DIVISION OF RESEARCH
Graduate School of Business Administration
The University of Michigan

November 1983

COMPOSITE PREDICTORS OF ACCOUNTING DATA

Working Paper No. 350

Eugene A. Imhoff, Jr.
Paul V. Pare

The University of Michigan

FOR DISCUSSION PURPOSES ONLY

Composite Predictors of
Accounting Data

Eugene A. Imhoff, Jr.*
Paul V. Paré**

November 10, 1983

* Professor of Accounting, The University of Michigan
**Assistant Professor of Accounting, Laval University

The authors are grateful to the Graduate School of Business, The University of Michigan, and the Faculty of Science and Administration, Université Laval for their financial support of this research. Please do not quote without permission. Comments are welcome.

Composite Predictors of
Accounting Data

This research investigates the performance of models that combine several different methods of predicting accounting data. Significant resources are consumed by corporate financial planners, security analysts, auditors, investment bankers, lenders, and others in an effort to predict and evaluate corporate sales and earnings data. Numerous investment services focus on these forward looking accounting data as a basis for investment decisions. Independent auditors are using forecasting methods in their analytical review procedures more frequently than ever before. Market-based studies have demonstrated that earnings forecasts contain information relevant to the valuation of securities. Hence, predictions of sales and earnings are of importance to the business community as a whole.

Here we empirically examine the potential contribution of composite prediction models to managerial forecasts of income statement data. While models may be used by corporate managers, the management forecast process is best described as a human subjective judgment process that may or may not consider times series or economic prediction models in forming judgments (Danos and Imhoff [1982]). We demonstrate, using security analysts' forecasts as surrogates for management forecasts, that the composite prediction models employed are capable of greater forecast accuracy than are some of the analysts' forecasts. More importantly, we present evidence which illustrates how econometric and time series forecasts can be used to compliment analysts' forecasts and form predictions that are more accurate still.

Background

A number of studies have attempted to identify models capable of accurately predicting accounting earnings (Foster [1977] Griffin [1977] Brown and Rozeff [1978] Lorek et al. [1976]). These forecast accuracy studies have often involved comparisons between models, analysts, and management (Imhoff [1978] Basi et al. [1976]). The models used in these studies have been ARIMA or other univariate time series models. While little work has been done to explore the use of econometric or associative models in forecasting accounting data, Brown and Ball [1967] have demonstrated that an entity's industry activity is capable of explaining a significant portion of earnings variability. Also, Gonedes [1973a] has suggested that accounting earnings are related to industry-wide and economy-wide factors, and the empirical results of Gonedes [1973b] and Magee [1974] have substantiated these relationships. We report results from our effort to predict earnings and sales using a combination of both extrapolative and econometric models.

This study evaluates both sales and earnings whereas prior forecast studies have focused on earnings. Sales represent the economic inflows of greatest importance to the entity, and predictions of sales have been somewhat neglected in the forecasting literature. As an example of the potential importance of sales forecasts, consider them in an audit context. From an auditor's perspective, analytical review of most income statement elements may be based on the relationships between elements of income (e.g., cost of sales) and of sales (Imhoff [1981]). However, testing the reasonableness of sales is an important first step in this process. Hence, the accurate predictions of sales may take on added significance to auditors in an analytical review context.

From management's perspective, sales are often the least controllable element of operations and the most difficult to predict. Corporate managers have noted the primacy of sales forecasts in the overall planning process (Danos and Imhoff, 1982). Once a level of sales activity is predicted, forecasting other elements of income is often a straightforward process. Moreover, in terms of the eventual bottom line effect of a given level of achieved sales, management may exercise some control by making various economic decisions (e.g., cost cutting measures) as well as noneconomic decisions involving accounting techniques (e.g., changes in estimates) during the year or at year-end.

From the viewpoint of an investor or creditor, it is from the sales stream that the cash flows needed to pay dividends or to repay debt and interest are generated. While accrual earnings are, in theory, a proxy for long run net cash flows of the entity, they are a noneconomic residual in the short run. Security analysts have long recognized the importance of the market share represented by sales and have attended to ex post and ex ante analysis of sales.

Because of the importance of sales data we were interested in learning whether efficiencies could be achieved by combining extrapolative and econometric forecasts of sales on both a quarterly and an annual basis. We were also interested in learning whether the combination of sales forecasts was more accurate than the combination of earnings forecasts. Given that sales are less subject to managerial influence than earnings, and are dependent on many exogenous factors beyond the scope of the firm's control, we expected sales to be less predictable than earnings.

Theory

The theory underpinning studies concerning the accuracy of forecasts of accounting data is elegant in its simplicity. The importance of historical accounting data has been demonstrated by the significant resources allocated to its production. Like most other costly historical data, predictions of their future values follow logically from the fundamental quest to reduce uncertainty for decision making purposes. Hence, we observe numerous producers and consumers of forward looking accounting data. Market-based research has empirically demonstrated the information content of historical accounting data (Ball and Brown [1968] Beaver [1968] Brown and Kennelly [1972]) as well as forecasts of accounting data (Imhoff and Lobo [1983] Waymire [1983] Penman [1980, 1982] Patell [1976] Gonedes, Dopuch, and Penman [1976]). Moreover, it has been demonstrated that the magnitude of forecast errors are related to the magnitude of unsystematic returns (Imhoff and Lobo [1983]). This last observation, while sensitive to the use of an appropriate error metric, is appealing in that it suggests that the forecast's accuracy is related to its value to investors in establishing market prices.

