Division of Research
Graduate School of Business Administration
The University of Michigan

August 1976

STATISTICAL ANALYSIS OF FORECAST ERRORS OF ECONOMETRIC MODELS

Working Paper No. 137

by
C.F. Ansley, W.A. Spivey, and W.J. Wrobleski

The University of Michigan

FOR DISCUSSION PURPOSES ONLY

Introduction

Increasing attention is being given to the evaluation of forecast errors of major U.S. econometric models. Recent studies are Duggal, Klein, and McCarthy [2], Eckstein, Green, and Sinai [3], Hirsch, Grimm, and Narasimham [8], Fromm and Klein [4,5], and Haitovsky, Treyz, and Su [6]. McNees [9] has also examined forecasts of eleven macroeconomic variables over a six-quarter horizon developed from eleven different econometric models.

In contrast to these studies, other forms of forecast evaluation have been developed (Nelson [10], Newbold and Granger [11], and Cooper and Nelson [1]) which combine econometric model forecasts with time series model forecasts, principally of the ARIMA type, to obtain a composite forecast or index. In these works the composite is a linear combination of forecasts in which the weights are estimated by standard least squares procedures with the constraint that such weights sum to 1. The latter property reflects the assumption that the constituent forecasts are unbiased estimates of actual values.

This investigation has three purposes: (1) to assess the bias in the forecasts of three econometric models--those of Data Resources Inc. (DRI), R. C. Fair, and the U.S. Department of Commerce (often called the BEA model); (2) to examine the correlation properties of the time series of errors from these models as well as the contemporaneous cross-correlation between forecast model errors; and (3) to examine the process of combining forecasts from these models by means of principal components analysis rather than by the methods used in Nelson [10], Newbold and Granger [11], and Cooper and Nelson [1]. The latter results are exploratory and are intended only as a first step in a continuing study of the problem of combining forecasts from several econometric and time series models.

The basic data used in this study consist of the actual and forecasted one quarter ahead changes in the levels of three macroeconomic variables--GNP in current dollars (GNP), unemployment rate (UR), and investment in residential structures (RS) developed from the DRI, Fair, and BEA models. The forecast period is from the third quarter of 1970 to the second quarter of 1975. The data are from those used in the McNees study [9]. The data from the DRI and BEA models in our study are used with the permission of the model proprietors; the Fair model is available in the literature. All of the forecasts were *ex ante* forecasts. Moreover, the forecasts do not necessarily represent the unmodified outputs of the individual models; in many cases the forecasts actually released have been adjusted by model proprietors to reflect judgmental considerations beyond those in the specification of exogenous variables. The forecasts thus represent the interaction of forecasters with their models.

Statistical Analysis of Forecast Errors

For a given quarter t, the underlying data which we analyze consist of the actual quarterly changes in the level of each of the three variables,

$$\nabla GNP_t = GNP_t - GNP_{t-1}$$

$$\nabla UR_t = UR_t - UR_{t-1}$$

$$\nabla RS_t = RS_t - RS_{t-1} \qquad t = 1,2,\ldots,20$$

and the *ex ante* forecast changes in these variables developed from the DRI, Fair, and BEA models, which we denote by

$$\hat{\nabla} GNP_t^{(j)}, \quad \hat{\nabla} UR_t^{(j)}, \quad \hat{\nabla} RS_t^{(j)}$$

where $j = 1,2,3$ indicates the DRI, Fair, and BEA models respectively. With this notation, one can denote an error for the forecast of the change in GNP made by forecaster j as

$$e(GNP_t^{(j)}) = \nabla GNP_t^{(j)} - \hat{\nabla} GNP_t^{(j)} .$$

To examine bias in the forecasts the means of the forecast errors and their standard deviations were calculated and are shown in Table 1. Only one is significant at the 10 percent level, supporting the hypothesis that forecasts of quarterly change over the given period are unbiased.

Table 1

Means and Standard Deviations of Forecast Errors

Variable and Forecaster	Mean	Std. Deviation	Value of t-Statistic
Gross National Product			
(1) DRI	0.815	8.733	0.417
(2) Fair	2.910	12.693	1.025
(3) BEA	1.390	8.133	0.764
Unemployment Rate			
(1) DRI	0.035	0.287	0.546
(2) Fair	0.010	0.358	0.125
(3) BEA	0.030	0.225	0.596
Investment in Residential Structures			
(1) DRI	0.705	1.539	2.049
(2) Fair	0.275	2.369	0.519
(3) BEA	0.430	1.643	1.170

