences by incorporating major policy changes, currency alignment, shifts in world demand, weather and other new information as it becomes available. This makes the Bayesian framework ideal. Brandt and Bessler (1983) also agreed with the idea of obtaining a parsimonious model that predicts out-of-sample data well, arguing that profligate models perform poorly at out-of-sample forecasting.

Dorfman (1998) later created a new Bayesian method to form composite qualitative forecasts and showed that forming composite forecasts from a set of forecasts in the Bayesian framework improved performance in an application to the hog prices. Dorfman and Sanders (2006) also introduced a systematic Bayesian approach to handle model specification uncertainty in hedging models, which can be applied to data on the hedging of corn and soybeans and on cross-hedging of corn oil using soybean oil futures.

In this paper, we are interested in investigating whether the forecasts of one commodity can help improve the forecasts of a second commodity. Hog, cattle, corn, and soybean are chosen in this paper because they are the four most common commodities that have been looked at the agricultural economics literature on forecasting. Essentially, we propose a new form of composite forecasting where model specification uncertainty is taken to include the possible inclusion of the forecasts from models of other, related commodities. We demonstrate this by constructing qualitative price forecasts for each commodity (hog, cattle, corn, and soybean), with a set of models some of which include price forecasts of other commodities.

2.3 Methodology

The Basics

In this paper, we used the Bayesian approach to deal with model specification uncertainty. To forecast hog price movements, we start with a set of possible forecasting models, estimate them all, and see which have the most posterior support from the data. This is done in two parts: the estimation of each model and the computation of each model's support.

For a given model j, assume a linear regression model:

$$y = X_j \beta_j + \epsilon_j, j = 1, \ldots, M, \tag{2.1}$$

where y is the vector of observations on hog prices assumed identical in all models, X_j is the matrix of the independent variables for the j^{th} model considered, ϵ_j is the vector of random errors for the j^{th} model, and j denotes the model in the set of M models considered. The differences between the models are restricted here to the matrix X of independent variables.

The prior distribution on the regression parameters β_j can be specified as

$$p(\beta_j) \sim N(b_{0j}, \sigma_j^2 V_{0j}), j = 1, \ldots, M, \tag{2.2}$$

where N represents the multivariate normal distribution, b_{0j} is the prior mean of the regression parameters for the j^{th} model and $\sigma_j^2 V_{0j}$ is the prior covariance matrix. The prior on σ_j^2 is specified as an inverse-gamma distribution, which is equivalent to a gamma distribution on σ_j^{-2},

$$p(\sigma_j^{-2}) \sim G(s_{0j}^{-2}, d_{0j}), j = 1, \ldots, M, \tag{2.3}$$

where G stands for the gamma distribution, s_{0j}^{-2} is the prior mean for the inverse error variance, and d_{0j} is the prior degrees of freedom. A higher value of d_{0j} indicates a more informative prior (Koop, 2003).

The likelihood function for each model can be specified as

$$L_j(y|\beta_j, \sigma_j^2, X_j) = (2\pi\sigma^2)^{-n/2} exp\{-0.5(y - X_j\beta_j)' \sigma_j^{-2}(y - X_j\beta_j)\}, j = 1, \ldots, M, \tag{2.4}$$

where the ϵ_j are assumed to follow a standard form of identically and independently distributed normal random variables.

Given these priors and the above likelihood function, the joint posterior distribution of β_j and σ_j^2 is derived according to Bayes Theorem that the posterior distribution is proportional to the prior distribution times the likelihood function. The joint posterior distribution is

$$p(\beta_j, \sigma_j^2|y, X_j) \sim NG(b_{pj}, V_{pj}, s_{pj}^2, d_{pj}), j = 1, \ldots, M, \tag{2.5}$$

where

$$V_{pj} = (V_{0j}^{-1} + X_j' X_j)^{-1}, \tag{2.6}$$

$$b_{pj} = V_{pj}(V_{0j}^{-1} b_{0j} + (X_j' X_j)\hat{\beta}_j), \tag{2.7}$$

$$d_{pj} = d_{0j} + n_j, \tag{2.8}$$

and

$$s_{pj}^2 = d_{pj}^{-1}[d_{0j}s_{0j}^2 + (n_j - k_j)s_j^2 + (\hat{\beta}_j - b_{0j})'(V_{0j} + (X_j'X_j)^{-1})^{-1}(\hat{\beta}_j - b_{0j})], \tag{2.9}$$

where NG represents the joint normal-gamma distribution, $\hat{\beta}_j$ and s_j^2 are the standard OLS quantities and n_j and k_j are the rows and columns of X_j , respectively. Equations (6) to (9) together help define the parameters in the distribution. $s_{pj}^2 V_{pj}$ is the posterior mean of the variance, b_{pj} is the posterior mean of the coefficients, which are the weighted averages of the parameters of the prior distribution and the parameters that are derived from the maximum likelihood estimator based on the data, and d_{pj} is the posterior degrees of freedom.