That forecasts of accounting data are useful in reducing uncertainties important to investors in establishing security prices is a theory which has both empirical content and intuitive appeal. That the accuracy of the predictions contributes to their value also has empirical support and analytical consistency in a risk averse world. Therefore, the pursuit of accurate forecast methods in accounting is theoretically well grounded, just as it is in statistics, economics, psychology, biology, physics, and other fields concerned with such phenomena as the weather, GNP, medical diagnosis of disease, molecular growth, and so on.[1]

Combining Forecasts

Combining forecast models to form a composite prediction has been recently investigated in nonaccounting contexts in a number of studies (Bates and Granger [1969] Newbold and Granger [1974] Dickinson [1973, 1975] Winkler and Makridakis [1983] and Makridakis and Winkler [1983]). The combining of forecast methods is efficient in the same way that combining securities in a portfolio to diversify investment risk is efficient. It has been demonstrated empirically that combining forecast models to form composite predictions significantly improves forecast accuracy and greatly reduces the variance of the prediction errors (Makridakis and Winkler [1983]). The simple averaging approach to composite prediction has also been applied to economic data with some success (Wall Street Journal [April 6, 1983, p. 48]). Yet, composite forecast methods have not yet been applied to accounting data in published forecast studies. This paper reports results of a combination process that suggests forecast accuracy of accounting data may be enhanced by combining extrapolative models with econometric models.

Combination Processes Examined

There are several ways to combine forecast models. The most straightforward is to take a simple (equal weighted) average of the

---

[1]We would argue, based on casual obervations of corporate financial planning systems, that the opportunities for academics to contribute to practice in business planning is far greater than in some of these other areas where the state of the art is much further along.

forecasts produced by the models. If a forecast generated by an extrapolative model were combined with a forecast from an econometric model, the combined forecast would be computed as:

$$C_t = \omega_1 E_t^1 + \omega_2 E_t^2 \tag{1}$$

where

C_t = combined forecast at time t

E_t^1 = forecast from extrapolative model at time t

E_t^2 = forecast from econometric model at time t

ω_i = weights assigned to forecasts, such that $\sum_{i=1}^{n} \omega_i = 1.0$.

In the equal weighted two-forecast case as above, $\omega_1 = \omega_2 = .5$. In the n forecast case, $\omega_1 = \omega_2 = \ldots = \omega_n = \frac{1}{n}$. Makridakis and Winkler [1983] have reported empirical evidence that supports the use of equal weighted combinations on the grounds of simplicity and overall forecast accuracy. However, they also report results from five different value weighted combination processes and demonstrate that some of the combinations are capable of achieving superior results (Winkler and Makridakis [1983]). Their combinations were based on ten different forecast methods consisting primarily of extrapolative smoothing models that could be computed in a completely automated mode.

Our combinations were all based on a time series model plus an econometric model. Based on the prior research regarding time series behavior of accounting data, the Box-Jenkins (ARIMA) method was used for the time series forecasts.[2] We also examined unequal weights in combining forecasts in order to allow the "best" forecast model to be more heavily weighted on a firm by firm basis. Cooper and Nelson [1975] have demonstrated that the optimum weights may be obtained by regressing the actual value of the variable to be predicted against the individual forecasts from the forecast models being considered. The least-squares estimate of ω_1 is given by

$$\omega_1 = (S_2^2 - S_1 S_2 r) \big/ (S_1^2 + S_2^2 - 2S_1 S_2 r) \tag{2}$$

where:

S_i = the estimated standard deviation of the forecast error of forecast model i (when i = 1,2)

r = the estimated correlation between forecast errors of the individual models.

While several other methods of computing weights have been examined (Ansley, Spivey, and Wrobleski [1976] Howrey, Hymans, and Greene [1980]), the applications to different forecast problems make it difficult to compare their results. We elected to use the Cooper and Nelson approach, modified by the fact that the actual data are available only prior to the forecast period.

The weights computed using equation (2) are based on the average performances of the individual forecast models over the sample period. It has been suggested that the optimal weights will vary over time for any given entity. For that reason, optimal composite predictors should

---

[2] The ARIMA method was applied on a firm by firm basis. Prior earnings forecast research has demonstrated the efficacy of ARIMA models. See Imhoff and Paré [1982] for a summary of these studies.

take into account the variability over time of the combining weights. If weights vary over time, expression (1) becomes

$$C_t = \omega_{1t} E_t^1 + \omega_{2t} E_t^2, \tag{3}$$

where C_t, E_t^1, and E_t^2 are as previously defined, and ω_{1t}, and ω_{2t} are the weights assigned to forecasts obtained from E^1 and E^2 at time t. The time-varying weights can be computed from the regression equation

$$A_t = \omega_{1t}E_t^1 + \omega_{2t}E_t^2 + e_t, \tag{4}$$

where A_t is the actual value of the forecasted variable at time t. If the weights are constrained to sum to one, equation (4) reduces to

$$A_t - E_t^2 = \omega_{1t}(E_t^1-E_t^2) + e_t. \tag{5}$$

Cooley and Prescott [1976] proposed the following structure for a regression model with a time-varying parameter:

$$Y_t = \beta_t X_t, \tag{6}$$

where β_t is subject to both permanent and transitory changes over time and varies according to

$$\beta_t = \beta_t^p + u_t \tag{7}$$

$$\beta_t^p = \beta_{t-1}^p + v_t, \tag{8}$$

where p denotes the permanent component of the parameter and the errors u_t and v_t are assumed to be normally distributed, with mean 0 and variances σ_u^2 and σ_v^2, respectively. From (8) it follows that at period N+1 the permanent component β_t can be writtten

$$\beta^p_{N+1} = \beta^p_t + \sum_{s=t+1}^{N+1} v_s; \tag{9}$$

that is, the permanent component at time N+1 is made up of the permanent component at period t plus all the permanent changes which occurred between t and N+1. Using (9), the time-varying regression model can be rewritten:

$$Y_t = \beta^p_{N+1} X_t + X_t u_t - X_t \sum_{s=t+1}^{N+1} v_s, \tag{10}$$

such that;

