

FORECASTING AGRICULTURAL PRICES USING A BAYESIAN COMPOSITE APPROACH

Christopher S. McIntosh and David A. Bessler

Abstract

Forecast users and market analysts need quality forecast information to improve their decision-making abilities. When more than one forecast is available, the analyst can improve forecast accuracy by using a composite forecast. One of several approaches to forming composite forecasts is a Bayesian approach using matrix beta priors. This paper explains the matrix beta approach and applies it to three individual forecasts of U.S. hog prices. The Bayesian composite forecast is evaluated relative to composites made from simple averages, restricted least squares, and an adaptive weighting technique.

Key words: composite forecasting, Bayesian, matrix beta, outperformance.

In a competitive market, participants strive to formulate optimal forecasts of uncertain prices. The agricultural decision maker or analyst can base such forecasts on information from any number of sources. For example, three possible sources of forecast information are futures prices, expert opinions, and historical cash prices. It becomes the task of the forecast user to synthesize and apply this information.

When faced with two or more forecasts of the same uncertain event, a typical reaction of an analyst is to attempt to identify which single forecast is best. The best forecast is then employed in the planning process while the others are ignored. Several authors (Bates and Granger; Clemen and Winkler; Bunn; Bessler and Brandt) have argued that rather than attempting to choose the best forecast, the analyst should form a composite forecast as a weighted average of the forecasts

available. Because different forecasts may contain useful, if not independent, information, a composite forecast will assuredly outperform the worst individual forecast and oftentimes show substantial improvement (in a mean squared error sense) over the best individual forecast. In this case, the analyst has something to gain by using the available information to form a composite forecast.

There are several approaches to combining forecasts. Clemen and Winkler suggest that a simple average of available models will perform well. While this approach is quite easy to apply, it assumes the weights are constant over the time period analyzed, ignoring the relative performance of the individual forecasts. Bates and Granger suggest deriving weights that minimize the composite forecast variance, where the variance is estimated from historical forecast performance of the individual forecasts. This method implicitly assumes that the forecast error processes will be stationary over time and generally requires a large number of observations. Other techniques of combining forecasts, including regression analysis and adaptive weighting schemes, can be used.

Some authors have suggested a Bayesian approach to forming composites based on the outperformance criterion (Bunn; Bessler and Chamberlain). This approach assigns weights to individual forecasts based on the user's prior beliefs regarding the relative performance of the individual forecasts over time and observed forecast performance. A Bayesian approach has the advantage of allowing the forecast user to control the degree to which the weights change, and, therefore, it need not be tied too closely to the data. The purpose of this paper is to explain and illustrate a Bayesian approach of combining forecasts us-

The authors are, respectively, Assistant Professor of Agricultural Economics at the University of Georgia and Professor of Agricultural Economics at Texas A&M University.

Thanks are due to Kusumben Mistry, Timothy A. Powell, and Robert Alan Davis for computer assistance and to anonymous reviewers for comments on earlier drafts.

Copyright 1988, Southern Agricultural Economics Association.

ing matrix beta priors. An earlier paper by Bessler and Chamberlain studied beta priors which are useful for combining *two* individual forecasts. Here we consider the empirical application of matrix beta priors to three real world forecasts. The performance of the Bayesian composite forecast is evaluated relative to composites constructed using a simple average, restricted ordinary least squares, and an adaptive weighting scheme.