For each model, after generating point forecasts using the posterior means of the parameters found above and the actual values of the independent variables, we convert the point forecasts into directional forecasts using the simple rule:

$$f_{jt} = \begin{cases} 1 & \text{if } \hat{y}_{jt} \text{ - } y_{j,t-1} > 0 \\ 0 & \text{if } \hat{y}_{jt} \text{ - } y_{j,t-1} \leq 0 \end{cases}, j = 1, \ldots, M, \tag{2.10}$$

where f_{jt} denotes a dichotomous variable denoting a price forecast of either up (1) or down (0) and y_{jt} denotes the commodity price at t time period for j^{th} model, respectively. The set of f_{jt} are our qualitative forecasts.

Model Specification Uncertainty

Now we describe the process for handling model specification uncertainty. First, a discrete prior weight is assigned to each model

$$p(M_j) = \mu_j, \sum_{j=1}^{M} \mu_j = 1. \tag{2.11}$$

Here we choose to use uninformative priors across the model specification, so all models are treated equally. In this case, $\mu_j = 1/M$, $\forall j$. Then, using the above results for the posterior distributions shown in (5), we derive the marginal likelihood functions by integrating out the parameter uncertainty to leave

$$p(y|M_j) = c_j[|V_{pj}|/|V_{0j}|]^{1/2}(d_{pj}s_{pj}^2)^{-d_{pj}/2}, \tag{2.12}$$

where

$$c_j = \frac{\Gamma(d_{pj}/2)(d_{0j}s_{0j}^2)^{d_{0j}/2}}{\Gamma(d_{0j}/2)\pi^{n/2}}, \tag{2.13}$$

and Γ is the Gamma function. The marginal likelihood measures how well the model fits on average, where the averaging is over parameter values with posterior support. As shown in equation (12), the smaller the posterior mean of the variance is, the larger the marginal likelihood will be, which indicates that the better the model fits, the larger the marginal likelihood will be. Combining (11) and (12) by Bayes Theorem, the posterior probability of each model is given by

$$p(M_j|y) \propto \mu_j[|V_{pj}|/|V_{0j}|]^{1/2}(d_{pj}s_{pj}^2)^{-d_{pj}/2} = \mu_j p(y|M_j), j = 1, \ldots, M. \tag{2.14}$$

Dividing each value in (14) by the sum of the unnormalized posterior probabilities across all M models produces normalized posterior model probabilities that sum to one. Denote these normalized posterior probabilities by

$$\omega_j = \frac{\mu_j p(y|M_j)}{\sum\limits_{j=1}^{M} \mu_j p(y|M_j)}, j = 1, \ldots, M. \tag{2.15}$$

These posterior probabilities ω_j are the key to evaluating both general model specification uncertainty and the advantage of including forecasts of other commodity prices in the forecasting model. Models which receive higher posterior probabilities are better supported by the data, indicating that those models are preferred choices and can be expected to yield better forecasting performance. We further obtain the posterior support for model traits by summing the posterior probabilities of each variable across models, thereby determining ideal model specification. The higher the summed posterior probability of one variable, the more support that variable has for being included in the model specification.

We also form a composite forecast using the posterior model probabilities to construct a weighted average of all the individual model forecasts:

$$\hat{f}_t = \begin{cases} 1 & \text{if } \sum_j \omega_j f_{jt} \geq 0.5 \\ 0 & \text{otherwise} \end{cases}, j = 1, \ldots, M, \tag{2.16}$$

where $\hat{f}_t$ represents the composite forecast at t time period. Because this is qualitative forecasting, if the sum of the posterior model probabilities on the set of models that predicted 1 is greater than 0.50, the composite forecast is a 1.

2.4 Data

Data on the four commodity prices are collected from the Chicago Mercantile Exchange (CME) Group, using monthly futures prices for lean hog futures ($/lb), live cattle futures ($/lb), corn futures ($/bushel), and soybean futures ($/bushel). Possible independent variables, including autoregressive (AR) processes and exogenous variables, are selected based on analyses of previous studies in the literature.

