NCCC-134
APPLIED COMMODITY PRICE ANALYSIS, FORECASTING AND MARKET RISK MANAGEMENT

Are Composite Forecasts More Accurate?
An Application to Livestock and Poultry

by

Richard Stillman, Mark Weimar, and Kenneth Nelson

Suggested citation format:

Stillman, R., M. Weimar, and K. Nelson. 1991. "Are Composite Forecasts More Accurate? An Application to Livestock and Poultry." Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, IL. [http://www.farmdoc.uiuc.edu/nccc134].

Are Composite Forecasts more Accurate? an Application to Livestock and Poultry

Richard Stillman, Mark Weimar and Kenneth E Nelson 1/

Introduction

The accuracy of USDA livestock nd poultry forecasts has been under study by the Economic Research Service (ERS), USDA and the Congressional General Accounting Office (GAO) and numerous universities. Some studies indicate bias and preventable errors in the official USDA livestock and poultry forecasts of which ERS analysts are major contributors. This research examines whether a composite of the ERS Annual Livestock Model (ERS-ALM) forecasts combined with time series models and the expert forecasts of the World Agricultural Outlook Board (WAOB), the Interagency Livestock Estimates Committee (IC) could have reduced the bias and improved precision of the forecasts during the 1980's.

Some literature has suggested that composite forecasts (where forecasts from several sources are combined to provide a forecast) perform better than single model or expert forecasts. We compare 1); a composite forecast produced from the ERS-ALM, the IC, and an autoregressive moving average model (ARIMA) and 2) a composite of the ERS-ALM, AR and ARIMA models with both the expert forecasts and forecasts of each model. Also, a naive model forecast will be examined.

We examine the annual one step ahead forecast errors of beef, pork and broiler production and market prices for cattle, hogs and broilers. Each of these forecasts are examined for bias, point accuracy using both a linear and quadratic loss function. The tests utilized will be Mean Error, Mean Absolute Percentage Error, Root Mean Square Percentage Error, and a test to determine the statistical differences between forecasts.

Previous Studies

Economists have long debated ways to improve the accuracy of forecasts. Composite forecasts methodologies have in many cases improved the accuracy of the forecast. Several studies have shown that the composite forecast from different sources provide a single more accurate forecast. Brandt and Bessler examined forecasts of quarterly hog prices over the late 1970's and early 1980's (2). Gerlow and Irvin also examined quarterly hog prices using combination forecasts (4). Both of these found the composite forecasting superior based on a quadratic loss criterion.

Methodology for evaluation of forecasts has also created differences of opinion among economists. Over the years a large number of tests have been devised to examine the accuracy of forecasts. In most cases the evaluation method is based on linear or quadratic loss functions. Brandt and

---

1/ Agricultural economists with the Commodity Economics Division, Economic Research Service, U.S. Department of Agriculture.

Bessler used Mean Squared Errors (MSE) and a statistical test of the differences between these errors [2,1]. Elam and Holder, in their evaluation of USDA analyst forecasts used Mean Errors (ME) to measure bias, Root Mean Square Errors (RMSE) and Mean Absolute Errors (MAE) to test accuracy in their evaluation [3]. Just and Rausser, in their examination of large scale econometric models utilized MSE which they decomposed into the bias, forecast variance and actual variance components [5]. Gerlow and Irvin utilized the MSE of the forecast as an evaluation tool as well as several economic criterion based on the forecasts ability to provide economic returns to the user [4].

Forecast Methods and Evaluation Criteria

Forecast Methods

We investigate several ways to make the USDA WAOB forecasts more accurate and reduce bias. Several models were examined over the period 1981 to 1990 to compare their ability to forecast beef, pork, and chicken production and prices. The forecasting methodologies examined were, the WAOB Interagency Committee (IC) forecast, the ERS Annual Livestock Model (ERS-ALM), an Autoregressive moving average (ARIMA) model, a simple weighted average (SWA1) of the previous forecasts, a simple weighted of the econometric and time series models (SWA2), and a naive forecast. Each of these methods yield a one year ahead forecast based on the information available on approximately January 1 of the year forecasted.

The IC forecasts were available in the January WASDE publication (USDA). These forecasts are based on consensus among analysts from different agencies within the USDA and are essentially composite forecasts using committee members judgement.

The ERS-ALM model was used to make one step ahead forecasts. These forecasts were performed by estimating the model starting with the period 1958 to 1980. The parameter estimates were then used to forecast 1981. The model was then reestimated adding 1981 to the estimation period. The new parameter estimates were then input into the model and used to forecast 1982. This procedure continued until one step ahead forecasts were obtained for the evaluation period of 1981 to 1990.