$$Y_1 = \beta^p_{N+1} X_1 + X_1 u_1 - X_1 v_1 - X_1 v_2 - \cdots - X_1 v_N$$

$$Y_2 = \beta^p_{N+1} X_2 + X_2 u_2 - X_2 v_2 - X_2 v_3 - \cdots - X_2 v_N$$

$$Y_N = \beta^p_{N+1} X_N + X_N u_N - X_N v_N.$$

The covariance matrix of the residuals is

$$\begin{array}{lll} X_1^2(\sigma_u^2 + N\sigma_v^2) & X_1X_2(N-1)\sigma_v^2 \cdots\cdots & X_1X_N\sigma_v^2 \\ X_2X_1(N-1)\sigma_v^2 & X_2^2(\sigma_u^2 + (N-1)\sigma_v^2) \cdots\cdots & X_2X_N\sigma_v^2 \\ \vdots & \vdots \qquad \vdots & \vdots \\ X_NX_1\sigma_v^2 & X_NX_2\sigma_v^2 \cdots\cdots & X_N(\sigma_u^2 + \sigma_v^2). \end{array}$$

The first step of the estimation procedure for time-varying weights is to compute the maximum-likelihood estimates of σ_u^2 and σ_v^2. The second step is to estimate the path of β_t, given all the observations on the independent variable Y_t. These estimates are computed using an optimal smoothing algorithm adapted from the engineering literature (Cooley, Rosenberg, and Wall [1977]).

Monte Carlo tests indicate that, in terms of estimation efficiency, time-varying regression is superior to ordinary least squares and generalized least squares when the structure of the model is subject to change over time (Cooley and Prescott [1973]). This varying-parameter technique has not been used to estimate weights in accounting data forecast combinations. We examine combinations of forecast models using all three weighting schemes: equal weights, value adjusted constant weights, time varying weights, and provide some insights regarding the efficiency of these methods as they apply to accounting data.

Research Method

Sample Selection

In order to evaluate the composite forecast approach we chose to track its performance over a number of years for a somewhat homogenous group of firms. A sample of 50 manufacturing companies satisfied the data requirements and provided a total of 800 observations of forecast performance. The data requirements included:

1. availability of quarterly sales and net earnings data since 1962
2. a December 31 year-end
3. publication of quarterly sales and earnings per share forecasts in Value Line Investment Survey for the period 1975-78
4. publication of industry data by the Federal Trade Commission in the Quarterly Financial Report since 1962.

The first condition guarantees a sufficient number of observations to estimate both the extrapolative and econometric models. The second condition follows from the variables, used in the econometric models.

These regression models are based on exogenous variables. If all firms have the same fiscal year, the predicted values of these exogenous variables used in the econometric models will have a controlled effect and their impact will be the same for all firms. Allowing different year-ends could have a confounding effect on the results. The third condition will permit a comparison of analysts' sales and earnings forecasts to serve as a benchmark for comparison with the composite predictors developed by the combination models. The fourth condition assures the availability of data to estimate the association between sales and earnings of a firm, its industry, and the economy.[3]

The four criteria were applied sequentially to the 2,466 firms making up the total population of the COMPUSTAT quarterly tape. First, all firms not classified in one of the industries listed in Appendix A (roughly 55 percent) were deleted. The next criterion (firms with a December 31 year-end) was necessary to insure a contemporaneous relationship between the financial results of individual firms and their industry's financial results as published by the Federal Trade Commission in the Quarterly Financial Report. An additional 439 firms did not satisfy this requirement. A total of 419 firms did not have quarterly sales and income data from 1962 to 1974. Only two firms were deleted for not having sales and net income forecasts published in the Value Line Investment Survey. Finally, 32 firms were deleted because they were either acquired, merged,

---

[3]Because of the absence of any industry data on a per share basis, all prediction models used sales and net earnings as the predicted variables. Later, when we compare these models to analysts' forecasts, we use analysts' sales and earnings per share forecasts since predictions of net earnings were not available. This difference was unavoidable.

or went bankrupt over the period of 1975 to 1978. Only 226 firms satisfied all the criteria and from that group, 50 were selected randomly to form the sample used in this study. Table 1 shows the industry composition of the sample based on S.I.C. two-digit industry codes.

Composite Predictors

The composite predictions consisted of the combination of an extrapolative model and an associative (econometric) model. The first component of the composite, the extrapolative forecast (E^1), was estimated using the Box-Jenkins (BJ) methodology (Box and Jenkins [1976]). For each firm in the sample, a BJ model was estimated to compute a forecast based strictly on previous sales or earnings data.

The second component of the combined forecast, the econometric forecast (E^2), was obtained from the regression model:

$$E^2_{ijt} = \alpha_0 + \alpha_1 I_{jt} + \alpha_2 M_t + \delta_1 Q_1 + \delta_2 Q_2 + \delta_3 Q_3 + \upsilon_{it} \qquad (11)$$

where;

E_{ijt} = reported earnings in period t by firm i in industry j,

I_{jt} = earnings in period t of all firms in industry j,

M_t = earnings at time t of all manufacturing firms,

Q_1 = dummy variable for quarter 1,

Q_2 = dummy variable for quarter 2,

Q_3 = dummy variable for quarter 3, and

υ_{it} = random disturbance term.

