CORNELL
AGRICULTURAL ECONOMICS
STAFF PAPER

An Appraisal of Composite Forecasting Methods

David W. Park
and
William G. Tomek

June 1986 No. 86-13

Department of Agricultural Economics
Cornell University Agricultural Experiment Station
New York State College of Agriculture and Life Sciences
A Statutory College of the State University
Cornell University, Ithaca, New York, 14853

AN APPRAISAL OF COMPOSITE FORECASTING METHODS

David W. Park and William G. Tomek

In forecasting, the challenge faced by analysts is to select the "optimal" forecast and forecasting procedure. One approach to selecting such forecasts is to use composite forecasting methods. This paper reviews the conceptual framework for forming optimal composite forecasts and attempts to identify circumstances when composite forecasts might work best. Then, Monte Carlo simulations and real world examples are used to compare composite forecasts with individual forecasts and to compare alternative methods of making composite forecasts. The appraisal is intended to provide insights into the potential benefits of composite forecasts and thereby provide some practical guides to the use of composite methods.

Price Forecasting Methodology

This section identifies the potential sources of forecast error, surveys alternative forecasting methods, and reviews the development of composite forecasting methodology.

Sources of Forecast Error

Errors associated with individual forecasting models may be categorized into four sources:

Sampling Error in Estimating Model Parameters. Since the parameters of forecasting models are unknown, they must be estimated based on sample data and/or prior information. Hence, the estimated parameter values are subject to sampling error. Since economic analysts typically cannot increase sample size or replicate experiments, little can be done to reduce sampling error.

True Error for the Forecast Period. While forecast error is assumed to be zero on average, a non-zero error is probable in any particular period. Such errors are associated with the random or non-systematic component of the time series under consideration which, by definition, is unpredictable and, hence, in principle cannot be reduced by additional modeling efforts. Alternative individual forecasts, therefore, may be subject to similar random errors, although in practice time-series methods may model the error term more carefully than does an econometric specification.

Erroneous Ancillary Forecasts. In the case of econometric models, ancillary forecasts of regressors are typically required in order to develop ex ante forecasts. Thus, forecasts can differ from the actual values because they were based on erroneous estimates of the regressors.

Model Misspecification and Structural Change. Since analysts rarely, if ever, know the true price-generating process, correct model specification poses a challenge. Specification error includes the exclusion of relevant variables or the inclusion of irrelevant variables in econometric models. In addition, the nature of the error term may be incorrectly specified in econometric or time-series. Finally, should structural change occur, a correctly specified model will produce erroneous forecasts unless both the timing as well as the nature of the change can be correctly anticipated.

David W. Park is an assistant professor of agricultural economics, Oklahoma State University; William G. Tomek is a professor of agricultural economics, Cornell University. Paper presented at the NCR-134 Conference, St. Louis, Missouri, April 22-23, 1986.

Unfortunately, given the lack of adequate theory and data, the identification of model misspecification and the assessment of its effects on forecast accuracy are often difficult to make.

Omitted variables will increase the variance of forecast error, but a case can be made that the resulting forecasts are still unbiased. The estimates of the true parameters of the correct model are generally biased when a variable is omitted, but a forecast depends on the estimates of the parameters of the misspecified model. These parameters are a linear combination of parameters of the correct model, and this combination depends on the relationship between the excluded and the included variables. For example, if $Y_t = \beta_1 X_{t1} + \beta_2 X_{t2} + e_t$ is the correct model, if X_{t2} is omitted and if $X_{t2} = \alpha X_{t1} + v_t$, then the fitted model is $Y_t = \beta_1{*} X_{t1} + e_t{*}$, where $\beta_1{*} = \beta_1 + \beta_2\alpha$. In principle, least squares provides an unbiased estimate of $\beta_1{*}$. The relationship between X_{t1} and X_{t2} (i.e., α) may or may not be stable and hence $\beta_1{*}$ may not be stable. The quality of forecasts using the misspecified model will depend on the stability of the relationship between the excluded and included variables in the forecast period. Also, omission of a variable will increase the estimated variance of the error term and may result in autocorrelated residuals.

While little can be done by the analyst to reduce the errors arising from the first two sources identified above, errors associated with ancillary forecasts and model misspecification may be reduced as the analyst gains additional insights into the processes generating the variables.[1]

Alternative Forecasting Methods

A common typology classifies forecasting techniques as being either qualitative, time-series, or causal methods. A number of studies have compared forecasting methods (e.g., Makridakis and Hibon, Groff, Kirby, Harris and Leuthold, Hogarth and Makridakis, and Leuthold, <u>et al.</u>). Such studies cannot identify the preferred procedure for all applications because the most accurate method will depend on the particular application -- the variable being forecast and the time period considered. In addition, particular comparisons are influenced by the expertise of the analyst. Nonetheless, such studies indicate that quantitative methods outperform qualitative methods and that sophisticated time-series methods outperform the simpler time-series methods. No general consensus exists regarding the relative performance of ARIMA and econometric methods. All methods are subject to potential problems of specification error and structural change, the seriousness of which varies with the time series being forecast.