ARIMA forecasts were obtained using the same methodology applied to the ERS-ALM forecasts. Because of model data requirements, the beginning period of the data used to estimate the ARIMA models began in 1948 as opposed to 1958 which was used in the ERS-ALM estimation. The lag length of the ARIMA models differed over the period of evaluation and was based on the within sample Akaike Information Criterion (AIC) and the Schwartz Information Criterion (SIC)[8]. The relatively short estimation data set required that both the Autoregressive and the Moving Average lags be restricted to 3 or less. These models were estimated for the sample period ending in 1980 and then used to forecast the next year. The models were then reestimated and the next years forecast was obtained.

The composite forecasts were computed as simple weighted averages. The first composite forecast (SWA1) averaged the three statistical model estimates and the IC forecast. The second composite forecast (SWA2) used only the statistical models. The naive forecast was simply the previous years actual value for the variable examined.

Evaluation Criterion

The criteria utilized to measure the accuracy of the forecasting methods are the Mean Error (ME), the Mean Absolute Percentage Error (MAPE), the Root Mean Square Percentage Error (RMSPE), and a test suggested by Davidson and MacKinnon [3].

Mean Error is a measure of bias in the forecast. The Mean Error is defined as the actual value minus the forecasted value and is expected to have a value of zero for an unbiased forecast.

$$ME=\frac{\sum_{i=1}^{n} Actual_i-Forecasted_i}{n}$$

The Mean Absolute Percentage Error measures the point accuracy of the forecast. This criterion is based on a linear loss function and weights all forecast error the same. Expressing this statistic in percentage terms allows an unitless comparison of forecasts among variables. The MAPE is defined as,

$$MAPE=\frac{\sum_{i=1}^{n} |Actual_i-Forecasted_i|}{n}$$

The Root Mean Squared Percentage Error also measures the point accuracy of the forecast. This criterion is based on a quadratic loss function and weights the largest error the greatest. The RMSPE is defined as,

$$RMSPE=\sqrt{\frac{\sum_{i=1}^{n} (Actual_i-Forecasted_i)^2}{n}}$$

Davidson and MacKinnon suggest several methods to test differences between models [3]. One of these tests is called the C test. The C test presents some problems with the actual power of the test. However, it was chosen as a crude test of significance differences among forecasts. Note that the C test is similar to methodology suggested by Granger and Newbold to weight composite forecasts [6]. The C test tests the null hypothesis,

$$H_0: Y_i=f_i(Z_i,\beta)+\epsilon_{0,i}$$

against the alternative hypothesis of the model specification,

$$H_1: Y_t = g_t(Z_t, \gamma) + e_{1,t}$$

The test involves the estimation of,

$$y_t = (1-\alpha)\hat{f}_t + \alpha\hat{g}_t + e_t$$

or,

$$y_t - \hat{f}_t = \alpha(\hat{g}_t - \hat{f}_t) + e_t$$

where,

$$\hat{g}_t = g_t(Z_t, \hat{\gamma})$$

and

$$\hat{f}_t = f_t(X_t, \hat{\beta})$$

One then tests whether $\alpha = 0$ and conversely that $\alpha \neq 1$. Note that the variance of α is bias and the probability of a Type I error is no greater than the size of the test.

Forecast evaluations

Evaluations for 1981 to 1990, were done on Commercial beef, pork, and broiler prices and production. Omaha choice steer price was chosen for the beef price. Seven market barrow and gilt price and 12-city broiler price were used for the pork and broiler prices respectively. Table 1 contains the results of the ME, RMSPE, and MAPE analysis and table 2 contains the tests of significant differences among the forecasts at the 5 percent level.

Commercial Beef Production

Of the beef production forecasts, none indicate bias, with no ME's significantly different from zero. Although most of the methods appeared to under predict the level of beef production. The range of the MAPE's were from 1.729 percent to 4.015 percent none of which are extremely large. A naive

Table 14. Forecast evaluations.