BAYESIAN COMPOSITE FORECASTING

Bayesian composite forecasting requires that the analyst assign priors (initial beliefs regarding the probability that one forecast will outperform the other[s] in a finite set of forecasting trials) to weight each of the alternative forecast in the composite. The initial priors reflect both the assessment of the relative performance of the forecasts over a finite number of trials and the degree of confidence in this assessment. For multiple forecasts, these beliefs may be summarized using a matrix of pairwise beta distributions with a Dirichlet diagonal. The beta distributions are characterized by two parameters that reflect the mean and variance of the distributions. The Dirichlet distribution is sometimes referred to as a multinomial beta distribution and is characterized by m parameters, where m is the number of forecasts being examined. Consider a matrix of priors, A , with elements a_{12} and a_{21} giving the analyst's assessment regarding the relative performance of forecast 1 against forecast 2. Prior weighting in favor of forecast 1 would be indicated by a_{12} greater than a_{21} . The variance of the distribution is a function of the magnitude of these parameters. The greater the values, the tighter the prior densities, reflecting a greater degree of certainty on the part of the analyst. These off-diagonal elements of the matrix A are an assessment of partial outperformance, in the sense that they pertain to only one pair of forecasts. The diagonal elements are the parameters of the Dirichlet distributions and are an assessment of total outperformance (all elements of the matrix are positive). The parameter a_{11} would be associated with the analyst's assessment of the likelihood of forecast 1 outperforming *all* other forecasts in the composite. Here again, the larger the parameter, the tighter is the

prior density.

In forming the composite forecast, the subjective priors are combined with performance data generated as a finite series of trials (observations). Each historical observation can be treated as a Bernoulli trial in which forecast i either does or does not outperform forecast j . The beta and Dirichlet distributions are conjugate for data from a Bernoulli distribution (DeGroot). This means that the posterior distribution resulting from the updating of the prior will be of the same form as the prior.

With the matrix beta defined as an $(m \times m)$ array of beta densities with a Dirichlet diagonal, posterior means of the beta distributions are given by the matrix K with

$$(1) k_{ijn} = (a_{ij} + s_{ijn}) / (a_{ij} + a_{ji} + n),$$

for $i \neq j$,

where s_{ijn} denotes the number of times forecast j has outperformed forecast i in n realizations (all ties are credited to both forecasts).

Likewise,

$$(2) k_{jin} = (a_{ji} + s_{jin}) / (a_{ji} + a_{ij} + n),$$

for $i \neq j$,

where $s_{ijn} + s_{jin} = n$, and a_{ij} , a_{ji} denote the priors given to forecast i and forecast j .

The posterior means of the m Dirichlet distributions are given by

$$(3) k_{iin} = (a_{ii} + s_{iin}) / (b_{ii} + n),$$

$i = 1, 2, \dots, m,$

where s_{iin} denotes the number of times forecast i outperformed *all* others in the n realizations, and m is the number of forecasts,

$$(4) b_{ij} = \sum a_{jj} \text{ for all } j \neq i,$$

where a_{ii} and a_{jj} are the prior Dirichlet parameters, and

$$(5) n = \sum s_{iin} \text{ for all } i.$$

To form the composite forecast, a vector of outperformance probability weights must be extracted from the matrix (K) of posterior means.¹ The rows of the posterior mean

¹ The posterior mean provides an unbiased point estimate and is representative of the Bayesian expectation.

matrix K are normalized, giving what is referred to as the outperformance probability matrix Q with elements:

$$(6) q_{ij} = k_{ij} / \sum_{j=1}^n k_{ij}.$$

The steady-state vector of Q gives the posterior mean vector (weight vector) of the outperformance matrix beta distribution.²

Prior Beliefs

To formulate a Bayesian composite forecast, the user must state his or her prior beliefs regarding the likelihood of one individual forecast outperforming one or all of the others. There are several ways to do this. If the user has no prior knowledge or opinions regarding relative forecast performance, a uniform prior may be employed giving equal weighting initially and allowing the performance of the forecasts to greatly influence the weights. If the user has some prior notion about how the forecasts should be weighted, an ad hoc procedure can be used to assign weights to the individual forecasts along with a "degree of certainty." These weights are then employed in the first step of the composite, with the "degree of certainty" determining the extent that the individual forecast performance is allowed to influence the successive weights. A third approach would be to actually elicit the user's subjective prior probabilities and fit the beta and Dirichlet distributions to these data (Bessler and Chamberlain).

