Do Composite Procedures Really Improve the Accuracy of Outlook Forecasts?

by

Evelyn V. Colino, Scott H. Irwin, and Philip Garcia

Suggested citation format:

Colino, E. V., S. H. Irwin, and P. Garcia. 2009. "Do Composite Procedures Really Improve the Accuracy of Outlook Forecasts?" Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.uiuc.edu/nccc134].

Do Composite Procedures Really Improve the Accuracy of Outlook Forecasts?

Evelyn V. Colino

Scott H. Irwin

Philip Garcia[1]

Paper presented at the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management
St. Louis, Missouri, April 20-21, 2009

---

[1] Evelyn V. Colino is a Graduate Research Assistant for the AgMAS Project in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign. Scott H. Irwin is the Laurence J. Norton Chair of Agricultural and Consumer Economics at the University of Illinois, Urbana, Illinois. Philip Garcia is the T.A. Hieronymus Distinguished Chair in Futures Markets. A number of individuals generously contributed outlook forecast data to this project: the late Gene Futrell, Darrel Good, Glenn Grimes, Chris Hurt, Dave Miller, John Lawrence, Ron Plain, and Laurian Unnevehr. We are indebted to each of them for their invaluable assistance. Funding support from the Aurene T. Norton Trust is gratefully acknowledged.

Do Composite Procedures Really Improve the Accuracy of Outlook Forecasts?

Practitioner's Abstract

This paper investigates whether the accuracy of outlook hog price forecasts can be improved using composite forecasts in an out-of-sample context. Price forecasts from four well-recognized outlook programs are combined with futures-based forecasts, ARIMA, and unrestricted Vector Autoregressive (VAR) models. Quarterly data are available from 1975.I through 2007.IV, which allow for a relatively long out-of-sample evaluation period after permitting model specification and appropriate composite-weight training periods. Results show that futures and numerous composite procedures outperform outlook forecasts. At intermediate horizons, OLS composite procedures perform rather well. The superiority of futures and composite forecasts decreases at longer horizons except for an equal-weighted approach. Importantly, with just few exceptions, nothing outperforms the equal-weight approach significantly in any program or horizon. Overall, findings favor the usage of equal-weighted composites, a result that is consistent with previous empirical findings and recent theoretical papers.

Key words: forecast combination, outlook, futures, time-series, out-of-sample

Introduction

U.S. public outlook forecasts are viewed as a valuable source of information and traditionally have played an important role in agricultural decision making. As a consequence, agricultural economists have devoted considerable effort to evaluate their forecast performance (e.g., Allen 1994). Previous studies by Colino and Irwin (2007), and Colino, Irwin, and Garcia (2008) address the accuracy of livestock price forecasts and assess whether outlook forecasts can be improved through combination with other forecasts. Colino and Irwin (2007) evaluate the accuracy of four outlook programs relative to futures prices in the hog and cattle markets. Results indicate that futures prices outperform outlook forecasts in most root mean squared error (RMSE) comparisons with several statistically significant differences. However, in a forecast encompassing framework, forecast combinations between outlook and futures are found to outperform futures alone.

Colino, Irwin, and Garcia (2008) investigate whether the predictability of outlook program hog price forecasts can be improved. Examining numerous time-series forecasts and futures-based forecasts, the findings suggest that these alternative forecasts can substantially improve the performance of outlook forecasts. Evidence from an encompassing analysis indicates that simple combinations of outlook and any of the alternative forecasts are able to reduce the errors of outlook alone by economically significant levels. Averaging of multiple forecast models also is shown to improve the accuracy of outlook forecasts.

These findings highlight the benefits of combining forecasts in livestock markets, but raise unanswered questions. Forecast combination is a well-established approach when two or more competing forecasts are available since a smaller variance generally can be obtained by combining individual forecasts. Also, pooling forecasts is rewarding when alternative forecasts contain useful information not included in the other, a concept closely related to forecast encompassing (e.g., Granger and Newbold 1986; Diebold and Lopez 1995). In this context, encompassing tests provide a straightforward but useful framework to identify the weights of the composite forecasts (Fang 2003). However, the same data or period often is used both to estimate the composite weights—as a product of encompassing regressions—and to assess the performance of the composite forecasts. Results from Colino and Irwin (2007) are subject to this limitation. A more complete out-of-sample evaluation of composite forecasts requires dividing the data into three periods: 1) one for estimating individual models, 2) another for investigating out-of-sample performance of individual models and identifying optimal composite weights, and 3) a final for out-of-sample comparison of forecast accuracy. A relevant question then is whether forecast combination in livestock markets is still a useful and valuable technique in a "more realistic" out-of-sample evaluation.

A second issue is related to how to combine forecasts. Theoretical research and applications of combinatorial forecast methods have grown significantly in the recent decades, from simple combination schemes to more sophisticated approaches (Bates and Granger 1969; Granger and Ramanathan 1984; McIntosh and Bessler 1988; Bessler and Chamberlain 1988; Clark and McCracken 2004, 2006; Stock and Watson 2004; Capistran and Timmermann 2009). A systematic comparison of alternative techniques would provide insight into which procedures lead to more accurate forecasts. This notion is consistent with Clemen (1989) who emphasizes, "…combining forecasts has been shown to be practical, economical, and useful…We no longer need to justify this methodology. We do need to find ways to make the implementation of the technique easy and efficient."

The purpose of the paper is to analyze whether the accuracy of outlook hog price forecasts can be improved through composite forecasts in a realistic out-of-sample context. Price forecasts from four well-recognized outlook programs are combined with forecasts based on futures and ARIMA and unrestricted Vector Autoregressive (VAR) models. The outlook programs are the same as those evaluated by Colino and Irwin (2007). The time-series models are relatively simple and performed well in the composite analysis conducted by Colino, Irwin and Garcia (2008). Quarterly data are available from 1975.I through 2007.IV for the analysis, allowing for a relatively long out-of-sample evaluation after accounting for model specification and appropriate composite-weight training periods.

Reasonably straightforward pooling techniques are considered. Several procedures have not been applied to agricultural markets previously and their accuracy is compared to more standard composite procedures. Composite methods include: equally-weighted average, equally-weighted average with a bias correction, weights based on constrained and unconstrained regressions, time-varying weights based on mean squared error, an odds matrix approach, and a shrinkage approach. These procedures represent nearly all categories of composite techniques identified by Timmerman (2006).

Individual Forecast Models

Four individual forecasts are used: an outlook forecast, a futures-based forecast, and forecasts from two time-series models, a VAR and an ARIMA model. These models provide a representation of the main forecast approaches available for agricultural markets. Outlook forecasts and forecasts based on futures prices are probably the most relevant instruments for agricultural decision makers when planning future action. Quarterly finished hog price forecasts from four different outlook programs are available for comparison. The outlook price forecasts evaluated are issued by University of Illinois in combination with Purdue University, Iowa State University, University of Missouri, and the Economic Research Service of the USDA. Information about sample periods, timing of release, target cash prices, and sources for each outlook forecast series are provided in Colino and Irwin (2007).