Turning to a consideration of the serial correlation of the forecast errors, we use as the test statistic the von Neumann ratio of the mean square successive difference of the errors to their variance (von Neumann [12]) which is, for forecaster j and for GNP,

$$VN(GNP^{(j)}) = \frac{n}{n-1} \frac{\sum_{t=2}^{20} (e(GNP_t^{(j)}) - e(GNP_{t-1}^{(j)}))^2}{\sum_{t=1}^{20} (e(GNP_t^{(j)}) - \bar{e}(GNP^{(j)}))^2} .$$

This statistic has mean and standard deviation

$$E[VN] = \frac{2n}{n-1}$$

and

$$\sigma[VN] = \sqrt{\frac{4n^2(n-2)}{(n+1)(n-1)^3}}$$

respectively. The acceptance of the null hypothesis that no serial correlation exists, against the alternate hypothesis that either positive or negative serial correlation is present in the errors, is based on an examination of the sampling distribution of the von Neumann statistic, approximations of which appear in Hart [7]. At the 10 percent level of significance and for $n = 20$ the acceptance region for this test is $1.37 \leqq VN \leqq 2.85$. The results are summarized in Table 2.

It is clear that, except for the errors associated with the unemployment time series in the DRI and BEA models, there is evidence of (positive) serial correlation among the errors for each of the variables and models.

Table 2

Results of von Neumann Test of Serial Correlation

Variable	Value of von Neumann Statistic	Decision Concerning Null Hypothesis of No Serial Correlation
Gross National Product		
(1) DRI	1.23	Reject
(2) Fair	0.80	Reject
(3) BEA	1.36	Reject
Unemployment Rate		
(1) DRI	1.62	Accept
(2) Fair	1.17	Reject
(3) BEA	1.88	Accept
Investment in Residential Structures		
(1) DRI	1.13	Reject
(2) Fair	0.65	Reject
(3) BEA	1.34	Reject

Up to this point we have studied relationships among errors within a model. The question of whether there is a correlation in forecast errors across models can be considered by examining the contemporaneous cross-correlations which appear in Table 3. Using a 10 percent level of significance (the acceptance region for the test is $-.378 \leqq r \leqq .378$), one sees that all cross-correlations are significant. Although the structures of the three models differ in many important respects--different sizes

Table 3

Cross-Correlations of Forecast Errors
for GNP, UR, and RS

Forecasters	GNP	UR	RS
DRI and Fair	0.797	0.774	0.649
Fair and BEA	0.819	0.420	0.729
BEA and DRI	0.807	0.610	0.496

of equation systems, different exogenous and endogenous variables, and different lag structures--one concludes nevertheless that the forecast errors for the same variables from different models are correlated.

Analysis by Principal Components

The three models studied in this paper share a similar model-building approach based upon a common underlying macroeconomic theory, and this approach appears to embody an agreement as to ways of formulating this theory in econometric equations. The significant cross-correlations noted above also suggest that a dependency structure exists between the outputs of forecasters. How can this dependency be analyzed further? We approach this problem by using methods of principal component analysis in an attempt to discern factors among the several forecasters which may be attributable in turn to various similarities and differences in model-building strategy and implementation.

Table 4 shows principal components calculated from all twenty observations (1970.3 - 1975.2). It should be emphasized that the principal components are based on the ex ante forecast quarterly changes in levels of the variables GNP, UR, and RS. Finally, our calculations utilize the sample covariance matrices of these observations rather than correlation matrices because of the commensurability of the units of measurement for each variable from each forecaster.

It is clear that for all variables studied the first two principal components alone explain nearly all the total sample variance for all three forecasters. In fact, the first principal component accounts for at least 83 percent of the total sample variance for each variable. An examination of the magnitudes of the coefficients in the first principal component of GNP and RS indicates that BEA has the largest, followed in order by DRI and Fair. For the variable RU, however, DRI has the largest coefficient, followed by BEA and Fair. It may also be interesting to observe that DRI and Fair have coefficients of nearly equal magnitude for the variable GNP.

An examination of the second principal component in Table 4 indicates that it contributes 13.64 percent for GNP, 7.16 percent for RU, and 5.43 percent for RS.

Table 4

Principal Components for n = 20 (1970.3-1975.2)

Forecaster	$\hat{\nabla GNP}$ Coefficients in Principal Components			$\hat{\nabla UR}$ Coefficients in Principal Components			$\hat{\nabla RS}$ Coefficients in Principal Components		
	First	Second	Third	First	Second	Third	First	Second	Third
DRI	0.513	0.489	0.706	0.711	0.597	0.372	0.576	0.619	-0.533
Fair	0.438	-0.856	0.275	0.359	0.146	-0.922	0.313	0.435	0.844
BEA	0.739	0.168	-0.653	0.604	-0.789	0.110	0.755	-0.653	0.057
Sample Variance of Component	195.61	32.15	7.84	15.14	1.22	0.66	0.39	0.02	0.01
Cumulative Percent of Total Sample Variance	83.03	96.67	100.00	88.94	96.10	100.00	91.64	97.07	100.00

However, the pattern of magnitudes of the coefficients of the second principal components shifts more markedly from variable to variable than is the case for the first principal components. The third principal components are relatively unimportant.