Composite Forecasting Methods

When faced with two or more forecasts, a typical reaction has been to attempt to discover the "best" forecast and discard the rest, but this approach may not be appropriate where the "objective is to make as good a forecast as possible since the discarded forecasts nearly always contain some

---

[1] Once a forecasting model has been developed, it is common for researchers to assume away errors associated with ancillary forecasts and model misspecification. More realistically, one would expect such sources of error to persist in even the most carefully formulated model due to the complexity of price-generating processes and data limitations.

useful independent information" (Bates and Granger, p. 451). The independent information may arise as a result of including variables in one forecast which are not included in other forecasts, making different assumptions regarding the functional relationship between the variables, or obtaining better measures of the error term structure in one method than other methods. These benefits may be especially large if the process generating the variable is complex and, hence, difficult to model.

Since forecasts from alternative methods include independent information, the resulting composite forecasts should prove to be more accurate than any individual forecasts. For example, assuming two individual forecasts are unbiased, minimizing the variance of the composite forecast gives the "optimal" composite weighting scheme.

$$(1)\quad k^* = (\sigma_2^2 - \sigma_{12}) / (\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}).$$

If the variances of the two individual forecasts are equal, then the optimal weight is 0.5. If the individual forecasts are uncorrelated, then the optimal weight is based on relative variances. Composite forecasts based on k* will have an error variance at least as small as the more accurate of the individual forecasts. This result, however, assumes that the variance and covariance parameters are known. In practice, such parameters are unknown, and the performance of composite forecasts may deviate from theoretical expectations.[2]

In general, if $0 < k^* < 1$, the benefits associated with composite forecasts are greatest in cases where individual forecasts are negatively correlated. However, in cases where $k^* > 1$, the composite variance can be significantly smaller than the variances of either of the individual forecasts should the individual forecast errors be highly positively correlated.

Relaxing the assumption of unbiasedness invalidates the formulas developed above, but not the concept underlying composite forecasts. In contrast, relaxing the assumption of stationarity does not invalidate the results developed, but it does pose substantial empirical problems as attempts are made to estimate k* using (1), since the unknown variance and covariance values are changing from period to period. In this case, (1) might more appropriately be expressed with all parameters subscripted by time.

Composite methods cannot, however, reduce errors associated with the true random component of the series being forecast, and in practice, the individual forecasts may contain little or no independent information. Indeed, our experience suggests that the errors associated with alternative individual forecasts often possess variances of similar magnitudes and are highly correlated. In this case, the denominator of (1) approaches zero, and consequently, estimates of k* are highly sensitive to imprecision in the estimates of variances and covariances. Thus, development of accurate estimates of k* with limited data can be difficult. As a result, the performance of composite relative to individual forecasts is not always superior.

---

[2] The difficulty in using (1) in developing empirical estimates of k^* has been recognized by Bates and Granger, Harris and Leuthold, and Winkler and Makridakis.

In addition to optimal weighting schemes based on (1), a number of alternative formulas have been proposed. Among these alternatives are methods based on adaptive smoothing, weighted mean squared errors, and absolute deviation.[3] Many of these methods are admittedly ad hoc, but the simplicity with which they can be understood and implemented makes them of interest to practitioners.

Equations (2) and (3) below present general formulas which accommodate various assumptions with regard to adaptive smoothing and unequal weighting of observations. k_T is a vector of composite weights applied in period T.[4] Σ is the variance-covariance matrix of error terms of alternative individual forecasts. α is the smoothing coefficient and is restricted to lie between 0 and 1. w is the coefficient for weighting recent periods more heavily (for $w > 1$) than distant periods. $e_{i,t}$ is the error of the ith forecast in period t.

$$(2) \quad k_T = \alpha\, k_{T-1} + (1 - \alpha)\, (\Sigma^{-1}\, 1) / (1'\Sigma^{-1}\, 1),$$

where

$$(3) \quad (\Sigma)_{ij} = \sum_{t=1}^{T-1} w^t\, e_{i,t}\, e_{j,t}.$$

If $\alpha = 0$, then the resulting composite weights are not adaptively smoothed. If $w = 1$, then errors in each period are weighted equally in developing composite weights.[5] If one assumes that each forecast has the sameerror variance, $\sigma_1^2 = \sigma_2^2$, a composite weight based on simple averages results(i.e., for a two-forecast composite, each forecast would receive a composite weight of 0.5). Finally, if the covariance between error terms, σ_{ij}, is assumed to be equal to zero, Σ becomes a diagonal matrix. Allowing for alternative assumptions regarding adaptive smoothing, unequal weighting, and covariance terms results in eight distinct composite schemes (i.e., 2 * 2 * 2).