Price forecasts generated from futures prices are constructed using the procedure developed by Hoffman (2005). For each calendar month, the model uses the nearest-to-maturity contract. An average of futures prices observed on the day the outlook forecasts is released is used to represent the quarterly average futures price. Lean hog futures prices are then converted to live hog units to make them comparable to outlook forecasts which are reported in live weight terms.[1]

The third and fourth forecast models are an AR(5) and an unrestricted VAR(5). These models have been widely used in the livestock forecasting literature, are relatively easy to compute, and exhibited good performance in the composite analysis conducted by Colino, Irwin, and Garcia (2008). The VAR(5) specification is highly consistent with previous hog price forecasting models and is the result of a thorough process of examining potential variables, structural changes undergone in the industry, and preliminary estimations of reduced VARs. It is a five-variable system with a five fixed-lag structure, which was determined by the Akaike's Information Criteria (AIC), Final Prediction Error (FPE), and Hannan and Quinn' Information Criterion (HQIC). The variables composing the VAR are live-hog prices, corn prices, number of sows farrowing, pork production, and beef prices. More information about the VAR specification and variables selected are available in Colino, Irwin, and Garcia (2008).

Each outlook forecast has a different target cash price, which also has varied over time. To allow for the different program-related target cash prices, futures prices are adjusted by a program-related basis. Historical basis levels are computed by averaging daily futures prices for each quarter and subtracting the quarterly target cash price specified by the outlook forecast. Following Garcia and Sanders (1996), ARMA basis forecasts are estimated and the futures prices are adjusted. Time-series models are specified and estimated for cash prices of the specific program.

Forecast Combination Methods

Theoretical and empirical research on forecast composites is extensive (Timmermann 2006). A key question is whether forecast combination weights should be estimated or simple averages should be used. Empirical evidence has shown that it is difficult to outperform simple procedures such as an arithmetic average (Clemen 1989; Makridakis and Hibon 2000; Stock and Watson 2004). The effect of parameter estimation error can be a significant determinant of the performance of composite forecasting models. For instance, least squares weight (slope) estimation procedures (e.g., Granger and Ramanathan 1984) require the estimation of covariances between forecast errors, which can introduce an additional source of error in weight estimation. Simple combination schemes like arithmetic averages, although perhaps biased, do not require the estimation of covariances across model errors so they offer a potentially attractive option (Timmermann 2006). This question can be viewed as a tradeoff between imposing equal weights for each individual forecast which could lead to bias—a clearly a suboptimal scheme, and estimating the weights which could lead to a loss in efficiency. This bias-efficiency trade-off gives rise to the so-called *forecast combination puzzle* recently analyzed by Smith and Wallis (2009) and Issler and Lima (2009).

After review of available approaches, we follow Capistran and Timmermann (2009) who assess nearly all categories of composite procedures outlined in Timmermann (2006). All approaches assume that the combination rule takes a linear additive form,

(1) $$f^{c}_{t+h/t}(P) = \sum_{i=1}^{k} w^{i}_{t} f^{i}_{t+h/t}$$

where, w_t^i is the estimated weight on model i at time t, K is the number of h-step-ahead forecasts of the hog price P, $f^i_{t+h|t} = \hat{P}^i_{t+h|t}$ is the i^{th} forecast model available at time t, and $f^c_{t+h|t}$ is the composite forecast. The task is to derive the optimal weight to be assigned to each forecast.

The first approach follows a naïve decision rule that uses the **best-previous forecast**. For each forecast origin t, we identify the individual forecast with the most accurate historical performance and use it to forecast into the future (Clark and McCracken 2006; Capistran and Timmermann 2009). In this approach, the forecast with the lowest mean-squared error (MSE) receives all the weight,

$$\hat{P}_{t+h/t} = \hat{P}^{i_t^*}_{t+h/t}$$

where

(2) $$i_t^* = \arg\min_{i=1,\ldots N_t} t^{-1} \sum_{\tau=1}^{t} (P_\tau - P^{i}_{\tau/\tau-h})^2$$

A second composite forecast uses an arithmetic average of the four individual forecasts. It is also called the **equal-weighted composite forecast**,

(3) $$\overline{P}_{t+h/t} = K_t^{-1} \sum_{i=1}^{k_t} \hat{P}^{(i)}_{t+h/t}$$

A third composite uses a **projection of the equal-weighted combination,**

(4) $$\tilde{P}_{t+h/t} = \hat{\alpha}_t + \hat{\beta}_t \overline{P}_{t+h/t}$$

This regression on the equal-weighted forecast includes a constant to adjust for potential biases in the individual forecasts as well as in the aggregate and allows the slope coefficient to differ from unity. The combination is an extension of (2) and only requires the estimation of α and β. It is a potentially useful technique since it uses information from all forecasts in the average, but adjusts for possible bias and noise in the aggregate forecast. Capistran and Timmermann (2009) found this procedure to possess good overall forecast performance.

Probably the most common procedure for estimating combination weights is through **least squares regressions** (Bates and Granger 1969; Nelson 1972; Granger and Ramanathan 1984). In matrix notation, the k-vector of weights, $\widehat{W}$ is estimated by regressing the actual values of the target variable, P_N on the K-vector of forecasts, $\hat{P}_{N|N-h}$ over the period $N = 1, \dots, T$

(5) $$\hat{W}_T = (\sum_{N=1}^{T-1} \hat{P}_{N+h/N} \hat{P}'_{N+h/N})^{-1} \sum_{N=1}^{T-1} \hat{P}_{N+h/N} P_{N+h}$$

Following Granger and Ramanathan (1984), three versions of the procedure are examined,

(6) $$P_{t+h} = w_t' \hat{P}_{t+h/t} + e_{t+h} \rightarrow s.t. w_t' \iota = 1$$

(7) $$P_{t+h} = w_t^o + w_t' \hat{P}_{t+h/t} + e_{t+h}$$

(8) $$P_{t+h} = w_t' \hat{P}_{t+h/t} + e_{t+h}$$

Equation (6) is a **constrained regression** which requires composite weights to sum up to unity and individual forecasts to be unbiased to guarantee the combined forecast also is unbiased. This procedure has been used in many studies and is directly related to Harvey, Leybourne, and Newbold's (1998) encompassing test. Equation (7) is an **unconstrained regression** since it allows for bias in the individual forecasts, which can be corrected by the constant w_t^0, and does not require the weights to sum to unity. Equation (8) is a simple variation of equation (6) without a constant. The equations are estimated by standard OLS.