Forecast Method	Mean	Mean Error	Mean Absolute Error	Root Mean Square Error
	million lbs		percent	
Actual Beef Production	23125.20			
IC	22385.00	740.200	3.952	4.253
ERS-ALM	23193.22	-68.020	4.015	5.044
ARIMA	22747.66	377.536	2.326	2.679
Naive	23010.30	114.900	1.892	2.213
SWA1	22756.14	369.064	2.578	3.030
SWA2	22879.85	245.353	2.458	3.033
Actual Pork Production	14938.50			
IC	14635.00	303.500	2.764	4.365
ERS-ALM	14264.49	674.012	4.486	5.261
ARIMA	14666.82	271.684	4.169	5.362
Naive	15052.60	114.100	4.489	5.582
SWA1	14569.10	369.402	3.299	4.287
SWA2	14547.13	391.369	3.631	4.498
Actual Broiler Production	14470.02			
IC	14305.50	164.515	1.534	1.966
ERS-ALM	14013.21	456.805	2.975	3.464
ARIMA	13803.80	666.215	4.419	4.870
Naive	13740.05	729.963	4.851	5.222
SWA1	13944.53	525.483	3.503	3.811
SWA2	13824.21	645.806	4.290	4.598
Actual Omaha Steer Price	65.60			
IC	68.39	-2.791	7.088	9.262
ERS-ALM	73.80	-8.209	11.497	21.792
ARIMA	63.91	1.677	6.008	6.961
Naive	64.55	1.043	5.403	6.458
SWA1	67.35	-1.754	7.454	8.686
SWA2	67.01	-1.409	7.586	9.899
Actual 7 Market Barrow and Gilt Price	48.60			
IC	49.70	-1.097	8.887	11.129
ERS-ALM	53.02	-4.419	12.803	16.740
ARIMA	46.67	1.923	10.967	12.869
Naive	47.17	1.437	11.116	13.221
SWA1	48.97	-0.376	9.720	10.863
SWA2	48.74	-0.137	10.904	11.980
Actual 12 city Broiler Price	52.62			
IC	50.87	1.749	8.985	11.167
ERS-ALM	51.43	1.185	9.726	11.689
ARIMA	51.13	1.481	9.463	10.835
Naive	52.05	0.574	9.107	10.533
SWA1	51.15	1.463	8.845	9.896
SWA2	51.25	1.368	9.139	10.037

Table 2. Tests of significance among forecast methods.

Forecast	Naive	ARIMA	ERS-ALM	WASDE	SWA1
Beef production					
Naive	--				
ARIMA		--			
ERS-ALM	S1/	S	--	S	S
WASDE	S	S		--	S
SWA1	S				--
SWA2	S				
Pork production					
Naive	--				
ARIMA		--			
ERS-ALM			--		
WASDE				--	
SWA1					--
SWA2					
Broiler production					
Naive	--				
ARIMA		--		S	
ERS-ALM			--	S	
WASDE				--	
SWA1				S	--
SWA2				S	
Steer price					
Naive	--				
ARIMA		--			
ERS-ALM	S	S			
WASDE	S	S	--	S	S
SWA1	S	S		--	
SWA2	S	S			--
Barrow and Gilt price					
Naive	--				
ARIMA		--		S	S
ERS-ALM	S	S		S	S
WASDE			--	S	S
SWA1				--	
SWA2					--
Broiler price					
Naive	--				
ARIMA		--			
ERS-ALM			--		
WASDE				--	
SWA1					--
SWA2					

1/ S in a cell indicates that the forecast in the top heading is superior to the forecast in the side heading at the 5 percent confidence level.

forecast had the smallest MAPE for beef production over the evaluation period. The ARIMA had the second lowest MAPE. These were followed in order by the SWA2, SWA1, IC, and ERS-ALM forecasts. RMSPE's for all models exhibited the same ordering as the MAPE's, with the exception of SWA1 being lower than SWA2. The RMSPE's range from 1.729 to 4.015.

The Naive, ARIMA, SWA1, and SWA2 forecasts were significantly superior to the ERS-ALM and the WASDE forecasts. In addition, the Naive forecasts were significantly superior to the SWA1 and SWA2 forecasts and the WASDA forecasts were significantly superior to the ERS-ALM forecasts.

Commercial Pork Production

Commercial pork production forecasts showed no significant bias as measured by Mean Error. Interagency Livestock Estimates Committee forecasts had the lowest MAPE's at 2.764 and the highest was the ERS-ALM at 4.486. Second and third best MAPE's were SWA1 and SWA2, followed by the ARIMA, Naive and ERS-ALM forecasts. Ordering in the RMSPE's were slightly different with the lowest RMSPE being the SWA1 and the IC forecast was the second lowest. The SWA2 forecast was the third lowest followed by the ERS-ALM, ARIMA, and Naive forecasts.

None of the pork production forecasts were statistically superior.

Commercial Broiler Production

Several of the broiler production forecasts exhibited a significant bias as measured by the ME. Both composite forecasts as well as the ARIMA and Naive model forecasts had ME that were significantly different from zero. MAPE analysis indicated that the IC forecasts were the lowest (1.534), followed by the ERS-ALM forecast (2.975). The SWA1, SWA2, ARIMA, and Naive forecasts followed in that order. The RMSPE's ranked the forecast in the same order as the MAPE's, with the range in values from 1.965 percent to 5.265 percent.

The WASDE broiler production forecasts were statistically superior to all other forecasts.