A COMPARISON OF FOUR COMPOSITE FORECASTS USING U.S. HOG PRICES

Individual Forecasts

Composite forecasts were formulated from three individual forecasts. All actual price data and individual forecast series were quarterly observations on and forecasts of the USDA seven-market-average hog price for barrows and gilts (200-220 lb.) from the third quarter of 1973 through the second quarter of 1986. The individual forecast data were: a) an

expert's forecast, b) the futures market price, and c) a one-quarter-lead ARIMA (6,0,0) forecast. The expert's forecasts were for one-quarter-ahead cash prices made by Glen Grimes, professor of Agricultural Economics at the University of Missouri. The futures forecast prices correspond directly to the expert forecasts. The futures forecasts for each period are the closing price quoted in the annual *Yearbook* of the Chicago Mercantile Exchange for the day Grimes' forecast was published and for the contract that expired as close as possible to the end of the one-quarter lead time. The ARIMA model was identified and fit using quarterly cash prices from first quarter 1958 through second quarter 1973 (USDA). This model was then used to forecast one-step-ahead quarterly cash prices over the period from third quarter 1973 through second quarter 1986 and updated after each realization using the Kalman procedure in the RATS software package (Doan and Litterman). The forecast data appear in Table 1.

Composite Forecasts

Four composite forecasts are evaluated for relative forecast accuracy by examining their performance over the period of first quarter 1975 through second quarter 1986. The composites examined are a) a matrix beta Bayesian composite, b) a simple average of three forecasts, c) a restricted ordinary least squares (ROLS) combination, and d) an adaptively weighted composite based on forecast error histories.

The matrix beta Bayesian composite forecast was calculated in two steps. The first step was to examine relative forecast performance over the first six realizations (third quarter 1973 through fourth quarter 1974). A uniform prior, implying no prior information (i.e., $a_{ij} = 1$ for all i and j), was employed for this first set of realizations. The ending weights indicated were 0.500, 0.261, and 0.239 for the expert, futures, and ARIMA forecasts, respectively. The second step was to employ the weights indicated by the initial set of realizations as a prior for the remaining forecast realizations (first quarter 1975 through second quarter 1986), along with

² The steady-state vector is the unique vector which satisfies the equation $Qp = p$. The solution can be obtained by solving $(I - Q)p = 0$, where I is an $(m \times m)$ identity matrix and $p_1 + p_2 + \dots + p_m = 1$. A software package (titled COMPFORE) for combining two to five forecasts using the matrix beta approach is available from the authors. COMPFORE is designed to run on MS-DOS computers with 256K RAM.

TABLE 1. QUARTERLY ACTUAL PRICES RECEIVED AND THREE FORECASTS OF SEVEN-MARKET-AVERAGE PRICES FOR BARROWS AND GILTS (200-220LB.), THIRD QUARTER 1973 THROUGH SECOND QUARTER 1986