Monte Carlo Simulation

A study of individual and composite forecasting methods based on simulation techniques provides a way to compare forecasting methods under known conditions. Thus, the primary objective of the simulation is to develop practical guidelines for implementing composite forecasting methods and to gain insights into circumstances when composite forecasts work best by observing the performance of composite forecasting methods relative to individual forecasting methods under varying conditions.

In the simulation, alternative time-series patterns are considered, but for simplicity, the study is restricted to the case of one dependent variable

---

[3] Adaptive smoothing and unequal weighting of observations represent two schemes for dealing with non-stationarity in forecast performance.

[4] All individual and composite forecasts are ex ante (i.e., only include information available at the time the forecast is prepared).

[5] In cases where relative forecast performance among alternative models is changing over time, it may be advantageous to weight recent periods more heavily than distant periods in computing composite weights.

(Y) and one independent variable (X). Y is defined as a linear transformation of X plus an independent error term. X is defined as the sum of trend, seasonal, and random components. The specific generating equations for X and Y are presented in (4) and (5), respectively.

(4) $X_t = \alpha_0 + \alpha_1 t + \alpha_2 \cos(2\pi t/P) + ex_t$,

where α_0 is an intercept term, $\alpha_1 t$ is a trend component, $\alpha_2 \cos(2\pi t/P)$ is a seasonal (harmonic) component, and ex_t is a normally distributed error term with zero mean and standard deviation given by XESDV.

(5) $Y_t = \beta_0 + \beta_1 X_t + ey_t$,

where ey_t is a normally distributed error term with zero mean and standard deviation given by YESDV.

In generating the X and Y time series, values of eight parameters (i.e., α_0, α_1, α_2, P, XESDV, β_0, β_1, and YESDV) were specified. The functions of these parameters and the specific values considered in the study are presented in table 1. Parameter values were selected with the objective of generating a variety of time-series patterns.

Table 1. Monte Carlo Simulation: Time-Series Generation Parameter Values.

Parameter	Function	Values
α_0	mean adjustment of X	100
α_1	trend coefficient of X	0, 1, 2
α_2	seasonal amplitude of X	0, 10, 20
P	periodicity of X	12
XESDV	standard deviation of error of X	5, 10, 15
β_0	mean adjustment of Y	10
β_1	coefficient on X to generate Y	2
YESDV	standard deviation of error of Y	5, 10, 15

For each series, 132 observations were generated. Of these, the first 120 observations were used to estimate the forecasting models. Then, the models were sequentially reestimated over the next 11 observations, and 12 one-step-ahead forecasts were made. For each set of parameter values, ten replications were performed.

The individual forecasting methods considered include three forecasts from an econometric model (i.e., an econometric model evaluated on an ex post basis, evaluated with forecasted exogenous variables, and based on lagged exogenous variables), an exponential smoothing model, and a naive (no-change) model. The econometric models, based on the assumption that the correct model specification was know, were fit using least squares. The exponential smoothing model was fit based on the minimization of mean squared error (MSE).[6]

In the descriptions of the five individual forecasting models presented below, FY_{t+1} and FX_{t+1} denote the forecast of Y_{t+1} and X_{t+1}, respectively,

[6] The smoothing parameter was approximated with a grid search procedure based on the criterion of minimum mean squared error.

made in period t. Greek and Roman letters refer to true and estimated parameter values, respectively.

Econometric Model - Ex Post Evaluation (F1). X_{t+1} is assumed known in period t.

(6) $FY_{t+1} = b_0 + b_1 X_{t+1}$.

Econometric Model - Forecast Value of X (F2). F2 differs from F1 in that X_{t+1} is not assumed to be known in period t. An ancillary forecast of X_{t+1} is provided by a simple linear trend model as presented in

(7) $X_t = \gamma_0 + \gamma_1 t + e_t$,

(8) $FX_{t+1} = g_0 + g_1 (t+1)$,

and the forecast is made from

(9) $FY_{t+1} = b_0 + b_1 FX_{t+1}$.

Econometric Model - Lagged Value of X (F3). As an alternative to providing an ancillary forecast of X_{t+1}, F3 lags X by one period in the model specification. While this constitutes an error in variables, the potential benefit is that X_{t+1} no longer needs to be forecast prior to predicting Y_{t+1}. The model for Y_t is assumed to be

(10) $Y_t = \beta_0 + \beta_1 X_{t-1} + e_t$,

and the forecast is made from

(11) $FY_{t+1} = b_0 + b_1 X_t$.