For the hog price forecasting models, the AR terms to be considered range from AR(3) to AR(12) and the exogenous variables include monthly disposable personal income (logged), monthly commercial cattle slaughter (thousand heads), monthly broiler-type poultry eggs hatched (million eggs), monthly number of sows farrowing (thousand heads), and monthly pork cold storage (million pounds). For the cattle price forecasting models, the independent variables considered are the same as the hog price forecasting model except pork storage is not included. In the corn price forecasting models, the AR terms to be considered range from AR(3) to AR(6) and the exogenous variables to be included are monthly corn export (thousand units), monthly corn inventory (million bushels), monthly lagged acres planted to corn (thousand acres), and monthly fuel ethanol production (million gallons). For the soybean price forecasting models, the independent variables considered are the same as in the corn model except the ethanol variable is not included. All the data of the exogenous variables for the four commodity price forecasting models come from the National Agricultural Statistics Service (NASS). Additionally, the monthly lagged acres planted to corn/soybean is a weighted average of lagged acres of corn/soybean based on the Palmer Drought Severity Index (PDSI) for each state. The data on the PDSI are provided by the National Oceanic and Atmospheric Administration (NOAA).

All data are monthly extending from January 1981 to December 2013. We use the first twenty-six years (January 1981-December 2006) for in-sample estimation, and then evaluate out-of-sample forecasting performance over the last 84 observations, which are from January 2007 to December 2013.

Table 2.1 shows the set of variables considered in the model specification and the total number of forecasting models estimated for each of the four commodity prices. In the hog price forecasting model, the hog price (PH) to be forecast is the monthly lean hog futures price ($/lb) as given by CME group. Among the exogenous variables considered for the hog price forecasting model, DSPI denotes the natural logarithm of monthly disposable personal income (billion dollars); CTSL denotes the monthly commercial cattle slaughter (thousand heads); HATCH denotes the monthly broiler-type poultry eggs hatched (million eggs); SF denotes the monthly number of sows farrowing (thousand heads); PKST denotes the monthly pork cold storage (million pounds). In the cattle price forecasting model, the cattle price (PCA) to be forecast is the monthly live cattle futures price ($/lb) as given by CME group. The independent variables considered are basically the same as in the hog model except the PKST variable. In the corn price forecasting model, the corn price (PC) to be forecast is the monthly corn futures price ($/bushel) as given by CME group. Among the exogenous variables considered for the corn price forecasting model, EXPORT^c denotes the monthly corn export (thousand units); INVENTORY^c denotes the monthly corn inventory (million bushels); ACRES^c denotes the monthly lagged acreages planted for corn (thousand acres); ETHANOL denotes the monthly fuel ethanol production (million gallons). In the soybean price forecasting model, the soybean price (PS) to be forecast is the monthly soybean futures price ($/bushel) as given by CME group. The independent variables considered are the same as in the corn model except the ETHANOL variable. The data of the exogenous variables for the four commodity price forecasting models are provided by NASS. In addition, ACRES^c/ACRES^s is a weighted average lagged acres of corn/soybean based on the Palmer Drought Severity Index (PDSI) given by NOAA for each of the 48 states.

2.5 Empirical Results

Beginning with the hog price forecasting models, Table 2.2 presents the posterior probabilities for the model specification. The probabilities shown in Table 2.2 are the probability that each of the variables listed belongs in the true model. These probabilities show that there is clear and overwhelming support for the inclusion of AR(3) (0.992), disposable personal income (1.000), egg hatching (0.977), sows farrowing (1.000), and pork storage (0.999) in the hog price forecasting model. Also, cattle forecasts have a 0.878 posterior probability of inclusion. Other variables have little to no posterior support for inclusion in the hog price forecasting model. In terms of helping to uncover a model specification, the Bayesian approach provides excellent guidance.

Table 2.3 presents the out-of-sample forecasting performance of the 84 qualitative forecasts for the thirteen best and six worst forecasting models among the 420 specifications estimated. Note that the best and worst performing forecasting models are decided based on percent of correct predictions over the 2007-2013 period. The best performing model correctly forecast

65 out of 84 out-of-sample price movements (77.38 percent). Interestingly, seven of thirteen best performance models have longer autoregressive processes (with 11 or 12 lags) than the posterior model probabilities suggested would be best. Also, note that all these best models include cattle slaughter while the posterior probability suggests no inclusion.

Table 2.4 displays the percentage of correct out-of-sample forecasts of the five most probable and five least probable models; these are the models with the highest and lowest posterior model probabilities. The five most probable models are those that one would be most likely to choose ex ante before seeing out-of-sample forecasting performance. As shown in Tables 2.3 and 2.4, the five most probable and five best performing models include either one or more commodity forecasts. Over the 2007-2013 period, the fifth most probable model has better forecasting performance than the mean and median level of the set of all hog price forecasting models. In terms of the composite qualitative forecasts, the weighted average of all the 420 individual model forecasts over 2007-2013 period, correctly forecast 57 out of 84 out-of-sample price movements (67.86 percent), that is close to the mean and median level, although not as well as the best performance.