As mentioned, errors in the combination weights tend to be high in procedures like least squares regressions requiring the estimation of the covariance matrix of forecast errors. An alternative is to ignore the correlations of forecast errors and to weight each forecast by some measure of relative performance. Consider equation (9) which reduces the regression-based approach to a procedure in which the weight for each forecast is the inverse of its MSE relative to the sum of MSE for all the models (Bates and Granger 1969; Newbold and Granger 1974),

(9) $$w_t^i = \frac{MSE_i^{-1}}{\sum_{i=1}^{k} MSE_i^{-1}}$$

Gupta and Wilton (1987, 1988) propose a non-parametric **odds matrix approach** based on a matrix of pair-wise odds ratios to incorporate prior forecast accuracy. Specifically, let π_{ij} be the probability that forecast model i will outperform forecast j in the next realization. Then, each element of the matrix O, $o_{ij} = \pi_{ij}/\pi_{ji}$, can be viewed as the odds that forecast *i* will outperform forecast j. Combination weights are the solution to the system of equations,

(10) $$(O - KI)w = 0$$

where I is the identity matrix and w is the weight vector. The solution (the estimated weight vector) is given by the eigenvector corresponding to the largest positive eigenvalue (that solves $Ow = \tau_{max^w}$) which when normalized is unique. To estimate the pair-wise probabilities π_{ij} we follow Gupta and Witlon's (1987) empirical application,

(11) $$\pi_{ij} = \frac{a_{ij}}{a_{ij} + a_{ji}}$$

where a_{ij} is the number of times forecast *i* had a smaller absolute error than forecast j in the past. The approach has been shown to be superior to many methods particularly in small samples.

Shrinkage methods are a set of combination approaches widely and rather successfully used in the macroeconomic literature. Shrinkage methods offer a trade-off between the bias in averaging weights and parameter error when estimating weights (Timmermann 2006). Probably the most common approach shrinks towards equal-weights or the average of forecasts (Stock and Watson 2004, Diebold and Pauly 1990). Let $\hat{w}_t^i$ be the least-squares weight estimator for model *i* from regression (6) up to period *t*. Stock and Watson's (2004) combination weights take the form,

$$\tilde{w}_t^i = \varphi_t \hat{w}_t^i + (1 - \varphi_t)(\frac{1}{k})$$

(12) $$\varphi_t = \max(0, 1 - \frac{\theta k}{N - 1 - k - 1})$$

where θ is the shrinkage parameter. For fixed values of *K* and *N*, as the values of θ become larger (lower φ_t), the greater the shrinkage towards equal-weights. As the sample size *N* increases relative to the number of forecasts *K*, more weight is given to the least-squares estimate. Following Capistran and Timmermann (2009) we consider two values for the θ shrinkage parameter, 0.25 and 1, to assess forecast performance.

Finally, for composite techniques that require historical data for each forecast, weights are recursively estimated using two procedures. The first uses all the data available to the time of the prediction. Weights also are estimated using a fixed-rolling window of 60 observations to allow estimated weights to reflect only information in the most recent 15-year period.

Data and Estimation Procedures

The data are divided into three periods: initial fitting, weight determination, and final assessment. Hence, in second period, we develop the composite weights that are used to forecast in the last period. An exception to this process is the equally-weighted composite that simply averages the forecasts for each period.

Quarterly data for outlook forecasts, hog prices, and related variables are available for 1975.I-2007.IV. Following Ashley (2003), a sample size of at least 50 observations for the out-of-sample evaluation was first specified at the end of the data. Next, a composite identification period of 32 observations was used to establish the initial weights. The remaining data were used for initial fitting of the models. Chronologically, models are initially specified and estimated for the 1975.I-1984.IV sub-period. Individual model forecasts are generated recursively (by adding the next observation to the estimation window for each forecast date) for 1985.I-1993.IV sub-period and used to estimate the combination weights. For the 1994.I-2007.IV period, out-of-sample forecasts for the individual models and composite are also estimated in which the weights are allowed to change at each forecast based on procedures specified. In the last period, composite forecasts are computed recursively using all data, and by a rolling window that specifies a fixed number of observations. The performance of all composite forecasts is compared to the performance of the individual forecasts for the 1994.I to 2007.IV period.

As an example of the estimation procedure, univariate and VAR models are initially specified and estimated over 1975.I to 1984.IV. Next, their out-of-sample forecasts are generated recursively for the period 1985.I-1993.IV. Combined with the ex-ante outlook and futures forecasts for the same period, initial optimal composite weights are estimated. These weights are used for the first out-of-sample composite forecast of 1994.I. For the 1994.II composite forecast, weights are re-estimated using data from 1985.I to 1994.I. The process is repeated period-by-period adding the latest observation and generating a forecast for the subsequent period. Once the combined number of out-of-sample forecasts reaches 60 observations (beginning in 1985.I), weights are re-estimated both recursively or by fixing rolling window of 60 observations. One-, two-, and three-quarter-ahead price forecasts are analyzed for all outlook programs except the USDA, where availability is limited to one- and two-quarter-ahead forecasts.

Results

Overall, the forecast combination analysis applied to the four outlook programs yielded strong similarities. For brevity, we focus the discussion primarily on forecasts released by Iowa State University and identify important similarities and differences found for the other outlook programs.

Descriptive Analysis

Prior to the composite analysis, it is useful to examine the accuracy of the individual forecast models. Table 1 shows the most accurate individual forecasts and the frequency of their superiority for the four programs at different horizons. For each quarter in the 1994.I-2007.IV out-of-sample period, the performance of the four individual models (outlook, futures, AR(5), and VAR(5)) is ranked based on historical MSE recursively computed from 1985.I. At one-quarter ahead, futures forecasts are always the most accurate individual model, except for the USDA where the VAR(5) is the best model. At more distant horizons, the VAR(5) is most accurate. Futures market performs quite well for short-term forecasts, but at more distant horizons its superiority declines and an unrestricted VAR(5) outperforms the other individual forecasts consistently across programs.

RMSE Analysis

Table 2 presents the RMSE over 1994.I-2007.IV for the individual and composite models compared to the Iowa outlook forecasts. RMSE for a price forecast at a given horizon is computed as

$$RMSE = \left[\frac{1}{N} \sum_{t=1}^{N} (P_t - f_t)^2 \right]^{1/2} \tag{13}$$

where, as before, P_t is the actual cash price in quarter *t*, f_t is the forecast price evaluated for quarter *t* and *N* is the number of forecast observations. For each forecast horizon the three forecast models with the smallest RMSE are in bold font. When compared individually, Iowa forecasts are generally superior to those from the time series models.[2] Iowa outperforms the AR(5) and VAR(5) models, on average, by 10.98% and 3.73% respectively, across horizons. However, futures forecasts beat Iowa at the first and second horizon by almost 13% on average. For the first and second horizons, when Iowa is combined with other forecasts significant improvements emerge. On average, Iowa RMSE is reduced -18.23%, -12.17%, and -3.19% at one-, two-, and three-quarter ahead through combining.