Exponential Smoothing Model (F4). The general form of the exponential smoothing model is presented in (12). The smoothing coefficient, δ, is set where the mean squared error (MSE) is minimized. For t = 1, FY_t is set equal to Y_t.

(12) $FY_{t+1} = \delta Y_t + (1 - \delta) FY_t$.

Naive (No-Change) Model (F5). The naive (no-change) model simply states that the forecast value in the next period is the current value of the series.

(13) $FY_{t+1} = Y_t$.

F1 and F2 are based on the correct specification of model structure, but differ in the manner in which forecasts are developed. F3 addresses the need for an ancillary forecast by lagging X. F4 is a simple univariate smoothing model. F5, a no-change forecast, provides a benchmark against which the performance of alternative individual and composite forecasting models can be compared.

General equations for developing composite forecasts are presented in equations (2) and (3) above. Depending on the specification of α and w and assumptions regarding σ_1^2, σ_2^2, and σ_{12}, a variety of alternative composite forecasts are possible. Table 2 summarizes the parameter specifications and assumptions of composite forecasts examined by the Monte Carlo study. C1 through C4 assume that $\sigma_{12} = 0$. C5, a simple average, assumes that $\sigma_1^2 = \sigma_2^2$. C6 through C9 assume that σ_{12} does not equal zero.

For the initial forecast, no historical data on performance exist for computing composite weights. Consequently, in the first forecast period,

Table 2. Composite Forecasting Methods: Parameter Specifications and Assumptions.[a]

Composite Method	α	w	σ_{12}	σ_i^2
C1	0.0	1.0	0.0	
C2	0.3	1.0	0.0	
C3	0.0	1.5	0.0	
C4	0.3	1.5	0.0	
C5				$\sigma_1^2 = \sigma_2^2$
C6	0.0	1.0		
C7	0.3	1.0		
C8	0.0	1.5		
C9	0.3	1.5		

[a]Levels of α and w were selected based on preliminary findings of Bates and Granger.

forecasts were combined using equal weights (i.e., $k_T = 0.5$). In subsequent periods, historical information on forecast accuracy was incorporated into the calculation of weights according to the particular weighting scheme employed.

Based on nine methods, composite forecasts were formed between F2 and F4 and between F3 and F4. Hence, five individual forecasts and 18 composite forecasts are developed for each data series. Performance evaluation is based on mean absolute percentage error (MAPE) since this provides a unit-free measure of performance.[7]

The performance of each forecasting method across all time series is summarized in table 3. Clearly, ex post evaluation of forecast accuracy greatly overstates the precision which could be expected in true (ex ante) forecasting situations. This is a well-known result, and ex post evaluations are not emphasized in the remainder of this paper.

On average, over all of the comparisons, composite methods C1 through C5 produce forecasts which are equal or superior to (using the MAPE criterion) individual methods.[8] These composites give essentially identical results regardless of the method for combining the individual forecasts. But composite methods C6 through C9 produce poor forecasts, presumably because of the difficulties of obtaining precise estimates of the covariance terms from a

---

[7] In subsequent empirical examples, mean squared error (MSE) is used.

[8] Individual forecasting method F2, when considered separately from other individual forecasting methods, performed comparably to composites based on C1 through C5.

few observations.[9] (If priors were imposed on the probable covariance ranges, the performance of methods C6 through C9 likely would be improved.)

The average performance reported in table 3 is disaggregated in table 4 by selected levels of XESDV and YESDV. Not surprisingly, the forecasting accuracy of all methods deteriorated as the level of noise in the Y and the X series increases. The performance of composite forecasts C1 - C5, however, improves relative to the individual forecasts. Figure 1 illustrates the

Table 3. Monte Carlo Simulation: Performance of Individual and Composite Forecasting Methods.[a]

Forecast	Composite Method	Included Forecasts	Mean of MAPE	Std. Dev. of MAPE
F1	n.a.	n.a.	0.049	0.050
F2	n.a.	n.a.	0.097	0.051
F3	n.a.	n.a.	0.102	0.055
F4	n.a.	n.a.	0.101	0.054
F5	n.a.	n.a.	0.106	0.057
F6	C1	F2 & F4	0.098	0.052
F7	C2	F2 & F4	0.098	0.052
F8	C3	F2 & F4	0.098	0.052
F9	C4	F2 & F4	0.098	0.052
F10	C5	F2 & F4	0.098	0.052
F11	C6	F2 & F4	0.120	0.101
F12	C7	F2 & F4	0.117	0.093
F13	C8	F2 & F4	0.121	0.101
F14	C9	F2 & F4	0.117	0.092
F15	C1	F3 & F4	0.097	0.051
F16	C2	F3 & F4	0.096	0.051
F17	C3	F3 & F4	0.097	0.051
F18	C4	F3 & F4	0.097	0.051
F19	C5	F3 & F4	0.097	0.051
F20	C6	F3 & F4	0.124	0.194
F21	C7	F3 & F4	0.121	0.235
F22	C8	F3 & F4	0.128	0.195
F23	C9	F3 & F4	0.122	0.236
F2-F5	n.a.	n.a.	0.102	0.054
C1-C5	n.a.	n.a.	0.097	0.051
C6-C9	n.a.	n.a.	0.121	0.167

[a]MAPE is expressed in decimal form.