Moving to the cattle price forecasting models, Table 2.5 presents the posterior probabilities in favor of variable inclusion in the cattle forecasting model. These results show that disposable personal income (0.998), cattle slaughter (0.965), sows farrowing (0.990), and hog price forecasts (0.998) have enormous support for inclusion in the cattle price forecasting model. Other variables have little to no posterior support for inclusion in the cattle price forecasting model.

Tables 2.6 and 2.7 show the correct prediction percentage of the best/worst performing models and the most/least probable models, respectively. It is found that the most probable and best performing models all have disposable personal income as an exogenous variable and favor shorter AR process. In Table 2.7, over the 2007-2013 period, the fifth most probable models have better forecasting performance than the mean and median level of the total 350 cattle price forecasting models. We also find that the composite forecasts, computed from the total 350 cattle models over the 2007-2013 period, have 43 correct forecasts out of 84 out-of-sample price movements (51.19 percent); however, they perform worse than the mean and median level as well as than the most probable model.

Next for the corn price forecasting models, Table 2.8 presents the posterior probabilities in favor of variable inclusion in the corn price forecasting model. AR(3) has a 0.999 posterior probability of inclusion, ethanol production has a 0.864 probability. No other variables have posterior support that reaches 0.20, so the model specification is also quite clear.

The best performing and most probable models all include either hog or cattle price forecasts, as shown in Tables 2.9 and 2.10. As measured by the percentage of correct forecasts over the 2007-2013 period, the best performance models correctly forecast 60 out of 84 out-of-sample price movements (71.43 percent). In Table 2.10, four of the five most probable models beat the mean and median level. However, the composite forecasts computed based on the total 308 corn models over the 2007-2013 period correctly forecast 50 out of 84 out-of-sample price movements (59.52 percent), similar to the mean and median level but worse than the most probable models.

Finally, the soybean price forecasting model specification results are presented in Table 2.11. The posterior probabilities show strong support for including AR(3) (0.9940) and soybean export (0.9591) in the soybean price forecasting model.

Table 2.12 presents the correct prediction percentage for the eight best and seven worst performing forecasting models, while Table 2.13 displays the correct prediction percentage for the five most and five least probable models. Over the 2007-2013 period, the best performance models has 63.10 percentage of correct predictions. Among the top five probable models, four of them have above average forecasting performance and particularly the third probable model has 60.71 percentage of correct predictions, fairly close to the best forecasting performance. Although the composite forecasts computed based on the total 252 soybean models beat the most probable model, unfortunately they perform worse than the mean and median level. This suggests that our Bayesian Model Averaging methodology does not works for the soybean price qualitative forecasting model as well as other commodity price qualitative forecasting models.

Overall, the most probable models for each commodity price display around average forecasting performance among the entire set of models estimated. Yet, while the forecasting performance of the most probable models is not what we might have hoped for, we find that within the lists of the best forecasting models for each of the four commodity prices, models that include commodity price forecasts are heavily represented. Of those thirty-two best performing models, twenty-nine include one or more commodity price forecasts. This suggests that it is worth pursuing how commodity price forecasts can be improved by the inclusion of other commodity price forecasts in the forecasting models.

2.6 Conclusions

The Bayesian Model Averaging methodology applied here for model specification to the forecasting of four important commodity prices provides clear signals for variable inclusion in the forecasting models, although the results of the Bayesian Model Averaging are somewhat mixed with regard to signaling which models are likely to have the best out-of-sample forecasting performance. Based on our findings, in general, the models with the highest model probabilities based on the in-sample data deliver around average out-of-sample forecasting performance. For price forecasting of hog and soybean, the composite qualitative forecasts computed under the Bayesian framework outperform the most probable model among the entire set of models estimated but that is not the case for cattle and corn price forecasting. Also, the fact that twenty-nine of the thirty-two best performing forecasting model, as measured by the percentage of correct out-of-sample forecasts, contain price forecasts for one or more different commodities suggests that the idea of improving commodity price forecasting by including the composite forecasts of other commodities in the model is a good one. Still, additional work is needed to evaluate multiple models based on the out-of-sample forecasting performance so that users of such forecasts can have some scientific basis for choosing a model specification (including possibly using a composite forecast). We believe the results here show that we are on the right track, but have not yet arrived at our desired destination.