To analyze whether differences in RMSEs are statistically different the modified Diebold-Mariano (MDM) test proposed by Harvey, Leybourne, and Newbold (1997) is computed. The MDM statistic tests the null hypothesis of equality of forecast performance between forecast 1 and forecast 2 based on a specified loss function, $E\big(g(e_{1t}) - g(e_{2t})\big) = 0$. Assuming a quadratic loss function, the test is based on the difference in squared errors for two forecasts at a given horizon,

$$d_t = g(e_{1t}) - g(e_{2t}) = e_{1t}^2 - e_{2t}^2 \tag{14}$$

The MDM test is then specified as

$$\begin{gathered} MDM = \left[\frac{N + 1 - 2h + N^{-1}h(h-1)}{N} \right]^{1/2} \left[V(\bar{d}) \right]^{-1/2} \left[\bar{d} \right] \\ V(\bar{d}) = \left[N^{-1} \left(\gamma_0 + 2\sum_{s=1}^{h-1} \gamma_s \right) \right] \end{gathered} \tag{15}$$

where $\bar{d}$ is the sample mean d_t, *h=1,2,3* is the forecast horizon, $\gamma_0 = N^{-1}\sum_{t=1}^{N}(d_t - \bar{d})^2$ is the variance of d_t, and $\gamma_s = N^{-1}\sum_{t=s+1}^{N}(d_t - \bar{d})(d_{t-s} - \bar{d})$ is the s^{th} auto-covariance of d_t, *(k=1,…h-1).* Auto-covariance terms are included to account for the overlap in two- and three-quarter ahead forecasts. The MDM test statistic follows a *t*-distribution with *n-1* degrees of freedom.

Results from the MDM test applied to Iowa for one-, two-, and three-quarter horizons are presented in Table 3.[3] The MDM test of significance is computed for each pair of forecast comparisons to analyze not only the statistical improvement of forecast combination relative to outlook but also to assess meaningful differences between alternative composite approaches. The non-empty cells in the upper part of a matrix correspond to those comparisons between the particular row- and column-forecasts that are statistically different from zero. The number in the parenthesis identifies the forecast with smaller RMSE. Results show that at one-quarter ahead, futures forecasts (#2), all composite forecasts (#6 - #22), and even the naïve approach of following the best previous model (#5) provide statistically smaller RMSEs than those from Iowa alone. At more distant horizons, the attractiveness of using futures and composite forecasts decrease with the exception of the equal-weighted combination (#6). Notably the equal-weight combination shows a significant superiority over Iowa at all forecast horizons. The equal-weighted approach, the composite forecasts based on the odds matrix and those based on the historical MSE show a statistically significant superiority that tends to increase at longer horizons.

To better understand the successful performance of forecast combination, the evolution of Iowa forecast errors over time is compared to those from the equal-weighted composite approach and the constrained-OLS regression. The constrained-OLS regression is of interest not only because of its connection to encompassing tests, but also because it reflects those approaches that require the estimation of error covariances. In contrast, the equal weighted composite reflects those combinatory techniques that ignore any correlation between individual forecasts. Plots of these forecast errors at one-, two-, and three-quarter ahead are reported in Figure 1. While all errors move closely together through time, it is possible to see how forecast errors of the equal-weighted composite approach tend to fall inside the range of the outlook and the OLS-regression errors. Equal-weighted composite forecast errors tend to be less variable and more precise estimations. This tendency becomes even more evident at two- and three-quarter ahead.[4]

Examination of the RMSEs for the other programs reveals several salient points (Tables 4, 5, and 6). First, futures are superior to the other three outlook forecasts at all horizons. Second, while a variety of composite forecasts have significantly smaller RMSEs than the other outlook forecasts, the equal-weighted approach is always smaller than outlook forecasts and significantly outperforms all outlook forecasts at all forecast horizons except for Missouri at one- and three-quarters ahead. Third, composites generated using OLS regressions perform well in several programs particularly at the first and second horizons (USDA, Purdue/Illinois, Missouri—2nd quarter), but their ability declines at the most distant horizon where the forecast errors are largest for all procedures. Added estimation error during periods of high variability may explain the decline in forecast ability relative to equal-weighting, inverse MSE, and posterior odds. These last results are consistent with previous findings (Clemen 1989; Makridakis and Hibon 2000; Stock and Watson 2004, Capistran and Timmerman 2009), which point to the use of an equal-weighted composite procedure. Overall, results suggest that most of the forecast combination models can significantly improve the performance of outlook forecasts in isolation with an average percentage reduction in RMSEs across all four programs of -19.43%, -19.22%, and -5.45% across horizons.

Table 7 summarizes the benefits of pooling the four individual hog price forecasts using an equal-weighted approach, which exhibited relative superiority among the forecast combination approaches. The table presents the percentage RMSE reductions (increments) obtained from the combinations relative to futures forecasts alone and relative to outlook forecasts alone. Percentage RMSE changes from futures forecasts are also of interest since they are usually considered the "gold standard" for comparison in agricultural price forecasting. At one-quarter ahead, the average composite forecast does not provide smaller RMSEs than the futures in isolation, except for the USDA outlook forecast. Excluding the USDA, average forecast errors across the three other programs are 5.41% larger than those from futures. Average equal-weighted forecasts do a better job at two- and three-quarter ahead. RMSE reductions relative to futures forecasts average -5.89% across outlook programs and horizons. The usefulness of the equal-weighted forecasts is more noticeable when outlook forecasts are the benchmark. Regardless of the outlook program, the equal-weighted composite forecast consistently provides smaller RMSEs than those obtained from outlook alone. Across programs, the average forecast error reductions are -16.39%, -18.17%, and -7.21% at one-, two-, and three-quarter ahead, respectively. Within programs, the largest reduction is obtained for the USDA, while Missouri receives the minimum benefits from the combinations. On average, the equal-weighted composite reduces the Missouri, Illinois/Purdue, Iowa, and the USDA forecast errors by -9%, -11.41%, -15.66%, and -19.62%, respectively.

Summary and Conclusions

This study analyzes whether the accuracy of outlook hog price forecasts can be improved through composite forecasts in a realistic out-of-sample context. Data are divided into three periods for individual model fitting, composite forecast training, and final evaluation. Price forecasts from four-well recognized outlook programs are combined with futures-based forecasts, an ARIMA model, and an unrestricted Vector Autoregressive (VAR) model using alternative combining techniques.

For the out-of-sample 1994.I-2007.IV period futures outperform outlook forecasts at most horizons. Based on the MDM test of significance in RMSE differences, futures is superior to outlook in 5 out of 11 cases. A variety of composite procedures also provide smaller RMSEs than outlook forecasts with numerous statistically significant differences. Performance of the futures and composite forecasts decreases at longer horizons, with the exception of the equal-weighted composite. The equal-weighted composite always has a smaller RMSE than outlook forecasts and is significantly superior in 9 out of 11 cases. On average, the equal-weighted composite forecast reduces outlook RMSEs by -16.39%, -18.17%, and -7.21% at first, second and third horizon. While it is difficult for the equal-weighted composite forecast to beat futures at the first horizon, it reduces futures errors by an average of -5.89% at the second and third horizons, and is statistically smaller in 2 cases. Finally, methods including the equal-weighted approach, the odds matrix, and the composite based on the historical MSE that do not require estimation of OLS regressions to generate the weights are significantly superior at the most distant horizon.