Following Brown and Ball [1967] and Gonedes [1973a], equation (11) was estimated in two stages. First, the earnings of all firms in industry j were regressed against the earnings of all manufacturing firms:

$$I_{jt} = \beta_0 + \beta_1 M_t + \beta_1 Q_1 + \beta_2 Q_2 + \beta_3 Q_3 + \varepsilon_{jt} \cdot \quad (12)$$

This first step isolates the factors which are specific to industry j and eliminates the potential problem of multicollinearity between the earnings of industry j and the earnings of all manufacturing firms (Gonedes [1973a, p. 414]). The second stage used equation (11), with ε_{jt} from (12) replacing I_{jt}, so that the final form of the general regression model was:

$$E^2_{ijt} = \alpha_0' + \alpha_1' \varepsilon_{jt} + \alpha_2' M_t + \delta_1' Q_1 + \delta_2' Q_2 + \delta_3' Q_3 + \upsilon_{it}' . \quad (13)$$

The combination of E^2 with E^1 considered three forms of the general model in (13): regular regression (ordinary least squares) (RR); regular cegression adjusted for serial correlation (AR); and, spectral regression (SR). This generated the three following $E^1 - E^2$ combinations of extrapolative and econometric models: C_1= BJ-RR; C_2= BJ-AR; C_3= BJ-SR. The weight assigned to each $E^1 - E^2$ forecast combination was based on its performance over the estimation period. In other words, if one model, say E^1, fits the series better, then more weight is given to this forecast in forming the composite predictor. It was expected that this weighting scheme would allow for improvements in accuracy over equal weighted combinations. The weights were computed on a case by case basis from the actual results available <u>prior to the estimation period</u>.

Three different approaches were used to compute the weights assigned to E^1 and E^2. The first weighting scheme assigns a constant weight to all four quarterly predictions generated by a model for a given year. The second approach assigns weights to quarters based on the model's performance for a specific quarter over the sample period. The third approach uses time-varying weights with the last four weights computed over the sample period used to form the composite predictors for the coming year. This gives a total of nine composite predictors, identified in the results of C1, C2, and C3 for combinations using constant weights, C1Q, C2Q, and C3Q for combinations using quarterly weights, and C1V, C2V, and C3V for combinations based on time-varying weights. Finally a simple average of the forecasts generated by the four basic models (BJ, RR, AR, SR) was also computed. This simple average composite predictor is identified as AVG.

In summary, a total of fourteen models were examined. The four basic models, the average of the four basic models, and nine pairwise combinations of the basic models were designated as:

1. BJ = Firm specific ARIMA model
2. RR = Ordinary least-squares
3. AR = Regression with correction for serial correlation
4. SR = Spectral regression
5. AVG = Simple average of BJ, RR, AR, and SR
6. C1 = Combination of BJ and RR using same weight for all quarters
7. C2 = Combination of BJ and AR using same weight for all quarters
8. C3 = Comgination of BJ and SR using same weight for all quarters
9. C1Q = Combination of BJ and RR using quarterly weights
10. C2Q = Combination of BJ and AR using quarterly weights
11. C3Q = Combination of BJ and SR using quarterly weights
12. C1V = Combination of BJ and RR using time-varying weights
13. C2V = Combination of BJ and AR using time-varying weights
14. C3V = Combination of BJ and SR using time-varying weights.

The sample data evaluated using these 14 models for the 16 quarters for each of the 50 firms resulted in 22,400 predictions of sales and earnings.

Hypotheses

The general hypothesis is that no combination of forecast models is superior to the individual models. More important, however, the tests of the generalized null were designed to determine if any single model, or combination of models, distinguished itself statistically from all others.

An appropriate design for comparing two forecasting models is to treat each pair of forecasts for a specific firm as a single case. The elements of each pair are the relative prediction errors from two forecasts. The matched pair is reduced to a single observation by taking the difference in the absolute relative errors. The problem then is reduced to a single sample test.

The usual parametric test is the paired t-test. This test assumes that difference scores are normally distributed. Since that assumption is likely to be violated, a nonparametric test--the Wilcoxon matched-pairs signed-ranks test--will be used to test the null hypotheses. The Wilcoxon test requires that the level of measurement be an ordered metric scale, i.e., the difference scores can be ranked across firms.

Empirical Results

Descriptive measures of the absolute relative prediction errors on net income are presented in Table 2. Table 3 shows the results for sales. The models are ranked according to the mean absolute relative error, starting with the model with the lowest mean.

The mean absolute relative errors for net income are affected by the presence of very large observations. This problem is due to the fact that, in some cases, the denominators used in computing the relative errors are very small.

To overcome the problem of large errors due to very small or negative denominators, an alternative is to use a different metric proposed by Imhoff and Paré [1982]. Their metric uses the standard deviation of the series being forecasted as the denominator to compute the relative prediction error. This measure has not only the advantage of avoiding the small denominator problem but it also makes comparisons across firms more meaningful. It can be argued that a forecast error may be better defined in the relative context of forecast errors of firms in that industry. For example, some industries are more sensitive to changes in general economic conditions. On an individual basis, the predictability of accounting numbers depends upon the volatility of the number being predicted. The greater the dispersion of a series the higher the expected difference between the actual and the predicted value of the variable being forecasted.

The descriptive measures of the absolute relative prediction errors based on the standardized metric are presented in Table 4 for net income and Table 5 for sales. When compared to the descriptive measures for net income based on different error metrics, the standardized metric reveals much more uniformity across models.

To compare the performance of the fourteen predictors, the models were ranked in terms of accuracy with a rank of 1 given to the model with the most accurate forecast and a rank of 14 given to the model yielding the least accurate forecast. If there is no difference in the predictive ability of the models then the mean ranks should be about equal. The Friedman analysis of variance by ranks was used to test the null hypothesis that there is no statistically significant difference in the average ranks of the forecasting models. Given that only one model having a significantly different average rank can lead to the rejection of this null hypothesis, the Wilcoxon test was also applied on a pairwise basis to identify the model or models that were significantly better in terms of predictive ability.

The results of the paired comparisons reported in tables 6 and 7 revealed that the composite predictor C2 consistently has a lower average rank than any of the other models for both sales and net income. This difference is statistically significant at the .05 level in all but one case. The only case where C2 was not ranked significantly better was in the comparison with C2Q which the most similar combination. In general, all nine of the composite predictors perform better than all but one of the four basic models. For net income, the combinations have a lower average rank in 34 of the 40 pairwise comparisons, with 30 of them being statistically different at the .05 level. For sales, 29 of the 33 cases for which the composite has a lower rank are statistically significant. Overall, the composite predictors perform as well as, or better than, the four basic models, and on an individual basis, C2 is a significantly better predictor of both sales and net income.