[9] Based on an analysis of variance model, the mean MAPE associated with F1, F2-F5, C1-C5, and C6-C9 were found to be significantly different at the 95 percent confidence level.

Table 4. Monte Carlo Simulation: Performance of Forecast Groups for Selected Standard Deviations of Variables.[a]

Forecast Group	Mean MAPE	Std. Dev. MAPE
	XESDV = 15, YESDV = 15	
ex post ind. forecast, F1	0.070	0.061
ex ante ind. forecast, F2 - F5	0.126	0.066
composite forecasts, C1 - C5	0.118	0.060
composite forecasts, C6 - C9	0.138	0.078
	XESDV = 5, YESDV = 15	
ex post ind. forecast, F1	0.043	0.032
ex ante ind. forecast, F2 - F5	0.085	0.042
composite forecasts, C1 - C5	0.083	0.041
composite forecasts, C6 - C9	0.101	0.100
	XESDV = 15, YESDV = 5	
ex post ind. forecast, F1	0.063	0.062
ex ante ind. forecast, F2 - F5	0.120	0.053
composite forecasts, C1 - C5	0.114	0.050
composite forecasts, C6 - C9	0.137	0.090
	XESDV = 5, YESDV = 5	
ex post ind. forecast, F1	0.024	0.025
ex ante ind. forecast, F2 - F5	0.074	0.037
composite forecasts, C1 - C5	0.073	0.037
composite forecasts, C6 - C9	0.081	0.044

[a]Performance averaged for all values of parameters other than standard deviations. MAPE is expressed in decimal form.