Overall, results favor the use of composite forecasting methods to reduce outlook programs forecast errors. For short-term forecasting, it is difficult to outperform futures forecasts, but composite procedures perform quite well at more distant horizons in terms of forecast-error reduction. Our evidence on the trade-off between bias and efficiency suggests that the losses to inefficiency outweigh potential bias. At most distant horizons where predictive accuracy is more problematic, forecasts generated by OLS regressions do not perform as well as composite forecasts generated by other less-statistical procedures. These results are of value to decision makers in agricultural markets. In contrast to much of the previous research in agricultural markets, and particularly within the context of outlook forecasts, our findings establish solidly the effectiveness of futures and composite methods in a rather realistic context. Further, the forecast models and the composite methods used are easily implementable and require only minimal upkeep to provide reductions in forecast errors.

References

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

Ashley, R. 2003. "Statistically Significant Forecasting Improvements: How Much Out-of-Sample Data is Likely Necessary?" *International Journal of Forecasting* 19:229-239.

Bates, J.M. and C.W.J. Granger. 1969. "The Combination of Forecasts." *Operational Research Quarterly* 20: 451-468.

Bessler, D.A. and P.J. Chamberlain. 1988. "Composite Forecasting with Dirichlet Priors." *Decision Sciences* 19: 771-781

Capistran, C. and A. Timmermann. 2009. "Forecast Combination With Entry and Exit of Experts." *Journal of Business and Economic Statistics (forthcoming)*

Clark, T.E., and M.W. McCracken. 2004. "Improving Forecast Accuracy by Combining Recursive and Rolling Forecasts." Research Division, Research Working Paper 04-10, Federal Reserve Bank of Kansas City.

Clark, T.E., and M.W. McCracken. 2006. "Averaging Forecasts from VARs with Uncertain Instabilities." Research Working Paper 06-12, Federal Reserve Bank of Kansas City, Economic Research Department.

Clemen, R.T. 1989. "Combining Forecasts: A Review and Annotated Bibliography." *International Journal of Forecasting* 5: 559-583.

Colino, E.V., and S.H. Irwin. 2007. "Outlook vs. Futures: Three Decades of Evidence in Hog and Cattle Markets." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, IL.

Colino, E.V., S.H. Irwin, and P. Garcia. 2008. "How Much Can Outlook Forecasts be Improved? An Application to the U.S. Hog Market." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO.

Diebold, F. X. and J.A. Lopez. 1995. "Forecast Evaluation and Combination." *Research Paper No. 9525, Federal Reserve Bank of New York.*

Diebold, F.X. and P. Pauly. 1990. "The Use of Prior Information in Forecast Combination." *International Journal of Forecasting* 6: 503-508.

Fang, Y. 2003. "Forecasting Combination and Encompassing Tests." *International Journal of Forecasting* 19: 87-94.

Garcia, P., and D. R. Sanders. 1996. "Ex Ante Basis Risk in the Live Hog Futures Contract: Has Hedgers' Risk Increased?" *Journal of Futures Markets* 16:421-440.

Granger, C.W.J., and P. Newbold. 1986. *Forecasting Economic Time Series, Second Edition.* San Diego, CA: Academic Press, Inc.

Granger, C.W.J. and R. Ramanathan. 1984. "Improved Methods of Combining Forecasts." *Journal of Forecasting* 3:197-204.

Gupta, S. and Wilton, P.C. 1987. "Combination of Forecasts: An Extension." *Management Science* 33: 356-372.

Gupta, S. and Wilton, P.C. 1988. "Combination of Economic Forecasts: An Odds-Matrix Approach." *Journal of Business and Economic Statistics* 6: 373-379.

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

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

Henriksson, R.D. and R.C. Merton. 1981. "On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills." *Journal of Business* 54: 513-533.

Hoffman, L.A. 2005. "Forecasting the Counter-Cyclical Payment Rate for U.S. Corn: An Application of the Futures Price Forecasting Model." U.S. Department of Agriculture, Economic Research Service, Electronic Outlook Report No. FDS05a01. [http://www.ers.usda.gov/publications/FDS/JAN05/fds05a01/fds05a01.pdf]

Issler, J.V. and L.R. Lima. 2009. "A Panel Data Approach to Economic Forecasting: The Bias-Corrected Average Forecast." *Journal of Econometrics (forthcoming)*

Makridakis, S. and M. Hibon. 2000. "The M3-Competition: Results, Conclusions and Implications." *International Journal of Forecasting* 16:451-476.

McIntosh, C.S. and D.A. Bessler. 1988. "Forecasting Agricultural Prices using a Bayesian Composite Approach." *Southern Journal of Agricultural Economics* 20:73-80.

Nelson, C.R. 1972. "The Prediction Performance of the FRB-MIT-PENN Model of the U.S. Economy." *The American Economic Review* 62: 902-917.

Newbold, P. and C.W.J. Granger. 1974. "Experience with Forecasting Univariate Time Series and the Combination of Forecasts." *Journal of the Royal Statistic Society* 137: 131-146.

Smith, J. and K.F. Wallis. 2009. "A Simple Explanation of the Forecast Combination Puzzle." *Oxford Bulletin of Economics and Statistics* 71: 331-355.

Stock, J.H. and M.W. Watson. 2004. "Combination Forecasts of Output Growth in a Seven-Country Data Set." *Journal of Forecasting* 23:405-430.

Sutton, R.W., and Albrecht, J.E. 1996. "An Explanation of Hog Futures Changes." Management Marketing Memo #333, Department of Agricultural and Applied Economics, Clemson University. [http://cherokee.agecon.clemson.edu/mmm343.htm]

Timmerman, A. 2006. "Forecast Combinations." *Handbook of Economic Forecasting, Volume 1*. G. Elliot, C.W.J. Granger, and A. Timmerman eds., Amsterdam, NA: North-Holland.

Endnotes

[1] An estimated ratio of 0.73673 is applied to lean-hog futures prices. This factor is obtained by dividing the average weight of lean hogs (180.5) by the average weight of live hogs (245) (Sutton and Albrecht 1996). The adjustment is necessary because the Chicago Mercantile Exchange shifted the hog contract delivery terms from a live weight to carcass weight basis beginning with the February 1997 contract.

[2] Note the rankings here may deviate from those presented in Table 1 which are based on historical MSEs beginning in 1985.I. The RMSEs presented in Table 2 (and Tables 4, 5, and 6) are limited to the 1994.I-2007.IV period.