Comparisons with Analysts

Analysts' forecasts of net income and sales were compared to forecasts for the same variables generated by the three best composite models. Pairwise comparisons, using the Wilcoxon test, reveal that there were no significant differences (at the .05 level) between analysts' forecasts of net income and the forecasts from C2, C2Q, and AR. However, for predictions of sales these three models have a lower average rank than analysts and that difference is statistically significant at the .05 level.[4] These

---

[4]The models also generated mean errors and standard deviations which were significantly lower than Value Line is, providing a better estimate with greater consistency.

results are somewhat surprising given that analysts generally have information regarding the future sales and earnings of an entity not incorporated in the composite predictions.

In Combination with Analysts

A final step in the analysis was to examine the effect of combining predictions from models with those of analysts. We expected that the use of combinations involving both human judges (analysts) and models could provide further synergy since models efficiently process historical data (identifying trends and relationships) while analysts use current news and events that are not processed by models to make their predictions.[5] We combined the predictions in several ways, taking simple averages (e.g., Value Line plus Box-Jenkins plus AR divided by 3) as well as composites plus simple averages (e.g., C2 plus Value Line divided by 2). The results of these combinations demonstrated that using Value Line forecasts with models offered further improvements for sales but not for earnings. Table 8 reports the paired comparisons for sales, revealing that the combination of Value Line + BJ + AR has a significantly superior performance over all other models.[6] The second best model was that which combined C2 with Value Line.[7]

The results for earnings were more ambiguous. The combination of Value Line with other models required the conversion of forecasts for all

---

[5]The tendency for human judges to overweight current data ("recency effect") and underweight the past enables models to contribute to the accuracy of predictions by "expert judges" (such as Value Line analysts). Our current research is exploring this phenomenon for systematic biases.

[6]Only the close contenders are reported in Table 8.

[7]Again, the means and standard deviations were both lower (for these two- and three-way combinations with Value Line) than the alternatives.

models from net income to earnings per share (EPS). This transformation process converted forecasts of net earnings to EPS forecasts for models that were less precise than their net earnings forecast counterparts. As a result, Value Line forecasts generated lower average ranks than all other forecasts or combinations of forecasts, and the differences were significant in most cases. The only combination of models with Value Line that did not result in a significantly higher average rank than did Value Line alone was the C2 + VL combination.

Summary

The results reported here suggest that forecasts of accounting data may be enhanced by using combinations of extrapolative and econometric models. Using Value Line forecasts as a benchmark, there were several combination models that performed very well over the four year period [1975-78]. Bear in mind that these generic models could be readily enhanced if the econometric component incorporated more firm specific information. We used very general econometric models based on publicly available industry data.

Winkler and Makridakis [1983] have suggested that combining econometric forecasts and expert forecasts with time series forecasts might provide greater gains in accuracy than simply combining different time series forecasts. Our research has empirically examined these two extensions to time series combinations. First, we examined the efficiencies of combining time series forecasts with economic forecasts. These results demonstrated the combination of Box-Jenkins (BJ) and regular regression adjusted for serial correlation (AR) to be more accurate than any other single or composite prediction of earnings and sales. Using Value Line as a benchmark, we found the combination model (BJ + AR) to be statistically superior to Value Line for sales forecasts, and no different for earnings forecasts.

Next we examined combinations using time series forecasts, econometric forecasts, and Value Line forecasts. For sales, two different combinations of BJ, AR, and Value Line were more accurate than either Value Line alone or the models and combinations of models alone. For earnings, the results were more ambiguous, with Value Line and a combination of Value Line and C2 (BJ + AR) resulting in more accurate forecasts than most other models alone or in combination.

These results suggest that combination procedures offer potential benefits for accurately predicting accounting data. Further refinements of these methods should prove beneficial to auditors (e.g., in reviews of forecasts, for analytical review, etc.), corporate financial planners (e.g., in predicting performance of product lines or divisions, etc.), security analysts, bank lenders, and others interested in the future performance of the entity. These procedures offer a base for further inquiry into combination prediction processes.

APPENDIX A

S.I.C. 2-DIGIT INDUSTRY CODES AND NAMES

20	Food and Kindred Products
21	Tobacco Manufacturers
22	Textile Mill Products
26	Paper and Allied Products
27	Printing and Publishing
28	Chemicals and Allied Products
29	Petroleum and Coal Products
30	Rubber and Miscellaneous Plastics Products
32	Stone, Clay, and Glass Products
33	Primary Metal Industries
34	Fabricated Metal Products
35	Machinery, except Electrical
36	Electrical and Electronic Equipment
37	Transportation Equipment
38	Instruments and Related Products

TABLE 1

S.I.C. INDUSTRY BREAKDOWN OF SAMPLE

S.I.C. Two-digit Industry Code	Industry Name	Number of Firms
20	Food and kindred products	3
21	Tobacco manufacturers	1
22	Textile mill products	1
26	Paper and allied products	4
27	Printing and publishing	1
28	Chemicals and allied products	9
29	Petroleum and coal products	7
30	Rubber and misc. plastics products	2
32	Stone, clay, and glass products	1
33	Primary metal industries	2
34	Fabricated metal products	2
35	Machinery, except electrical	6
36	Electrical and electronic equipment	4
37	Transportation equipment	5
38	Instruments and related products	2
		50

TABLE 2

DESCRIPTIVE MEASURES OF ABSOLUTE
RELATIVE PREDICTION ERRORS
VARIABLE: NET INCOME
METRIC: (A-F)/F
N=800

Models	Minimum	Maximum	Mean	Std. Dev.
AVG	.00070	60.362	.71855	2.9085
AR	.00069	61.096	.71893	2.9966
C2Q	.00007	79.528	.80674	4.2379
C1	.00024	48.700	.84340	3.2620
C2V	.00054	50.067	.84449	3.4236
RR	.00005	85.013	.85688	4.2406
C2	.00031	198.991	.90119	7.4507
C1Q	.00014	167.432	.96015	6.6066
BJ	.00056	91.549	.96120	4.6809
SR	.00038	280.285	.99622	10.1005
C3V	.00059	134.598	1.04896	6.8179
C3	.00039	323.570	1.63564	15.1691
C3Q	.00029	957.927	1.89613	33.9076
C1V	.00019	1235.676	3.48549	49.4732