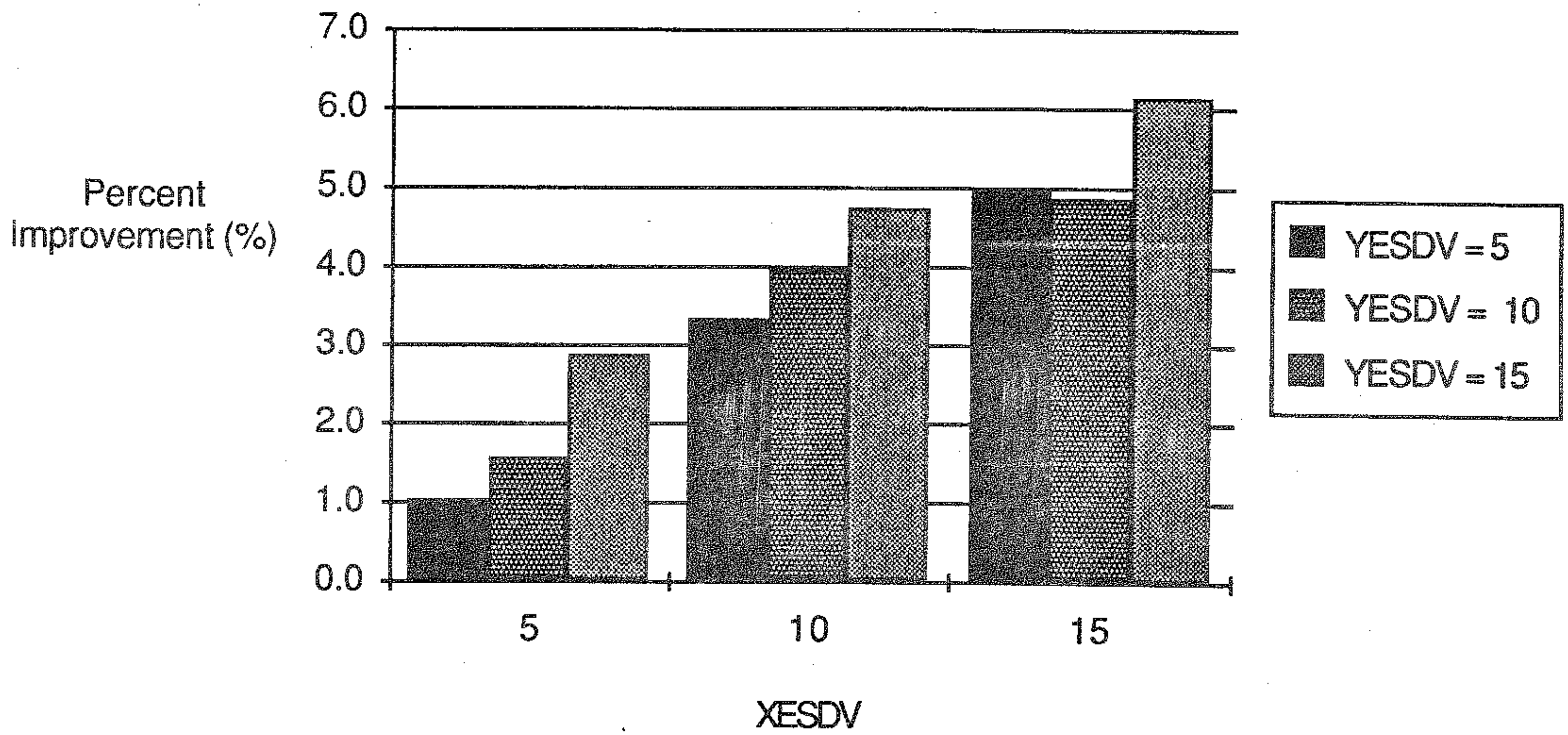

Figure 1. Percent Improvement in MAPE of Composite Relative to Individual Forecasts by XESDV and YESDV

percent improvement in MAPE associated with composites forecasts based on C1 through C5 relative to individual forecasts F2 through F5, collectively. For example, when XESDV and YESDV are 5, MAPE for C1 - C5 is 7.3 percent and for F2 - F5 is 7.4 percent, representing a 1.0 percent improvement in accuracy of composite relative to individual forecasts. However, when XESDV and YESDV are 15, the MAPE for composite forecasts is 6.1 percent smaller than for individual forecasts.

In our simulations, methods C1 through C5 give essentially identical MAPEs for a given pair of individual forecasts (as suggested in table 3). But the particular forecasts used in the composites do make a difference, and the performance of the individual forecasts is variable. These types of results are illustrated in tables 5 and 6. For example, F3 gives better forecasts than F2 and F4 when XESDV is small while F2 and F4 perform better than F3 when XESDV is large (table 5). Composites based on F3 and F4 are preferred when the variables have little random variation (table 6).

Thus, in comparing tables 5 and 6, one can see that the composites did not improve upon the best of the individual forecasts, and obtaining a composite forecast with a small MAPE depended on using an individual forecast with a small MAPE. For example, when X and Y are relatively variable, a composite between F3 and F4 is worse than F2 alone, and the composite of F3 - F4 does not improve upon the F3 forecast. Such results are not encouraging about the use of simulations to provide specific guides to doing composite forecasts. The results do indicate, however, that the composite is likely to be as good as or better than its constituent parts.

With respect to selecting an individual forecast, the simulations suggest that if explanatory variables have large random variability, as could be the case for a quantity variable based on production, the use of a lagged specification to "solve" the problem of ancillary forecasts is a poor choice. Using the correct specification with a simple ancillary forecast from trend outperformed the (incorrect) lag specification, which did not require a method for forecasting the explanatory variable. This comparison assumed, of course, that the analyst could forecast a part of the systematic component of X (trend), but did not assume that the ancillary forecast captured the entire systematic part of X (the seasonal component).

The simulations did not appraise a wide variety of time-series methods. A larger number of methods, however, were considered in two real-world applications, which are considered in the next section.

Empirical Examples

To explore further the performance of composite forecasts, two monthly time series were examined: slaughter steer prices and soybean oil prices. Individual forecasting models included a naive (no-change) model (NAIVE), an exponential smoothing model (EXPO), a stepwise autoregressive model (STEPAR), a simple trend-seasonal model (TS), a lagged econometric model (LAG),[10] an ARIMA model

---

[10] The LAG model consisted of regressing price on lagged prices.

Table 5. Monte Carlo Simulation: Performance of Selected Ex Ante Individual Forecasts.[a]

Forecast Method	Mean MAPE	Std. Dev. MAPE
	XESDV = 15, YESDV = 15	
F2	0.115	0.057
F3	0.132	0.070
F4	0.120	0.062
	XESDV = 5, YESDV = 15	
F2	0.086	0.044
F3	0.078	0.036
F4	0.088	0.046
	XESDV = 15, YESDV = 5	
F2	0.114	0.051
F3	0.126	0.054
F4	0.116	0.051
	XESDV = 5, YESDV = 5	
F2	0.075	0.041
F3	0.070	0.032
F4	0.079	0.042

[a]Performance averaged for all values of parameters. MAPE is expressed in decimal form.

Table 6. Monte Carlo Simulation: Performance of Selected Composite Forecasts.