[3] MDM test results for the other three outlook forecasts are available from the authors.

[4] In addition, the Henriksson and Merton (1981) test of directional accuracy was performed. Results show that, with only four exceptions, all price forecasts evaluated, across outlook programs and horizons, reject the null of no timing ability implying they all have value for decision makers. In terms of the percentage of directionally correct forecasts, outlook forecasts, except for Missouri, are outperformed by all individual and composite forecasts. Notably, the simple average has a correct direction in its predictions 76.4% of the time, which is a stronger performance compared to the average of all composite forecasts of 74.9% across programs and horizons.

Table 1. Best individual forecasts in recursive out-of-sample evaluations, 1994.I-2007.IV

	Forecast horizon		
Outlook Program	**1-qtr.-ahead**	**2-qtr.-ahead**	**3-qtr.-ahead**
Univ. of Illinois/Purdue Univ.	Futures - (100%)	VAR(5) - (100%)	VAR(5) - (88.5%)
Iowa State University	Futures - (100%)	VAR(5) - (100%)	VAR(5) - (100%)
University of Missouri	Futures - (100%)	VAR(5) - (100%)	VAR(5) - (60%)
USDA	VAR(5) - (100%)	VAR(5) - (95%)	

Notes: Each cell shows the individual forecast (outlook, futures, AR(5), and VAR(5)) model that is most frequently ranked first in terms of forecast accuracy in the out-of-sample period. Numbers in parentheses represent the frequency of superiority over time.

Table 2. RMSE of hog price forecasts during the out-of-sample evaluation period, Iowa St. University, 1994.I-2007.IV

	Forecast model	Forecast horizon					
		1-qtr.-ahead		2-qtr.-ahead		3-qtr.-ahead	
#1	Iowa State University	4.51		6.42		7.40	
#2	Futures	**3.55**	**	6.12		7.56	
#3	VAR(5)	4.80		6.25		7.94	
#4	ARMA(5,0)	5.24		7.10		7.84	
#5	Best previous model	**3.55**	**	6.25		7.94	
#6	equal-weight composite	3.87	*	5.61	**	**6.84**	*
#7	MSE-weight composite	3.77	**	5.71	*	6.88	
#8	MSE-weight composite - rolling window	3.77	**	5.71	**	6.89	
#9	Unrestricted OLS composite	3.65	**	5.54		7.33	
#10	Unrestricted OLS composite - rolling window	3.70	**	**5.45**		7.30	
#11	Unrestricted OLS composite -noconstant	3.58	**	5.71		7.47	
#12	Unrestricted OLS composite -noconstant - rolling window	3.62	**	5.67		7.54	
#13	Restricted OLS composite	**3.55**	**	5.74		7.35	
#14	Restricted OLS composite - rolling window	3.59	**	5.72		7.34	
#15	Projection on the equal-weight composite	3.84	*	**5.40**	**	6.89	
#16	Projection on the equal-weight composite - rolling window	3.84	*	**5.42**	**	6.98	
#17	Shrinkage - 0.25 composite	3.56	**	5.74		7.34	
#18	Shrinkage - 0.25 composite - rolling window	3.60	**	5.71		7.33	
#19	Shrinkage - 1 composite	3.56	**	5.75		7.28	
#20	Shrinkage - 1 composite - rolling window	3.60	**	5.69		7.27	
#21	Odds matrix composite	3.77	**	5.62	**	**6.87**	
#22	Odds matrix composite - rolling window	3.76	**	5.62	**	**6.87**	

Notes: All figures are reported as $/cwt. One, two, and three stars indicate statistical significance in the RMSE differences between outlook forecast and the alternative forecast model at the 10%, 5% and 1% level, respectively, based on the Modified Diebold-Mariano (MDM) test.

Table 3. Modified Diebold-Mariano (MDM) test results, Iowa St. University, 1994.I-2007.IV

Forecast model	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22
1-quarter Ahead																						
#1		(2) **			(5)**	(6)*	(7)**	(8)**	(9)**	(10)**	(11)**	(12)**	(13)**	(14)**	(15)*	(16)*	(17)**	(18)**	(19)**	(20)**	(21)**	(22)**
#2			(2)***	(2)***																		
#3					(5)***	(6)***	(7)***	(8)***	(9)***	(10)***	(11)***	(12)***	(13)***	(14)***	(15)**	(16)**	(17)***	(18)***	(19)***	(20)***	(21)***	(22)***
#4					(5)***	(6)***	(7)***	(8)***	(9)***	(10)***	(11)***	(12)***	(13)***	(14)***	(15)***	(16)***	(17)***	(18)***	(19)***	(20)***	(21)***	(22)***
#6							(7)**	(8)**					(13)*				(17)*		(19)*		(21)***	(22)***
#13																				(13)*		
#17																				(17)*		
2-quarter Ahead																						
#1						(6)**	(7)*	(8)**							(15)**	(16)**					(21)**	(22)**
#2				(2)*											(15)**	(16)**						
#3				(3)**	(5)***							(12)*	(13)*	(14)*			(17)*	(18)*	(19)*	(20)*		
#4					(5)**	(6)***	(7)***	(8)***	(9)***	(10)***	(11)***	(12)***	(13)***	(14)***	(15)***	(16)***	(17)***	(18)***	(19)***	(20)***	(21)***	(22)***
#5												(12)*	(13)*	(14)*			(17)*	(18)*	(19)*	(20)*		
#6							(6)**	(6)**														
#7																					(21)**	(22)**
#8																					(21)**	(22)**
#9										(10)**												
3-quarter Ahead																						
#1						(6)*																
#2						(6)*																
#3						(6)**	(7)**	(8)**			(11)*		(13)**	(14)*	(15)*		(17)**	(18)*	(19)**	(20)**	(21)**	(22)**
#4						(6)**	(7)**	(8)**							(15)*						(21)**	(22)**
#5						(6)**	(7)**	(8)**			(11)*		(13)**	(14)*	(15)*		(17)**	(18)*	(19)**	(20)**	(21)**	(22)**
#6											(6)**	(6)***	(6)**	(6)**			(6)**	(6)**	(6)**	(6)**		(6)*
#7											(7)**	(7)***	(7)**	(7)**			(7)**	(7)**	(7)**	(7)**		
#8											(8)**	(8)***	(8)**	(8)**			(8)**	(8)**	(8)**	(8)**		
#11																					(21)**	(22)**
#12																		(18)*	(19)**	(20)**	(21)**	(22)**
#13																	(17)***		(19)***		(21)**	(22)**
#14																		(18)***		(20)***	(21)**	(22)**
#15																(15)**						
#17																			(19)***		(21)**	(22)**
#18																				(20)***	(21)**	(22)**
#19																					(21)*	(22)*
#20																					(21)**	(22)*

Notes: At each cell, the number in parenthesis represents the superior forecast model. One, two, and three stars indicate statistical significance at the 10%, 5% and 1% level, respectively, based on the MDM test. No statistically significant differences between forecasts were found for the models not included.