Omaha Steer Price

None of the Omaha steer price forecasts exhibited any bias as measured by the ME. The rankings of the MAPE from lowest to highest were Naive, ARIMA, IC, SWA1, SWA2, and ERS-ALM. RMSPE analysis indicates that the Naive and ARIMA forecasts are the two best in that order. However, the SWA1 and SWA2 forecasts rank above the IC forecast.

The Naive, ARIMA, SWA1, and SWA2 forecasts were significantly superior to the ERS-ALM, WASDE, SWA1, and SWA2 forecasts. The ERS-ALM forecasts were significantly inferior all other forecasts.

Seven Market Barrow and Gilt Price

No significant bias was detected in the 7-market barrow and gilt price forecasts. The SWA1 forecast had the lowest MAPE, followed by the IC forecasts. Forecasts from SWA2 model had the third best MAPE, followed in

order by the ARIMA, Naive, and the ERS-ALM. The SWA1 also had the lowest RMSPE. The RMSPE for the IC forecasts was the second best.

The WASDE and SWA1 forecasts were statistically superior to the Naive, ARIMA, and ERS-ALM forecasts. In addition all other forecasts were statistically superior to the ERS-ALM forecasts.

Twelve city broiler prices

Forecasts of the 12-city broiler price had no significant bias. Composite forecasts SWA1 had the lowest MAPE, followed by the IC forecasts. Descending order of the rankings of the other MAPE's were, SWA2, ARIMA, ERS-ALM, and Naive. The lowest RMSPE for 12-city broiler price forecasts belonged to the SWA1 forecasts.

None of the forecasts were statistically superior.

Summary and conclusions

The evaluation indicates that composite forecasts perform well, but are not always the unambiguously superior forecasts. Comparing both bias and linear and quadratic loss functions indicate that on average the composite forecast performed the best among methodologies. However, it should be noted that all of the forecasting methods performed fairly well in point accuracy. Production forecasts in all cases had MAPE less than 5 percent. Price forecast MAPE's, with few exceptions, remained below 10 percent. RMSPE also indicate that the forecasting performance of all of the methodologies were in most cases below 10 percent. A few of the forecasts were significantly superior. However, no one methodology used is clearly a superior forecasting tool.

No statistically significant bias was found in the IC forecasts, even though the IC beef production forecasts were all lower than the actual values. The Interagency Committee did about as good of a job of forecasting as the composite models. However, these are in a sense already composite forecasts. ARIMA models also do a fairly good job of forecasting even with the minimum number of observations available for estimation. The poorest forecasting tool was the ERS-ALM. However, this tool was designed more for policy analysis and not one step ahead forecasts. The ERS-ALM also suffered from very low degrees of freedom in the estimation of the model, which effected the accuracy of the parameters.

One result stands out: the superior performance of the Naive model in forecasting beef production and prices. The MAPE for the Naive forecast of beef production was only 1.729 percent and the RMSPE was 2.047 percent. Naive forecasts for Omaha steer prices had a MAPE of 5.460 percent and a RMSPE of 6.61 percent. Over the period of the 1980's the optimal forecast of the next years beef production and prices would have been last years values. This is interesting when compared to the relatively volatile 1970's.

Using these composite forecasts appears to offer only small improvement for the USDA. This research suggests that the IC forecasts could be augmented by the use of some simple time series models along with the analysts judgment.

The results also suggest that if the ERS-ALM is to be used to make point forecasts some improvement is necessary. Overall the forecasts of the IC do not appear to be very inaccurate. The opportunity for improvement seems to be greatest for the price forecasts. Production estimates appear to be about as accurate as one could expect.

References

(1) Ashley, R., C.W.J. Granger, R. Schmalensee. "Advertising and Aggregate Consumption: an Analysis of causality". Econometrica, No. 48, (1980). 1149-1167.

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

(3) Davidson, Russel and James G. MacKinnon. "Several Tests for Model Specification in the Presence of Alternative Hypotheses" Econometrica. Vol. 49, No. 3, May, 1981.

(4) Elam, E.W. and S.H. Holder. "An Evaluation of the Rice Outlook and Situation Price Forecasts". Southern Journal of Agricultural Economics Dec 1985, 155-161.

(5) M.E. Gerlow and S. H. Irwin. "Economic Evaluation of Commodity Price Forecasting Models". Unpublished paper, Ohio State University.

(6) Granger, C.W.J., and Paul Newbold. Forecasting Economic Time Series. Academic Press, New York, 1977.

(7) R.E. Just and Gordon C. Rausser. "Commodity Price Forecasting with Large-Scale Econometric Models and the Futures Market". American Journal of Agricultural Economics. Vol. 63 no. 2, May 1981, 197-208.

(8) SAS/ETS User's Guide: Version 6. SAS institute Inc. Cary, NC, 1988.