TABLE 3

DESCRIPTIVE MEASURES OF ABSOLUTE
RELATIVE PREDICTION ERRORS
VARIABLE: SALES
METRIC: (A-F)/F
N=800

Models	Minimum	Maximum	Mean	Std. Dev.
C2	.00006	.45821	.06522	.05777
C2Q	.00002	.46816	.06980	.06397
C1	.00008	.53710	.07235	.06476
C3	.00010	.51500	.07313	.06526
AR	.00024	.45196	.07470	.06378
C3Q	.00046	.46963	.07554	.06790
AVG	.00002	.44960	.07712	.06586
C1Q	.00002	.48863	.07737	.06848
C2Q	.00020	1.31786	.07893	.08486
BJ	.00001	.70975	.08032	.07528
C3V	.00011	.92579	.08269	.08325
C1V	.00002	1.40076	.08671	.09463
SR	.00005	.47909	.10192	.08359
RR	.00002	.48663	.11046	.09108

TABLE 4

DESCRIPTIVE MEASURES OF ABSOLUTE
RELATIVE PREDICTION ERRORS
VARIABLE: NET INCOME
METRIC: (A-F)/σ USING INCOME
N=800

Models	Minimum	Maximum	Mean	Std. Dev.
C2	.00130	17.712	.9922	1.3380
AR	.00311	17.702	1.0197	1.3279
C2Q	.00024	17.713	1.0201	1.4099
AVG	.00239	17.604	1.0360	1.3135
C1	.00087	17.589	1.0483	1.3855
C3	.00060	17.582	1.0539	1.3821
C1Q	.00044	17.577	1.0702	1.4236
C3Q	.00102	17.569	1.0735	1.4274
BJ	.00265	17.738	1.1271	1.4188
C3V	.00276	17.628	1.1278	1.5087
C2V	.00253	17.719	1.1632	1.5908
C1V	.00031	17.630	1.1691	1.5875
SR	.00133	17.482	1.1729	1.3987
RR	.00019	17.492	1.1734	1.4007

TABLE 5

DESCRIPTIVE MEASURES OF ABSOLUTE
RELATIVE PREDICTION ERRORS
VARIABLE: SALES
METRIC: $(A-F)/\sigma_A$
N=800

Models	Minimum	Maximum	Mean	Std. Dev.
C2	.00029	2.0468	.30223	.27960
C2Q	.00012	2.0771	.32187	.30302
C1	.00033	2.2761	.33467	.31568
C3	.00041	2.2143	.33794	.31753
AR	.00088	1.9959	.34249	.29147
C3Q	.00188	2.0815	.34841	.32775
AVG	.00011	1.9887	.35559	.31594
C1Q	.00011	2.1381	.35623	.32997
C2V	.00086	2.2253	.36040	.34995
BJ	.00007	2.7040	.36930	.35967
C3V	.00038	2.4140	.38687	.39212
C1V	.00008	2.4062	.39687	.38748
SR	.00024	2.2958	.46463	.39597
RR	.00008	2.6391	.50929	.43731

TABLE 6

SIGNIFICANCE LEVELS[1] (WILCOXON TEST) OF PAIRWISE COMPARISONS
VARIABLE: NET INCOME
METRIC: (A-F)/

Models	BJ	RR	AR	SR	AVG	C1	C2	C3	C1Q	C2Q	C3Q	C1V	C2V	C3V
Average Rank[2]	7.93	8.53	7.28	8.48	7.26	7.16	6.63	7.16	7.54	6.87	7.43	7.78	7.40	7.58
BJ	-													
RR	<.05 (BJ)	-												
AR	<.01 (AR)	<.001 (AR)	-											
SR	<.01 (BJ)	.84 (SR)	<.001 (AR)	-										
AVG	<.05 (AVG)	<.001 (AVG)	.38 (AR)	<.001 (AVG)	-									
C1	<.001 (C1)	<.001 (C1)	.73 (C1)	<.001 (C1)	.71 (C1)	-								
C2	<.001 (C2)	<.001 (C2)	<.05 (C2)	<.001 (C2)	<.001 (C2)	<.001 (C2)	-							
C3	<.001 (C3)	<.001 (C3)	.55 (C3)	<.001 (C3)	.89 (AVG)	.67 (C3)	<.001 (C2)	-						
C1Q	<.01 (C1Q)	<.001 (C1Q)	.21 (AR)	<.001 (C1Q)	.35 (AVG)	<.05 (C1)	<.001 (C2)	<.10 (C3)	-					
C2Q	<.001 (C2Q)	<.001 (C2Q)	<.10 (C2Q)	<.001 (C2Q)	<.01 (C2Q)	<.05 (C2Q)	.20 (C2)	<.05 (C2Q)	<.01 (C2Q)	-				
C3Q	<.01 (C3Q)	<.001 (C3Q)	.12 (AR)	<.001 (C3Q)	.40 (AVG)	<.05 (C1)	<.001 (C2)	<.05 (C3)	.13 (C3Q)	<.01 (C2Q)	-			
C1V	.96 (C1V)	<.01 (C1V)	<.01 (AR)	<.01 (C1V)	<.05 (AVG)	<.001 (C1)	<.001 (C2)	<.001 (C3)	<.01 (C1Q)	<.001 (C2Q)	<.01 (C3Q)	-		
C2V	<.05 (C2V)	<.001 (C2V)	<.05 (AR)	<.001 (C2V)	.95 (C2V)	.27 (C1)	<.001 (C2)	.28 (C3)	.77 (C1Q)	<.001 (C2Q)	.91 (C2V)	<.05 (C2V)	-	
C3V	.27 (C3V)	<.01 (C3V)	<.05 (AR)	<.001 (C3V)	<.05 (AVG)	<.05 (C1)	<.001 (C2)	<.01 (C3)	.15 (C3V)	<.001 (C2Q)	<.10 (C3Q)	.11 (C3V)	.28 (C2V)	-

[1] In parentheses: model with lower average rank in pairwise comparisons.