Table 4. RMSE of hog price forecasts during the out-of-sample evaluation period, Univ. of Illinois/Purdue Univ., 1994.I-2007.IV

		Forecast horizon		
	Forecast model	**1-qtr.-ahead**	**2-qtr.-ahead**	**3-qtr.-ahead**
#1	Univ. of Illinois/Purdue Univ.	4.59	6.87	7.59
#2	Futures	3.79	5.75 **	7.07
#3	VAR(5)	4.66	5.97	7.55
#4	ARMA(5,0)	5.06	6.76	7.33
#5	Best previous model	3.79	5.97	7.65
#6	Equal-weight composite	3.90 ***	5.59 ***	**6.58** **
#7	MSE-weight composite	3.77 ***	5.59 ***	**6.58** **
#8	MSE-weight composite - rolling window	3.77 ***	5.58 ***	**6.58** **
#9	Unrestricted OLS composite	3.72 *	**5.35** **	6.95
#10	Unrestricted OLS composite - rolling window	3.76	**5.35** **	7.00
#11	Unrestricted OLS composite - noconstant	**3.63** **	**5.32** ***	7.01
#12	Unrestricted OLS composite - noconstant - rolling window	**3.67** **	5.37 **	7.22
#13	Restricted OLS composite	3.73 **	5.49 ***	7.01
#14	Restricted OLS composite - rolling window	3.74 **	5.54 **	7.05
#15	Projection on the equal-weight composite	3.94 *	5.70 **	6.79 *
#16	Projection on the equal-weight composite - rolling window	3.96 *	5.76 **	6.89 *
#17	Shrinkage - 0.25 composite	3.72 **	5.49 ***	6.99
#18	Shrinkage - 0.25 composite - rolling window	3.75 **	5.54 **	7.04
#19	Shrinkage - 1 composite	**3.71** **	5.48 ***	6.95
#20	Shrinkage - 1 composite - rolling window	3.74 **	5.53 ***	6.99
#21	Odds matrix composite	3.79 ***	5.54 ***	6.58 **
#22	Odds matrix composite - rolling window	3.79 ***	5.53 ***	6.60 **

Notes: All figures are reported as $/cwt. One, two, and three stars indicate statistical significance in the RMSE differences between outlook forecast and the alternative forecast model at the 10%, 5% and 1% level, respectively, based on the Modified Diebold-Mariano (MDM) test.

Table 5. RMSE of hog price forecasts during the out-of-sample evaluation period, University of Missouri, 1994.I-2007.IV

		Forecast horizon		
	Forecast model	**1-qtr.-ahead**	**2-qtr.-ahead**	**3-qtr.-ahead**
#1	University of Missouri	3.99	6.34	7.13
#2	Futures	3.60	5.61 **	7.08
#3	VAR(5)	4.61	6.17	7.86
#4	ARMA(5,0)	5.12 ***	7.13	7.90
#5	Best previous model	3.60	6.17	8.18 *
#6	Equal-weight composite	3.76	5.51 ***	**6.56**
#7	MSE-weight composite	3.63	5.58 **	6.60
#8	MSE-weight composite - rolling window	3.64	5.58 **	6.60
#9	Unrestricted OLS composite	3.55	**5.23**	6.99
#10	Unrestricted OLS composite - rolling window	3.53	**5.10** *	6.90
#11	Unrestricted OLS composite - noconstant	**3.51**	5.31 *	7.14
#12	Unrestricted OLS composite - noconstant - rolling window	**3.49**	**5.25** **	7.10
#13	Restricted OLS composite	3.55	5.46 *	7.12
#14	Restricted OLS composite - rolling window	3.52	5.43 *	7.10
#15	Projection on the equal-weight composite	3.57	5.35 **	**6.57**
#16	Projection on the equal-weight composite - rolling window	3.58	5.37 **	**6.60**
#17	Shrinkage - 0.25 composite	3.55	5.45 *	7.10
#18	Shrinkage - 0.25 composite - rolling window	3.52	5.42 *	7.08
#19	Shrinkage - 1 composite	3.54	5.45 *	7.03
#20	Shrinkage - 1 composite - rolling window	**3.52**	5.41 **	7.01
#21	Odds matrix composite	3.68	5.55 **	6.61
#22	Odds matrix composite - rolling window	3.68	5.54 **	6.62

Notes: All figures are reported as $/cwt. One, two, and three stars indicate statistical significance in the RMSE differences between outlook forecast and the alternative forecast model at the 10%, 5% and 1% level, respectively, based on the Modified Diebold-Mariano (MDM) test.

Table 6. RMSE of hog price forecasts during the out-of-sample evaluation period, USDA, 1994.I-2007.IV

		Forecast horizon	
	Forecast model	**1-qtr.-ahead**	**2-qtr.-ahead**
#1	USDA	6.27	8.03
#2	Futures	4.82 ***	6.42 ***
#3	VAR(5)	4.55 **	5.80 ***
#4	ARMA(5,0)	4.90 **	6.52 *
#5	Best previous model	4.55 **	6.02 **
#6	Equal-weight composite	4.36 ***	5.75 ***
#7	MSE-weight composite	**4.28** ***	5.61 ***
#8	MSE-weight composite - rolling window	4.29 ***	5.61 ***
#9	Unrestricted OLS composite	4.29 ***	**5.41** ***
#10	Unrestricted OLS composite - rolling window	4.30 ***	**5.43** ***
#11	Unrestricted OLS composite - noconstant	4.30 ***	**5.52** ***
#12	Unrestricted OLS composite - noconstant - rolling window	4.32 ***	5.55 ***
#13	Restricted OLS composite	4.31 ***	5.57 ***
#14	Restricted OLS composite - rolling window	4.35 ***	5.60 ***
#15	Projection on the equal-weight composite	4.42 ***	5.77 ***
#16	Projection on the equal-weight composite - rolling window	4.44 ***	5.78 ***
#17	Shrinkage - 0.25 composite	4.31 ***	5.56 ***
#18	Shrinkage - 0.25 composite - rolling window	4.35 ***	5.60 ***
#19	Shrinkage - 1 composite	4.29 ***	5.55 ***
#20	Shrinkage - 1 composite - rolling window	4.33 ***	5.58 ***
#21	Odds matrix composite	**4.27** ***	5.57 ***
#22	Odds matrix composite - rolling window	**4.27** ***	5.58 ***