Date	Actual	Expert	Futures	ARIMA
	----- \$/cwt. -----			
7303	49.04	35.50	42.16	39.44
7304	40.96	40.00	38.79	52.13
7401	38.40	44.00	44.15	47.20
7402	28.00	32.00	34.80	28.60
7403	36.59	37.00	34.15	30.34
7404	39.06	39.00	40.65	26.89
7501	39.35	42.00	43.20	49.24
7502	46.11	42.00	45.65	30.51
7503	58.83	52.00	54.45	51.20
7504	52.20	61.00	61.85	59.34
7601	47.99	47.00	46.33	49.73
7602	49.19	47.00	46.80	48.92
7603	43.88	47.00	50.10	52.11
7604	34.25	36.00	33.43	36.35
7701	39.08	34.50	35.90	38.05
7702	40.87	35.50	36.97	44.88
7703	43.85	42.50	39.24	40.64
7704	41.38	37.50	37.86	46.18
7801	47.44	39.50	41.23	47.65
7802	47.84	48.50	47.74	45.43
7803	48.52	46.50	43.98	47.90
7804	50.05	48.50	51.93	46.14
7901	51.98	48.50	48.94	53.22
7902	43.04	45.00	46.98	49.32
7903	38.52	40.00	37.70	42.44
7904	36.39	35.00	37.25	38.41
8001	36.31	38.50	41.10	36.99
8002	31.18	35.50	36.03	34.91
8003	46.23	37.50	42.90	35.23
8004	46.44	43.50	49.08	50.52
8101	41.13	45.00	49.78	48.38
8102	43.62	42.00	53.30	38.66
8103	50.42	52.00	53.28	49.81
8104	43.63	49.50	49.25	41.03
8201	48.17	44.50	46.18	41.15
8202	56.46	51.00	57.33	52.21
8203	61.99	58.00	59.25	55.86
8204	55.12	59.50	59.23	56.61
8301	55.00	57.50	55.80	60.65
8302	46.74	53.50	51.34	53.23
8303	46.90	46.50	40.07	42.66
8304	42.18	41.50	42.09	43.69
8401	47.68	48.50	51.51	47.64
8402	48.91	49.50	52.89	47.48
8403	51.21	56.50	52.59	54.69
8404	47.65	45.50	44.94	50.24
8501	47.33	49.50	51.95	51.05
8502	43.09	46.50	50.86	44.33
8503	43.62	48.50	44.64	43.11
8504	45.05	42.50	43.77	42.35
8601	43.30	44.50	42.19	47.82
8602	47.14	43.50	42.04	42.99

TABLE 2. MATRIX BETA PRIOR PARAMETERS USED IN FORMULATING THE BAYESIAN COMPOSITE FORECAST

	Prior Parameter Matrix ^a		
	Expert	Futures	ARIMA
Expert	26.000	13.572	12.428
Futures	26.000	13.572	12.428
ARIMA	26.000	13.572	12.428

^a These parameters correspond to weights of 0.500, 0.261, and 0.239 for the individual forecasts. The "degree of confidence" on a scale of 1-99 (with 99 being extremely confident) assigned to these weights was 20 and reflects the authors' subjective assessment of outperformance probabilities based on the first six forecast realizations.

a "moderate" degree of certainty.³ The prior beta matrix is shown in Table 2.

The ROLS approach uses regression analysis to determine the weights given the individual forecasts in forming a composite. A coherence restriction (that the weights must sum to one) was imposed. The ROLS regression model was fit over the period of third quarter 1973 to fourth quarter 1974. Forecasts were then made for one-step (quarter) ahead beginning with the first quarter of 1975. The model was updated at each step using the Kalman procedure in the RATS software package.

The adaptive weighting technique was based on forecast error histories. The weights were updated after every forecast realization based on the formula:

$$(7) w_{i,T-1} = \left[\begin{array}{c} k \\ \sum_{j \neq 1} \end{array} \begin{array}{c} T \\ \sum_{t=1}^T \end{array} \left(\sum_{j=1}^k e_{j,t}^2 \right) \right] /$$

$$(k-1) \sum_{i=1}^k \sum_{t=1}^T e_{i,t}^2,$$

where k is the number of individual forecasts, i is the individual forecast ($i=1, \dots, k$), T is the total number of realizations to date, and t is the time period in which the forecast was made. The composite was formulated beginning in the first quarter 1975 through the second quarter 1986 based on forecast data starting in the third quarter of 1973.

Table 3 contains the composite forecast values obtained from each of the four techniques and their respective mean squared forecast errors (MSFE) for the period from the first quarter of 1975 through the second

quarter of 1986. The adaptively weighted forecast achieved the lowest MSFE (13.379) of the composite methods over the set of forecasts examined with the Bayesian composite only slightly worse (13.458). The MSFE of the simple average and ROLS composite forecasts were 13.643 and 15.447, respectively. All of the composite forecasts achieved a lower MSFE than any of the individual forecasts, which achieved MSFEs of 15.48, 18.37, and 25.59 for the expert, futures, and ARIMA forecasts, respectively, over the same period.