[2] Average rank of individual models. Friedman S statistic = 172.88 (p<.001).

TABLE 7

SIGNIFICANCE LEVELS[1] (WILCOXON TEST) OF PAIRWISE COMPARISONS
VARIABLE: SALES
METRIC: (A-F)/σ

Models	BJ	RR	AR	SR	AVG	C1	C2	C3	C1Q	C2Q	C3Q	C1V	C2V	C3V
Average Rank[2]	7.77	9.54	7.26	9.07	7.21	6.95	6.08	7.08	7.48	6.47	7.29	7.88	7.12	7.78
BJ	-													
RR	<.001 (BJ)	-												
AR	.13 (AR)	<.001 (AR)	-											
SR	<.001 (BJ)	<.001 (SR)	<.001 (AR)	-										
AVG	.59 (AVG)	<.001 (AVG)	.76 (=)	<.001 (AVG)	-									
C1	<.001 (C1)	<.001 (C1)	.19 (C1)	<.001 (C1)	<.01 (C1)	-								
C2	<.001 (C2)	<.001 (C2)	<.001 (C2)	<.001 (C2)	<.001 (C2)	<.001 (C2)	-							
C3	<.001 (C3)	<.001 (C3)	.37 (C3)	<.001 (C3)	<.05 (C3)	.72 (C1)	<.001 (C2)	-						
C1Q	<.05 (C1Q)	<.001 (C1Q)	.54 (AR)	<.001 (C1Q)	.88 (AVG)	<.001 (C1)	<.001 (C2)	<.01 (C3)	-					
C2Q	<.001 (C2Q)	<.001 (C2Q)	<.001 (C2Q)	<.001 (C2Q)	<.001 (C2Q)	<.01 (C2Q)	<.05 (C2)	<.001 (C2Q)	<.001 (C2Q)	-				
C3Q	<.01 (C3Q)	<.001 (C3Q)	.84 (C3Q)	<.001 (C3Q)	.22 (C3Q)	<.05 (C1)	<.001 (C2)	.13 (C3)	<.01 (C3Q)	<.001 (C2Q)	-			
C1V	.35 (BJ)	<.001 (C1V)	<.01 (AR)	<.001 (C1V)	<.01 (AVG)	<.001 (C1)	<.001 (C2)	<.001 (C3)	<.01 (C1Q)	<.001 (C2Q)	<.001 (C3Q)	-		
C2V	<.05 (C2V)	<.001 (C2V)	.47 (AR)	<.001 (C2V)	.25 (C2V)	.62 (=)	<.001 (C2)	.54 (=)	.14 (C2V)	<.001 (C2Q)	.55 (C2V)	<.001 (C2V)	-	
C3V	.84 (=)	<.001 (C3V)	<.05 (AR)	<.001 (C3V)	.19 (AVG)	<.001 (C1)	<.001 (C2)	<.001 (C3)	<.05 (C1Q)	<.001 (C2Q)	<.01 (C3Q)	.67 (C1V)	<.01 (C2V)	-

[1] In parentheses: model with lower average rank in pairwise comparisons.

[2] Average rank of individual models. Friedman S statistic = 493.51 ($p<.001$).

TABLE 8

RANK TEST FOR SALES

	$\frac{VL+BJ+AR}{3}$	$\frac{VL+C2}{2}$	C2	VL	AR	BJ
$\frac{VL+BJ+AR}{3}$	-	-	-	-	-	-
$\frac{VL+C2}{2}$	$\frac{VL+BJ+AR}{3}$*	-	-	-	-	-
C2	$\frac{VL+BJ+AR}{3}$**	$\frac{VL+C2}{2}$*	-	-	-	-
VL	$\frac{VL+BJ+AR}{3}$**	$\frac{VL+C2}{2}$**	C2*	-	-	-
AR	$\frac{VL+BJ+AR}{3}$**	$\frac{VL+C2}{2}$**	C2**	AR	-	-
BJ	$\frac{VL+BJ+AR}{3}$**	$\frac{VL+C2}{2}$**	C2**	VL	AR	-

* Significant at the .05 level
** Significant at the .01 level

References

Ansley, C. F., W. A. Spivey, and W. J. Wrobleski, "Statistical Analysis of Forecast Errors of Econometric Models," Proceedings of the Business and Economic Statistics Section, American Statistical Association, 1976, pp. 203-207.

Ball, R. and P. Brown, "An Empirical Evaluation of Accounting Income Members," Journal of Accounting Research (Autumn, 1968), pp. 159-78.

Basi, Bart A., Kenneth J. Carey, and Richard D. Twark, "A Comparison of the Accuracy of Corporate and Security Analysts' Forecasts of Earnings," The Accounting Review (April, 1976), pp. 244-254.

Bates, J. M., and C. W. J. Granger, "The Combination of Forecasts," Operational Research Quarterly (December, 1969), pp. 451-468.

Beaver, W. "The Information Content of Annual Earnings Announcements" Empirical Research in Accounting: Selected Studies: 1968 Supplement to Journal of Accounting Research pp. 67-92.

Beaver, W., R. Clarke, and W. Wright "The Association Between Unsystematic Security Returns and the Magnitude of Earnings Forecast Error." Journal of Accounting Research (Autumn, 1979), pp. 316-340.

Box, G. E. P., and G. M. Jenkins, Time Series Analysis, Forecasting and Control, rev. ed. (Holden-Day, 1976).