Table 2.1: Variables Used to Predict Commodity Prices

Dependent Variable	Lags	Exogenous Variables
PH (cents per pound) (420 models)	AR(3) $\sim$ AR(12)	DSPI; CTSL; HATCH; SF ;PKST
PCA (cents per pound) (350 models)	AR(3) $\sim$ AR(12)	DSPI; CTSL; HATCH; SF
PC (cents per bushel) (308 models)	AR(3) $\sim$ AR(6)	EXPORT^c; INVENTORY^c; ACRES^c; ETHANOL
PS (10 cents per bushel) (252 models)	AR(3) $\sim$ AR(6)	EXPORT^s; INVENTORY^s; ACRES^s

Table 2.2: Hog Price Forecasting Model Specification (420 Models)

Model Traits	Post Probability
Include AR(3)	0.992
Include DSPI	1.000
Include CTSL	0.024
Include HATCH	0.978
Include SF	1.000
Include PKST	0.999
Include Cattle Forecasts	0.878
Include Corn Forecasts	0.070
Include Soybean Forecasts	0.054
No Forecasts	<0.001

Table 2.3: Top 13 and Bottom 6 Hog Price Forecasting Models by the Percentage of Correct Out-of-Sample Forecasts

Top 13 Models	% Forecasts Correct	Post Probability
1) AR(3)+DSPI$_t$+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+Cattle Forecasts$_t$+Soybean Forecasts$_t$	0.774	<0.001
2) AR(4)+DSPI$_t$+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+Cattle Forecasts$_t$+Soybean Forecasts$_t$	0.762	<0.001
3) AR(11)+DSPI$_t$+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +PKST$_{t-1}$	0.750	<0.001
3) AR(9)+DSPI$_t$+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+Cattle Forecasts$_t$+Soybean Forecasts$_t$	0.750	<0.001
3) AR(7)+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Cattle Forecasts$_t$+Soybean Forecasts$_t$	0.750	<0.001
3) AR(4)+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Cattle Forecasts$_t$+Soybean Forecasts$_t$	0.750	<0.001
3) AR(12)+DSPI$_t$+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+Cattle Forecasts$_t$+Corn Forecasts$_t$	0.750	<0.001
3) AR(11)+DSPI$_t$+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+Cattle Forecasts$_t$+Corn Forecasts$_t$	0.750	<0.001
3) AR(11)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$ +Cattle Forecasts$_t$+Corn Forecasts$_t$	0.750	<0.001
3) AR(12)+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Cattle Forecasts$_t$+Corn Forecasts$_t$	0.750	<0.001
3) AR(9)+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Soybean Forecasts$_t$	0.750	<0.001
3) AR(12)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$ +Corn Forecasts$_t$	0.750	<0.001
3) AR(11)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$ +Corn Forecasts$_t$	0.750	<0.001
Bottom 6 Models	**% Forecasts Correct**	**Post Probability**
1) AR(4)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$	0.512	<0.001
2) AR(6)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$ +Cattle Forecasts$_t$	0.524	<0.001
3) AR(5)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$	0.536	<0.001
3) AR(6)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$	0.536	<0.001
3) AR(5)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$ +Cattle Forecasts$_t$	0.536	<0.001
3) AR(7)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$ +Cattle Forecasts$_t$	0.536	<0.001
Mean	0.683	
Median	0.690	
Composite Forecasts	0.679	

Table 2.4: Top 5 and Bottom 5 Hog Price Forecasting Models by Posterior Probability

5 Most Probable Models	Post Probability	% Forecasts Correct
1) AR(3)+DSPI$_t$+HATCH$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$ +Cattle Forecasts$_t$	0.856	0.679
2) AR(3)+DSPI$_t$+HATCH$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$ +Corn Forecasts$_t$	0.060	0.607
3) AR(3)+DSPI$_t$+HATCH$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$ +Soybean Forecasts$_t$	0.049	0.643
4) AR(3)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$ +Cattle Forecasts$_t$	0.011	0.548
5) AR(3)+DSPI$_t$+CTSL$_{t-1,t-2}$+SF$_{t-1}$+PKST$_{t-1}$ +Corn Forecasts$_t$	0.007	0.726
5 Least Probable Models	**Post Probability**	**% Forecasts Correct**
1) AR(12)+CTSL$_{t-1}$+HATCH$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Corn Forecasts$_t$+Soybean Forecasts$_t$	<0.001	0.726
2) AR(12)+DSPI$_t$+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$ +SF$_{t-1}$+PKST$_{t-1}$+Corn Forecasts$_t$ +Soybean Forecasts$_t$	<0.001	0.690
3) AR(12)+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Cattle Forecasts$_t$+Soybean Forecasts$_t$	<0.001	0.702
4) AR(11)+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Corn Forecasts$_t$+Soybean Forecasts$_t$	<0.001	0.714
5) AR(12)+CTSL$_{t-1,t-2}$+HATCH$_{t-1,t-2}$+SF$_{t-1}$ +PKST$_{t-1}$+Cattle Forecasts$_t$+Corn Forecasts$_t$	<0.001	0.750
Mean		0.683
Median		0.690
Composite Forecasts		0.679

Table 2.5: Cattle Price Forecasting Model Specification (350 Models)