The MSFEs of the composite forecasts were examined using a test for significant difference in forecast accuracy developed by Ashley et al. This test decomposes the MSFE into its bias and variance components and provides additional insight into relative forecast accuracy. These statistics are summarized in Table 4. The Bayesian composite had significantly lower bias (at the .05 level) than the adaptively weighted, simple average, and restricted least squares composites. The adaptively weighted composite had significantly lower bias than the restricted least squares composite. No significant differences between composite forecast variances were detected.

CONCLUSIONS

The results indicate that, given these data and a quadratic loss performance metric (MSFE), the analyst would have been better off using a composite forecast rather than attempting to identify a "best" individual forecast. Attempting to choose a single "best" forecast either for an individual period or over a number of periods would likely have resulted in decreased forecast accuracy.

³ The prior used in forming the Bayesian composite was based on the weights obtained from the first six observations. The "moderate" degree of certainty was assigned using the "ad hoc" procedures described in the text and reflects the authors' subjective assessment of outperformance probabilities of the individual forecasts given the information provided by the first six observations. A higher "degree of certainty" would have resulted in a lower composite forecast mean squared error than reported here, while a lower degree of certainty would have resulted in a higher mean squared error.

TABLE 3. COMPOSITE FORECASTS OF QUARTERLY HOG PRICES, FIRST QUARTER 1975 THROUGH SECOND QUARTER 1986

Date	Bayesian	Simple Average	Restricted OLS	Adaptive Weighting
	----- \$/cwt. -----			
7501	44.044	44.814	44.624	44.081
7502	40.264	39.386	45.252	41.215
7503	52.483	52.549	54.794	52.885
7504	60.870	60.731	61.985	61.081
7601	47.458	47.686	46.432	47.281
7602	47.388	47.573	46.938	47.310
7603	49.076	49.736	51.151	49.341
7604	35.395	35.259	33.626	35.023
7701	35.718	36.149	35.872	35.771
7702	38.194	39.115	37.347	38.050
7703	41.103	40.792	39.039	40.730
7704	39.724	40.515	38.526	39.555
7801	41.964	42.795	41.952	41.956
7802	47.511	47.225	47.270	47.478
7803	46.081	46.127	44.336	45.829
7804	48.916	48.855	51.063	49.285
7901	49.810	50.221	49.694	49.788
7902	46.702	47.100	47.482	46.811
7903	39.945	40.047	38.595	39.677
7904	36.533	36.888	37.343	36.690
8001	38.933	38.864	40.269	39.166
8002	35.511	35.479	35.776	35.565
8003	38.521	38.542	40.856	39.032
8004	47.033	47.701	49.357	47.384
8101	47.367	47.719	49.201	47.694
8102	44.646	44.655	48.640	45.537
8103	51.838	51.696	52.284	51.927
8104	47.203	46.593	48.119	47.272
8201	44.105	43.944	44.662	44.261
8202	53.270	53.523	54.254	53.670
8203	57.823	57.702	58.268	57.921
8204	58.656	58.447	58.873	58.724
8301	57.805	57.985	57.162	57.675
8302	52.726	52.690	52.294	52.620
8303	43.354	43.075	42.207	43.092
8304	42.274	42.426	42.143	42.284
8401	49.274	49.217	49.873	49.398
8402	50.067	49.957	50.828	50.235
8403	54.733	54.594	54.349	54.587
8404	46.592	46.892	46.021	46.483
8501	50.719	50.834	50.986	50.813
8502	47.316	47.228	48.214	47.539
8503	45.800	45.414	45.839	45.661
8504	42.861	42.873	43.043	42.921
8601	44.671	44.838	44.083	44.598
8602	42.883	42.842	42.735	42.788
MSFE	13.458	13.643	15.447	13.379

TABLE 4. SIGNIFICANCE LEVELS FOR T-TESTS ON THE INDIVIDUAL COEFFICIENTS OF THE REGRESSION $\Phi_t = \beta_1 + \beta_2 [\Sigma_t + m(\Sigma_t)]^a$