Brown, Lawrence D., and Michael S. Rozeff, "The Superiority of Analyst Forecasts as Measures of Expectations: Evidence from Earnings," The Journal of Finance (March, 1978), pp. 1-16.

Brown, Philip, and Ray Ball, "Some Preliminary Findings on the Association between the Earnings of a Firm, Its Industry, and the Economy," Empirical Research in Accounting: Selected Studies, 1976, Supplement to the Journal of Accounting Research, pp. 55-77.

Brown, P. and J. Kennelly "The Informational Content of Quarterly Earnings," Journal of Business (July, 1972), pp. 403-21.

Cooley, T. F., and E. C. Prescott, "Tests of an Adaptive Regression Model," Review of Economics and Statistics (May, 1973), pp. 248-256.

----, "Estimation in the Presence of Stochastic Parameter Variation," Econometrica (January, 1976), pp. 167-184.

Cooley, T. F., B. Rosenberg, and K. D. Wall, "A Note on Optimal Smoothing for Time Varying Coefficient Problems," Annals of Economic and Social Measurement (Fall, 1977), pp 453-456.

Cooper, J. P., and C. R. Nelson, "The Ex Ante Prediction Performance of the St. Louis and FRB-MIT-PENN Econometric Models and Some Results on Composite Predictors," Journal of Money, Credit, and Banking (February, 1975), pp. 1-32.

Danos, P. and E. A. Imhoff, Jr. "Forecast Systems, Construction and Attestation" Auditing: A Journal of Practice and Theory (Winter, 1982), pp. 23-34.

Dickinson, J. P., "Some Statistical Results in the Combination of Forecasts" Operating Research Quarterly (1973) pp. 253-260.

Fishman, George S., Spectral Methods in Econometrics (Harvard University Press, 1969).

Foster, G., "Quarterly Accounting Data: Time-Series Properties and Predictive-Ability Results," The Accounting Review (January, 1977), pp. 1-21.

Gonedes, Nicholas J., "Evidence on the Information Content of Accounting Numbers: Accounting-Based and Market-Based Estimates of Systematic Risk," Journal of Financial and Quantitative Analysis (June, 1973a), pp. 407-443.

----, "Properties of Accounting Numbers: Models and Tests," Journal of Accounting Research (Autumn, 1973b), pp. 212-237.

----, Nicholas Dopuch, and Stephen H. Penman, "Disclosure Rules, Information-Production, and Capital Market Equilibrium: The Case of Forecast Disclosure Rules," Journal of Accounting Research (Spring, 1976), pp. 89-137.

Griffin, P. A., "The Time-Series Behavior of Quarterly Earnings: Preliminary Evidence," Journal of Accounting Research (Spring, 1977), pp. 71-83.

Howrey, E. P., S. Hymans, and M. N. Greene, "The Use of Outside Information in Econometric Forecasting," Discussion Paper, Department of Economics, The University of Michigan, 1980.

Imhoff, Eugene A., "The Representativeness of Management Earnings Forecasts," The Accounting Review (October, 1978), pp. 836-850.

Imhoff, Jr., E. A. "Analytical Review of Income Elements" Journal of Accounting, Auditing, and Finance (Summer, 1981), pp. 333-351.

Imhoff, Jr., E. A. and G. Lobo "Market Reaction to Revisions in Analysts' Composite Earnings Forecasts" Working Paper-The University of Michigan, June 1, 1983 draft.

Imhoff, Eugene A., and Paul V. Paré, "Analysis and Comparison of Earnings Forecast Agents," Journal of Accounting Research (Autumn, 1982 Pt.I), pp. 429-439.

Jaggi, Bikki, "Further Evidence on the Accuracy of Management Forecasts Vis-a-Vis Analysts' Forecasts," The Accounting Review (January, 1980), pp. 96-101.

Lorek, Kenneth S., Charles L. McDonald, and Dennis H. Patz, "A Comparative Examination of Management Forecasts and Box-Jenkins Forecasts of Earnings," The Accounting Review (April, 1976), pp. 321-330.

Mabert, V. A., and R. C. Radcliffe, "A Forecasting Methodology as Applied to Financial Time Series," The Accounting Review (January, 1974), pp. 61-75.

Magee, Robert P., Industry-Wide Commonalities in Earnings," Journal of Accounting Research (Autumn, 1974), pp. 270-287.

Makridakis, S. and R. L. Winkler, "Average of Forecasts: Some Empirical Results," Management Science (1983 in press).

Nelson, Charles R., "The Prediction Performance of the FRB-MIT-PENN Model of the U.S. Economy," American Economic Review (December, 1972), pp. 902-917.

----, Applied Time Series Analysis (Holden-Day, 1973).

Newbold, P., and C. W. J. Granger, "Experience with Forecasting Univariate Time Series and the Combination of Forecasts," Journal of the Royal Statistical Society, A, 137 (1974), pp. 131-146.

Patell, James M., "Corporate Forecasts of Earnings per Share and Stock Price Behavior: Empirical Tests," Journal of Accounting Research (Autumn, 1976), pp. 246-276).

Penman, S. "An Empirical Investigation of the Voluntary Disclosure of Corporate Earnings Forecasts," Journal of Accounting Research (Spring, 1980), pp. 132-160.

Penman, S. "Insider Trading and the Dissemination of Firms' Forecasts Information," Journal of Business (October, 1982), pp. 479-503.

Waymire, G. "Some Evidence on Management Decisions to Publicly Issue Earnings Forecasts in 'Disclosure Package'" Working Paper-The University of Chicago, June, 1983 draft.

Winkler, Robert L., "Probabilistic Predictions: Some Experimental Results," Journal of American Statistical Association (December, 1971), pp. 675-685.

R. L. Winkler and S. Makridakis, "The Combination of Forecasts" Journal of the Royal Statistical Society (1983 in press).