Model Traits	Post Probability
Include AR(3)	0.321
Include AR(6)	0.451
Include DSPI	0.998
Include CTSL	0.965
Include HATCH	0.047
Include SF	0.990
Include Hog Forecasts	0.998
Include Corn Forecasts	0.002
Include Soybean Forecasts	0.001
No Forecasts	<0.001

Table 2.6: Top 7 and Bottom 3 Cattle Price Forecasting Models by the Percentage of Correct Out-of-Sample Forecasts

Top 7 Models	% Forecasts Correct	Post Probability
1) AR(5)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$ +Corn Forecasts$_t$	0.655	<0.001
1) AR(5)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$+Soybean Forecasts$_t$	0.655	<0.001
3) AR(5)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$	0.643	<0.001
3) AR(4)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$	0.643	<0.001
3) AR(5)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Soybean Forecasts$_t$	0.643	<0.001
3) AR(5)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$ +Hog Forecasts$_t$+Corn Forecasts$_t$	0.643	<0.001
3) AR(5)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Corn Forecasts$_t$+Soybean Forecasts$_t$	0.643	<0.001
Bottom 3 Models	**% Forecasts Correct**	**Post Probability**
1) AR(4)+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	0.429	<0.001
2) AR(3)+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$+Corn Forecasts$_t$	0.440	<0.001
2) AR(3)+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Corn Forecasts$_t$	0.440	<0.001
Mean	0.559	
Median	0.560	
Composite Forecasts	0.512	

Table 2.7: Top 5 and Bottom 5 Cattle Price Forecasting Models by Posterior Probability

5 Most Probable Models	Post Probability	% Forecasts Correct
1) AR(6)+DSPI$_t$+CTSL$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	0.424	0.536
2) AR(3)+DSPI$_t$+CTSL$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	0.305	0.512
3) AR(5)+DSPI$_t$+CTSL$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	0.161	0.524
4) AR(4)+DSPI$_t$+CTSL$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	0.058	0.488
5) AR(6)+DSPI$_t$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Hog Forecasts$_t$	0.025	0.595
5 Least Probable Models	**Post Probability**	**% Forecasts Correct**
1) AR(12)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$ +SF$_{t-1,t-2}$+Corn Forecasts$_t$+Soybean Forecasts$_t$	<0.001	0.631
2) AR(12)+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Corn Forecasts$_t$+Soybean Forecasts$_t$	<0.001	0.524
3) AR(11)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$ +SF$_{t-1,t-2}$+Corn Forecasts$_t$+Soybean Forecasts$_t$	<0.001	0.631
4) AR(12)+DSPI$_t$+CTSL$_{t-1}$+HATCH$_{t-1}$ +SF$_{t-1,t-2}$+Hog Forecasts$_t$+Soybean Forecasts$_t$	<0.001	0.607
5) AR(11)+CTSL$_{t-1}$+HATCH$_{t-1}$+SF$_{t-1,t-2}$ +Corn Forecasts$_t$+Soybean Forecasts$_t$	<0.001	0.524
Mean		0.559
Median		0.560
Composite Forecasts		0.512

Table 2.8: Corn Price Forecasting Model Specification (308 Models)

Model Traits	Post Probability
Include EXPORT$_t^c$	<0.001
Include INVENTORY$_t^c$	0.125
Include ETHANOL$_t$	0.864
Include ACRES$_t^c$	0.011
Include AR(3)	0.999
Include Hog Forecasts	0.009
Include Cattle Forecasts	0.098
Include Soybean Forecasts	<0.001
No Forecasts	0.894

Table 2.9: Top 4 and Bottom 13 Corn Price Forecasting Models by the Percentage of Correct Out-of-Sample Forecasts