Composites of:[a]	Mean MAPE	Std. Dev. MAPE
	XESDV = 15, YESDV = 15	
F2 & F4	0.116	0.059
F3 & F4	0.121	0.061
	XESDV = 5, YESDV = 5	
F2 & F4	0.077	0.041
F3 & F4	0.070	0.033

[a]Composite forecasts based on C1. All methods of combining, C1 through C5, give essentially the same results. MAPE is expressed in decimal form.

(ARIMA), and an econometric model (ECON).[11] Composite forecasts based on two and three individual forecasts were developed based on methods C1 through C9 discussed earlier.

The individual price forecasting models were initially fitted to the period from 79.01 to 84.01. Then, over the following 11 months up to and including 84.12, each of the models was reestimated. In each period, one-step-ahead forecasts were prepared. The mean squared errors of individual forecasting methods over the 12 month period ending 85.01 are summarized in table 7.

In forecasting steer prices, an econometric forecasting model is the most accurate individual method considered followed by a naive (no-change) model and an ARIMA model. In contrast, in forecasting soybean oil prices, the most accurate forecasts are provided by an ARIMA model with an econometric model providing the least accurate forecasts. Due to the complexity of the soybean oil market and the relative simplicity of the econometric model, poor performance probably is due to model misspecification.

For soybean oil prices, however, the ranking of individual forecast methods is sensitive to the criterion used to make the ranking. If MAPE rather than MSE is used, then the ranking becomes STEPAR, ECON, LAG, ARIMA, and NAIVE. This sensitivity is related to large price variability and large forecast errors in May and June 1984. Prices increased sharply in May and fell in June. This caused errors in forecasts and changed the subsequent parameter estimates as models were updated as well as causing the rankings to be sensitive to the criterion used to make the ranking. Perhaps of more interest is the fact that the temporary blip in price appears to have been related to a temporary action by Brazil and Argentina to limit exports. Such an action could not have been anticipated by any of the models, including an econometric model with a variable representing the impact of exporters (because it would have been impossible to make an accurate ancillary forecast of this variable). Thus, none of the individual forecasts captured the price increase, and the corresponding composite forecasts, of course, do not solve this type problem.

Table 7. Mean Squared Error of Individual Forecasting Methods, Omaha Choice 900-1,100 Pound Slaughter Steer Prices and Soybean Oil Prices, Crude, Tanks, f.o.b. Decatur, 84.02 - 85.01.

Forecast	Cattle Price ($/cwt.)^2	Soybean Oil Price (cents/lb.)^2
NAIVE	2.81	9.91
EXPO	5.45	n.a.
STEPAR	4.23	10.28
TS	3.04	n.a.
LAG	n.a.	10.29
ARIMA	2.86	9.83
ECON	2.63	11.77

[11] See Park for a detailed description of each forecasting model considered.

This incident emphasizes that, while individual forecasts may contain some independent information which when combined gives improved composite forecasts, all forecasts can miss important events that influence prices. All forecasts tend to be extrapolative in nature and are only as good as the underlying information. Indeed, our experience suggests that a variety of individual forecasting methods can make similar errors, especially when a large unanticipated event occurs. Over the forecast period, both steer and soybean oil prices behaved erratically, and many of the methods missed critical turning points. In this context, the naive (no-change) model performs relatively well by a MSE criterion.

Composite forecasts based on two and three individual forecasts were developed. All possible combinations of individual forecasts and composite methods were examined. Table 8 summarizes the average mean squared error of each composite method.

Composite methods using covariance terms performed poorly relative to alternative methods in forecasting both steer and soybean oil prices. Theoretically, one would expect such methods to be superior, but as discussed earlier, covariances between errors of alternative individual forecasts are difficult to estimate precisely.

Table 8. Mean Squared Error of Composite Forecasting Methods, Omaha Choice 900-1,100 Pound Slaughter Steer Prices and Soybean Oil Prices, Crude, Tanks, f.o.b. Decatur, 84.02 - 85.01.

Method	Two-Forecast	Three-Forecast
Cattle	$(\$/cwt.)^2$	$(\$/cwt.)^2$
C1	3.01	2.81
C2	2.93	2.74
C3	3.07	2.88
C4	2.97	2.79
C5	2.97	2.79
C6	3.73	6.46
C7	3.22	4.39
C8	4.31	6.73
C9	3.36	4.23
Soybean Oil	$(cents/lb.)^2$	$(cents/lb.)^2$
C1	9.95	9.79
C2	10.07	10.02
C3	9.93	9.73
C4	10.05	9.99
C5	9.90	9.73
C6	16.23	32.96
C7	14.36	21.24
C8	18.90	39.02
C9	15.60	23.62

Empirical results with regard to the use of adaptive and weighted composite methods were inconclusive. With both steer and soybean oil prices, the differences in relative performance of adaptive vs. non-adaptive and weighted vs. non-weighted composite methods were relatively small.