	Level of Significance ^b							
	Bayesian		Simple Average		Restricted OLS		Adaptive Weighting	
	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2
Relative To:								
Bayesian	----	----	0.05	0.36	0.03	0.14	0.03	(0.35)
Simple Average	(0.05)	(0.36)	----	----	0.08	0.18	(0.43)	(0.32)
Restricted OLS	(0.03)	(0.14)	(0.08)	(0.18)	----	----	(0.03)	(0.08)
Adaptive Weighting	(0.03)	0.35	0.43	0.32	0.03	0.08	----	----

^a $\Phi_t = e_{1t} - e_{2t}$, where e_{it} is the t th forecast error from forecast i , $\Sigma_t = e_{1t} + e_{2t}$, and $m(\Sigma_t)$ is the mean of the Σ_t . β_1 is a measure of the difference in bias between the two forecasts; β_2 is a measure of the difference in variance. For a detailed description of this test, see Ashley et al. This test has been corrected for the sign of $m(\Sigma_t)$.

^bThe values *not* in parentheses indicate that the parameter estimates were positive. A positive parameter estimate indicates that the composite forecast named by the column heading has a significantly higher bias or variance than the composite forecast named by the row heading. For example, the simple average composite has significantly higher bias than the Bayesian composite at the 0.05 level of significance. Numbers in parentheses indicate that the parameter estimates were negative.

No general conclusions regarding the relative performance of various composite forecasts can be made on the basis of this one sample. However, the application serves to illustrate the relative performance of four composite methods in an applied context. For this particular data set, the Bayesian composite forecast performed better (achieving a slightly lower MSFE and significantly lower bias) than the simple average and restricted least squares approaches. Although the Bayesian composite achieved a slightly higher MSFE than the adaptively weighted forecast, it had significantly lower bias. Differences in the magnitude of MSFE between the Bayesian, adaptively weighted, and simple average com-

posites were slight.

Like the simple average of individual forecasts, the Bayesian approach does not require historical data on forecast errors. Unlike the simple average, the Bayesian approach allows the analyst to incorporate his or her own beliefs regarding relative forecast performance and considers the actual performance of the individual forecasts over time in forming the composite. These properties may be unimportant when one has two or three reliable forecasts. However, when one is faced with several forecasts of unknown quality and has little historical evidence on each, then the Bayesian composite is a useful alternative to the simple average.

REFERENCES

- Ashley, R., C. W. J. Granger, and R. Schmalensee. "Advertising and Aggregate Consumption: An Analysis of Causality." *Econometrica*, 48:5 (1980): 1149-67.
- Bates, J. M., and C. W. J. Granger. "The Combination of Forecasts." *Oper. Res. Quart.*, 20(1969): 451-68.
- Bessler, D. A., and J. A. Brandt. "Forecasting Livestock Prices with Individual and Composite Methods." *Applied Economics*, 13(1981):513-22.
- Bessler, D. A., and P. J. Chamberlain. "On Bayesian Composite Forecasting." *Omega*, 15(1987):43-48.
- Bunn, D. W. *The Synthesis of Forecasting Models in Decision Analysis*. Stuttgart: Burkhäuser Verlag Basel, 1978.
- Chicago Mercantile Exchange. *Yearbook*. Chicago, IL, 1973-87.
- Clemen, R. T., and R. L. Winkler. "Combining Economic Forecasts." *J. Bus. and Econ. Stat.*, 4(1986):39-46.

- DeGroot, M. H. *Optimal Statistical Decisions*. New York: McGraw-Hill Book Company, 1970.
- Doan, T. A., and R. B. Litterman. *Regression Analysis of Time Series*. Minneapolis, MN: VAR Econometrics, 1984.
- Grimes, G. *Missouri Agricultural Outlook Letter*. University of Missouri, Columbia, MO. Issued monthly.
- U.S. Department of Agriculture. *Livestock and Meat Statistics*. Washington, D.C., 1958-87.