Top 4 Models	% Forecasts Correct	Post Probability
1) AR(3)+INVENTORY^c_t+ETHANOL$_{t-1}$ +Cattle Forecasts$_t$	0.714	<0.001
1) AR(4)+INVENTORY^c_t+ETHANOL$_{t-1}$ +Cattle Forecasts$_t$	0.714	<0.001
1) AR(3)+INVENTORY^c_t+ETHANOL$_{t-1}$ +Cattle Forecasts$_t$+Soybean Forecasts$_t$	0.714	<0.001
1) AR(4)+INVENTORY^c_t+ETHANOL$_{t-1}$ +Cattle Forecasts$_t$+Soybean Forecasts$_t$	0.714	<0.001
Bottom 13 Models	**% Forecasts Correct**	**Post Probability**
1) AR(5)+ACRES^c_t+Hog Forecasts$_t$	0.524	<0.001
1) AR(6)+ACRES^c_t+Hog Forecasts$_t$	0.524	<0.001
1) AR(5)+ACRES$^c_{t-1}$+Hog Forecasts$_t$	0.524	<0.001
1) AR(6)+ACRES$^c_{t-1}$+Hog Forecasts$_t$	0.524	<0.001
1) AR(3)+EXPORT^c_t+ETHANOL$_{t-1}$ +Cattle Forecasts$_t$	0.524	<0.001
1) AR(3)+ACRES^c_t+Hog Forecasts$_t$ +Soybean Forecasts$_t$	0.524	<0.001
1) AR(4)+ACRES^c_t+Hog Forecasts$_t$ +Soybean Forecasts$_t$	0.524	<0.001
1) AR(5)+ACRES^c_t+Hog Forecasts$_t$ +Soybean Forecasts$_t$	0.524	<0.001
1) AR(6)+ACRES^c_t+Hog Forecasts$_t$ +Soybean Forecasts$_t$	0.524	<0.001
1) AR(3)+ACRES$^c_{t-1}$+Hog Forecasts$_t$ +Soybean Forecasts$_t$	0.524	<0.001
1) AR(4)+ACRES$^c_{t-1}$+Hog Forecasts$_t$ +Soybean Forecasts$_t$	0.524	<0.001
1) AR(5)+ACRES$^c_{t-1}$+Hog Forecasts$_t$ +Soybean Forecasts$_t$	0.524	<0.001
1) AR(6)+ACRES$^c_{t-1}$+Hog Forecasts$_t$ +Soybean Forecasts$_t$	0.524	<0.001
Mean	0.595	
Median	0.595	
Composite Forecasts	0.595	

Table 2.10: Top 5 and Bottom 5 Corn Price Forecasting Models by Posterior Probability

5 Most Probable Models	Post Probability	% Forecasts Correct
1) AR(3)+ETHANOL$_t$	0.396	0.643
2) AR(3)+ETHANOL$_{t-1}$	0.373	0.619
3) AR(3)+INVENTORY$_t^c$	0.088	0.583
4) AR(3)+ETHANOL$_t$+Cattle Forecasts$_t$	0.044	0.643
5) AR(3)+ETHANOL$_{t-1}$+Cattle Forecasts$_t$	0.041	0.631
5 Least Probable Models	**Post Probability**	**% Forecasts Correct**
1) AR(6)+ACRES$_t^c$+INVENTORY$_t^c$ +Cattle Forecasts$_t$+Soybean Forecasts$_t$	<0.001	0.607
2) AR(6)+ACRES$_t^c$+INVENTORY$_t^c$ +Hog Forecasts$_t$+Soybean Forecasts$_t$	<0.001	0.583
3) AR(6)+ACRES$_t^c$+ETHANOL$_t^c$ +Cattle Forecasts$_t$+Soybean Forecasts$_t$	<0.001	0.619
4) AR(6)+ACRES$_t^c$+INVENTORY$_t^c$ +Soybean Forecasts$_t$	<0.001	0.583
5) AR(6)+ACRES$_t^c$+INVENTORY$_t^c$ +Hog Forecasts$_t$+Cattle Forecasts$_t$	<0.001	0.595
Mean		0.595
Median		0.595
Composite Forecasts		0.595

Table 2.11: Soybean Price Forecasting Model Specification (252 Models)

Model Traits	Post Probability
Include AR(3)	0.999
Include EXPORT^s	0.970
Include INVENTORY^s	0.028
Include ACRES^s	0.002
Include Hog Forecasts	0.607
Include Cattle Forecasts	0.367
Include Corn Forecasts	0.029
No Forecasts	<0.001

Table 2.12: Top 8 and Bottom 7 Soybean Price Forecasting Models by the Percentage of Correct Out-of-Sample Forecasts

Top 8 Models	% Forecasts Correct	Post Probability
1) AR(3)+EXPORT_t^s+INVENTORY_{t-1}^s +Corn Forecasts_t	0.631	<0.001
1) AR(5)+EXPORT_t^s+Cattle Forecasts_t +Corn Forecasts_t	0.631	<0.001
1) AR(6)+EXPORT_t^s+Cattle Forecasts_t +Corn Forecasts_t	0.631	<0.001
1) AR(6)+EXPORT_{t-1}^s+Cattle Forecasts_t +Corn Forecasts_t	0.631	<0.001
1) AR(5)+EXPORT_t^s+INVENTORY_{t-1}^s +Cattle Forecasts_t+Corn Forecasts_t	0.631	<0.001
1) AR(6)+EXPORT_t^s+INVENTORY_{t-1}^s +Cattle Forecasts_t+Corn Forecasts_t	0.631	<0.001
1) AR(5)+EXPORT_t^s+ACRES_t^s+Cattle Forecasts_t +Corn Forecasts_t	0.631	<0.001
1) AR(6)+EXPORT_t^s+ACRES_t^s+Cattle Forecasts_t +Corn Forecasts_t	0.631	<0.001
Bottom 7 Models	**% Forecasts Correct**	**Post Probability**
1) AR(6)+ACRES_t^s+Hog Forecasts_t	0.440	<0.001
2) AR(6)+ACRES_{t-1}^s+Hog Forecasts_t	0.452	<0.001
3) AR(3)+ACRES_t^s+Hog Forecasts_t	0.464	<0.001
3) AR(4)+ACRES_t^s+Hog Forecasts_t	0.464	<0.001
3) AR(3)+ACRES_{t-1}^s+Hog Forecasts_t	0.464	<0.001
3) AR(4)+ACRES_{t-1}^s+Hog Forecasts_t	0.464	<0.001
3) AR(5)+ACRES_{t-1}^s+Hog Forecasts_t	0.464	<0.001
Mean	0.552	
Median	0.548	
Composite Forecasts	0.536	