Due to the difficulty associated with the estimation of composite parameters, composite forecasts based on simple averages often performed relatively well. Composite forecasts based on simple averages represented the most accurate and second most accurate composite method considered in forecasting soybean oil prices and steer prices, respectively.

The inclusion of a third individual forecast in a composite forecast generally improved accuracy. The exception is in the case of covariance-based composite methods where accuracy diminished.

Our results indicate that composite forecasts often are superior and seldom inferior to their constituent individual forecasts,[12] and in cases where one individual forecast is substantially more accurate than another, the resulting composite often has accuracy which is only slightly poorer than the better individual forecast.[13] Additionally, since composite methods provide a way to address forecasting errors associated with model misspecification, the benefits of composite methods should be greatest in complex markets where relatively simple forecasting models are applied (i.e., in cases where model misspecification is likely to be important). From an examination of steer and soybean oil price forecasting, where the soybean oil market is more complex and the forecasting models more simplistic, this appears to be the case.

Summary

Despite theoretical arguments, the ultimate test of the benefits of composite forecasting methods is in terms of practical applications. Our empirical results suggest that composite methods provide a means of developing robust forecasting models of agricultural commodity prices. Composite methods perhaps have the greatest advantage when forecasting series generated by complex processes and when the problems of specification error and of making ancillary forecasts of regressors are large. Composite methods based on covariance terms probably should be avoided. Developing accurate estimates of parameters for the composite weights is difficult and whether or not precise estimates can be developed in practical applications needs further study.

Composite forecasts developed using simple averages of individual forecasts perform well in comparison with more sophisticated approaches. The results regarding the use of adaptive and weighted composite techniques are inconclusive. Three-forecast composites, in general, have greater accuracy than two-forecast composites. However, the improvement in accuracy is modest

---

[12] In forecasting soybean oil and steer prices, two-forecast composites based on C1 through C5 were superior to both constituent individual forecasts 56 and 45 percent of the time, respectively. Two-forecast composites based on C1 through C5 never were poorer than the least accurate of the constituent individual forecasts.

[13] In such cases, the poorer individual forecast often is given a small weight in the composite and, consequently, does not adversely affect forecast performance to any large extent.

in most cases. The greatest relative improvement is associated with the poorer two-forecast composites.

Composite forecasts, however, are not a cure-all for poor models and data, and composite methodology is subject to a number of caveats:

1. The relative performance of a specific composite method varies according to time series and period considered. While, on average, a particular method may be superior, in a specific application its use may not reduce forecast error. Use of composite methods should be viewed as a risk management tool to be applied with the objective of reducing forecast error over longer-planning horizons.

2. The use of composite techniques does not reduce the amount of care which should be exercised in developing individual forecasting models. A poorly formulated individual forecast, when included in a composite forecast, adds little and may detract in terms of improved accuracy.

3. Forecasts, by their very nature, are extrapolative. Hence, all individual forecasts may fail to capture a fundamental change in the generating process, and composite forecasts, likewise, provide little in terms of capturing such changes.

4. The improvements in accuracy associated with sophisticated vs. simple composite methods and three vs. two-forecast composites are often modest and may not warrant the additional costs associated with their development.

In sum, composite forecasting techniques are a pragmatic means of dealing with conflicting forecasts in a systematic fashion, but for a composite to have important benefits, it must be based on individual forecasts that contain correct, independent information.

References

Bates, J. M. and C. W. Granger. "The Combination of Forecasts." Op. Res. Qtrly. 20(1969):451-68.

Groff, Gene K. "Empirical Comparison of Models for Short Range Forecasting." Management Science 20(1973):22-30.

Harris, Kim S. and Raymond M. Leuthold. "A Comparison of Alternative Forecasting Techniques for Livestock Prices: A Case Study." North Central J. Agr. Econ. 7(1985):40-50.

Hogarth, Robin M. and Spyros Makridakis. "Forecasting and Planning: An Evaluation." Management Science 27(1981):115-37.

Kirby, Robert M. "A Comparison of Short and Medium Range Statistical Forecasting Methods." Management Science 13(1966):202-10.

Leuthold, R. M., A. J. A. MacCormick, A. Schmitz, and D. G. Watts. "Forecasting Daily Hog Prices and Quantities: A Study of Alternative Forecasting Techniques." J. Amer. Statist. Assoc. 65(1970):90-107.

Makridakis, Spyros and Michele Hibon. "Accuracy of Forecasting: An Empirical Investigation." J. Royal Statist. Soc. A 142(1979):97-145.

Park, David W. "An Appraisal of Composite Forecasting Methods with Applications to Agricultural Commodity Price Series." Ph.D. thesis, Cornell University, 1986.

Winkler, Robert L. and Spyros Makridakis. "The Combination of Forecasts." J. Royal Statist. Soc. A 146(1983):150-57.