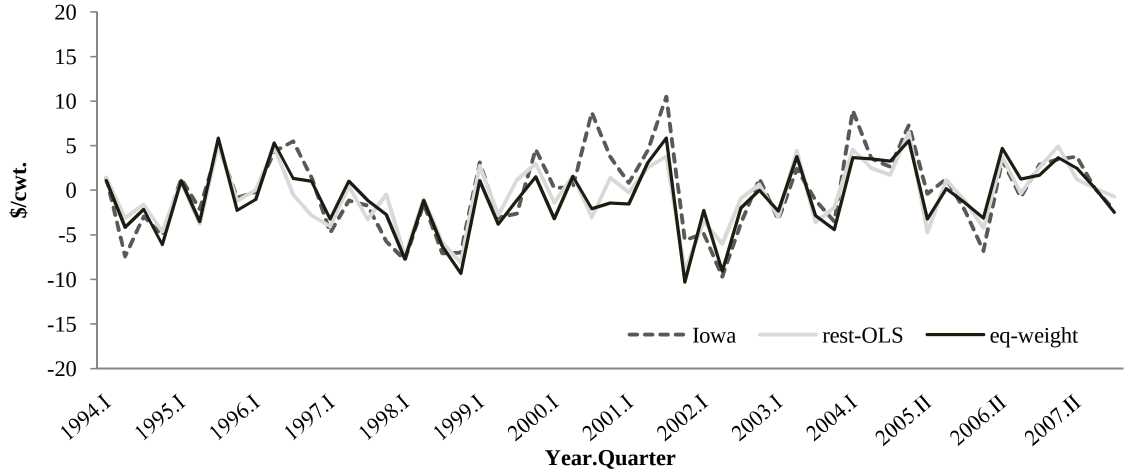
Notes: All figures are reported as $/cwt. One, two, and three stars indicate statistical significance in the RMSE differences between outlook forecast and the alternative forecast model at the 10%, 5% and 1% level, respectively, based on the Modified Diebold-Mariano (MDM) test.

Table 7. Summary of RMSE reductions, 1994.I-2007.IV

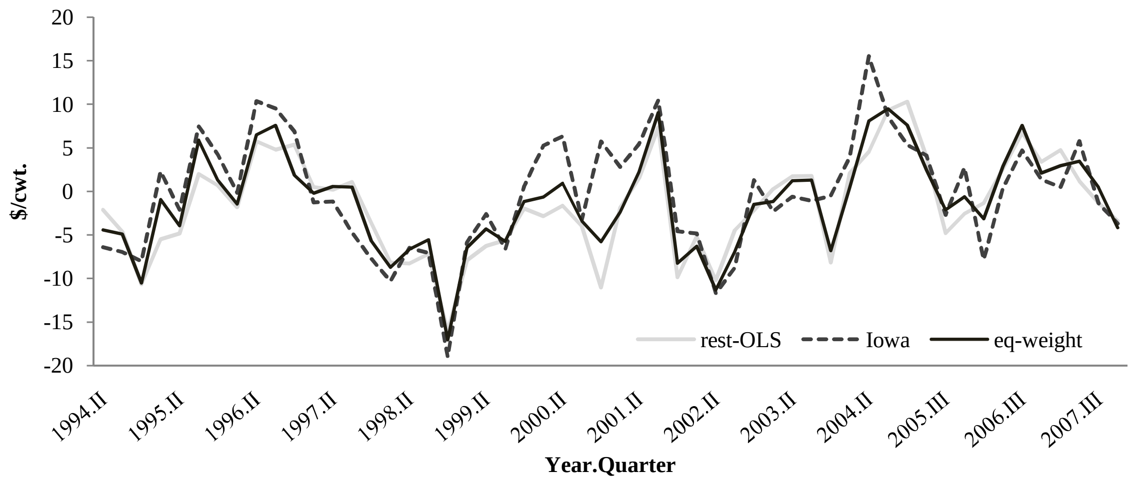
	1-qtr.-ahead			2-qtr.-ahead			3-qtr.-ahead		
	RMSE	RMSE reduction from futures	RMSE reduction from outlook	RMSE	RMSE reduction from futures	RMSE reduction from outlook	RMSE	RMSE reduction from futures	RMSE reduction from outlook
Univ. of Illinois/Purdue Univ.	4.59			6.87			7.59		
Futures	3.79			5.75			7.07		
equal-weight composite forecast	3.90	2.72%	-15.16%	5.59	-2.65%	-18.58%	6.58	-6.92%	-13.25%
Iowa State University	4.51			6.42			7.40		
Futures	3.55			6.12			7.56		
equal-weight composite forecast	3.87	9.19%	-14.04%	5.61	-8.33%	-12.59%	6.84	-9.59%	-7.61%
University of Missouri	3.99			6.34			7.13		
Futures	3.60			5.61			7.08		
equal-weight composite forecast	3.76	4.31%	-5.90%	5.51	-1.94%	-13.13%	6.56	-7.28%	-7.97%
USDA	6.27			8.03					
Futures	4.82			6.42					
equal-weight composite forecast	4.36	-9.45%	-30.46%	5.75	-10.38%	-28.41%			

Notes: RMSEs are reported as $/cwt.

Panel A. One-quarter ahead



Panel B. Two-quarters ahead



Panel C. Three-quarters ahead

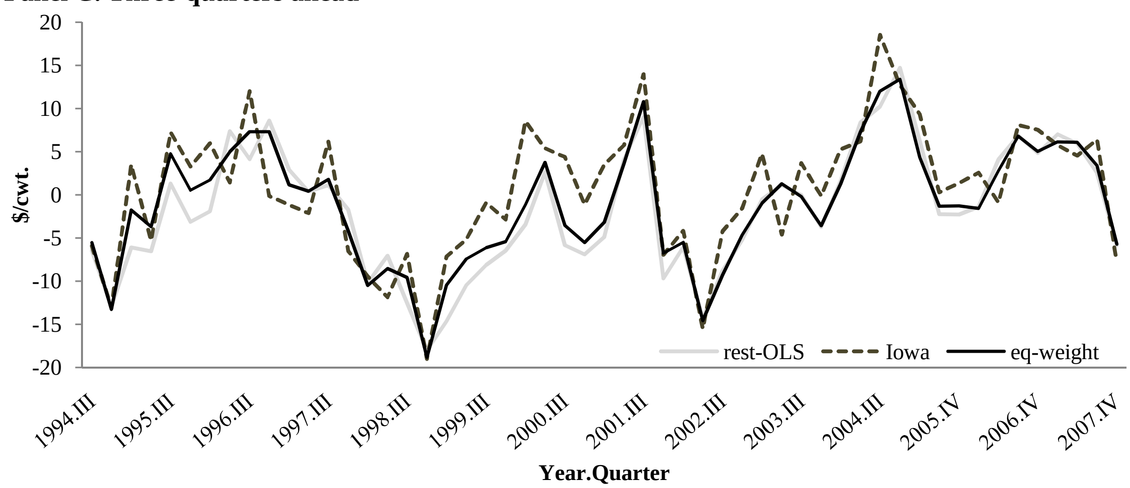


Figure 1. Out-of-sample forecast errors for Iowa State University, equal-weight, and constrained-OLS composite approaches, one-, two-, and three-qtr. ahead ($/cwt.), 1994.I-2007.IV