Table 2.13: Top 5 and Bottom 5 Soybean Price Forecasting Models by Posterior Probability

5 Most Probable Models	Post Probability	% Forecasts Correct
1) AR(3)+EXPORT$_t^s$+Hog Forecasts$_t$	0.295	0.524
2) AR(4)+EXPORT$_{t-1}^s$+Hog Forecasts$_t$	0.292	0.548
3) AR(3)+EXPORT$_t^s$+Cattle Forecasts$_t$	0.204	0.607
4) AR(3)+EXPORT$_{t-1}^s$+Cattle Forecasts$_t$	0.149	0.548
5) AR(3)+EXPORT$_t^s$+Corn Forecasts$_t$	0.015	0.571
5 Least Probable Models	Post Probability	% Forecasts Correct
1) AR(6)+ACRES$_{t-1}^s$+INVENTORY$_t^s$	<0.001	0.524
2) AR(6)+ACRES$_t^s$+EXPORT$_t^s$	<0.001	0.560
3) AR(6)+ACRES$_{t-1}^s$+INVENTORY$_t^s$ +Cattle Forecasts$_t$+Corn Forecasts$_t$	<0.001	0.560
4) AR(6)+ACRES$_{t-1}^s$+INVENTORY$_t^s$ +Hog Forecasts$_t$+Corn Forecasts$_t$	<0.001	0.607
5) AR(5)+ACRES$_{t-1}^s$+INVENTORY$_t^s$	<0.001	0.536
Mean		0.552
Median		0.548
Composite Forecasts		0.536

Bibliography

Allen, P.G. "Economic Forecasting in Agriculture." *International Journal of Forecasting* 10(1994):81-135.

Brandt, J.A. "Forecasting and Hedging: An Illustration of Risk Reduction in the Hog Industry." *American Journal of Agricultural Economics* 67(1985):24-31.

Brandt, J.A. and D.A. Bessler. "Composite Forecasting: An Application with U.S. Hog Prices." *American Journal of Agricultural Economics* 63(1981):135-140.

Brandt, J.A. and D.A. Bessler. "Price Forecasting and Evaluation: An Application in Agriculture." *Journal of Forecasting* 2(1983):237-248.

Cromarty, W.A. and W.M. Myers. "Needed Improvements in Application of Models for Agriculture Commodity Price Forecasting." *American Journal of Agricultural Economics* 57(1975):172-177.

Dorfman, J.H. "Bayesian Composite Qualitative Forecasting: Hog Prices Again." *American Journal of Agricultural Economics* 80(1998):543-551.

Dorfman, J.H. and D.R. Sanders. "Generalized Hedge Ratio Estimation with an Unknown Model." *Review of Futures Markets* 15(2006):167-188.

Feather, P.M. and M.S. Kaylen. "Conditional Qualitative Forecasting." *American Journal of Agricultural Economics* 71(1989):195-201.

Gerlow, M.E., S.H. Irwin, and T.R. Liu. "Economic Evaluation of Commodity Price Forecasting Models." *International Journal of Forecasting* 9(1993):387-397.

Goodwin, B.K. and R. Schnepf. "Determinants of Endogenous Price Risk in Corn and Wheat Futures Markets." *Journal of Futures Markets* 20(2000):753-774.

Kastens, T.L. and T.C. Schroeder. "Efficiency Tests of July Kansas City Wheat Futures." *Journal of Agricultural and Resource Economics* 21(1996):187-198.

Koop, G. *Bayesian Econometrics.* Chichester, England: Wiley, 2003.

Leuthold, R.M., A.J.A. MacCormick, A. Schmitz, and D.G. Watts "Forecasting Daily Hog Prices and Quantities: A Study of Alternative Forecasting Techniques." *Journal of the American Statistical Association* 65(1970):90-107.

Marsh, J.M. "A Rational Distributed Lag Model of Quarterly Live Cattle Prices." *American Journal of Agricultural Economics* 65(1983):539-547.