Thus we see that although the models are highly dissimilar in specific structure and detail only two principal components account for at least 96 percent of the total sample variance for all variables studied. Unfortunately, we are unable at this time to give an interpretation to the first two components which would be helpful in studying the forecast generation process. If an extension of this analysis to include other macroeconomic models should reveal that the first two principal components dominate in a similar manner, however, such a finding might render the search for an interpretation worthwhile.

Regression on Principal Components

We turn now to the following problem: can combinations of forecasts be formed which explain the actual values better than the individual forecasts themselves? Nelson [10] and Cooper and Nelson [1] have responded to such a question by using regression analysis in which the actual values were regressed directly on the individual forecasts (which also included time series forecasts). In contrast to this approach, we engage in a regression type of analysis in which the explanatory variables are the first two principal components.

We believe that our approach is superior because it accommodates the presence of multicollinearity among the forecasts produced by the various forecasters (the principal component vectors being mutually orthogonal). For example, in following the approach of Nelson we found that least squares multiple regression of the actual change on the forecast changes produced models having no significant regression coefficients because of multi-collinearity.

For purposes of exposition we present the results of the regression on the first two principal components for one variable only, the unemployment rate. Principal components were calculated for the fifteen quarters extending from the third quarter of 1970 to the first quarter of 1974, as well as for intervals of 16, 17, 18, and 19 quarters (all of which have the same initial quarter). The results are summarized in Table 5.

It should be noted that in these regressions the principal components were normalized to sum to 1, the constants were constrained to be zero, and the regression coefficients were estimated by least squares in order to satisfy the constraint that the regression coefficients sum to 1.

Table 5

Residuals for Principal Component Regression Compared to Forecast Errors for UR

Forecaster	1970.3-1974.1		1970.3-1974.2		1970.3-1974.3		1970.3-1974.4		1970.3-1975.1		1970.3-1975.2	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
DRI	0.013	0.155	-0.025	0.214	-0.024	0.208	0.011	0.249	0.047	0.289	0.035	0.287
Fair	-0.093	0.183	-0.113	0.193	-0.094	0.201	-0.044	0.287	0.011	0.368	0.010	0.358
BEA	0.013	0.164	-0.006	0.177	-0.012	0.173	0.028	0.237	0.026	0.231	0.030	0.225
Composite Based on Principal Components	-0.017	0.121	-0.040	0.158	-0.036	0.154	0.011	0.225	0.029	0.225	0.030	0.219

Because these composites yield a smaller standard error than the individual forecasts, it might be that such poolings of the individual forecasters would lead to forecasts which are superior to that of any forecaster. The authors plan further investigation of this issue.

References

[1] Cooper, J. P. and Nelson, C. R. "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 7 (Feb. 1975): 1-32.

[2] Duggal, V. G.; Klein, L. R.; and McCarthy, M. D. "The Wharton Model Mark III: A Modern IS-LM Construct." International Economic Review 15 (Oct. 1974): 572-94.

[3] Eckstein, O.; Green, E. W.; and Sinai, A. "The Data Resources Model: Uses, Structure, and Analysis of the U.S. Economy." International Economic Review 15 (Oct. 1974): 595-615.

[4] Fromm, G., and Klein, L.R. "A Comparison of Eleven Econometric Models of the United States." American Economic Review 63, 2 (May 1973): 385-93.

[5] Fromm, G., and Klein, L. R. "The NBER/NSF Model Comparison Seminar: An Analysis of Results." Annals of Economic and Social Measurement 5 (1976): 1-28.

[6] Haitovsky, Y.; Treyz, G.; and Su, V. Forecasts with Quarterly Macroeconometric Models. National Bureau of Economic Research, 1974.

[7] Hart, B. I. "Tabulation of the Probabilities of the Ratio of the Mean Square Difference to the Variance." Annals of Mathematical Statistics 13 (1942): 207-14.

[8] Hirsch, A. A.; Grimm, B. T.; and Narasimham, G. V. L. "Some Multiplier and Error Characteristics of the BEA Quarterly Model." International Economic Review 15 (Oct. 1974): 616-31.

[9] McNees, S. K. "An Evaluation of Economic Forecasts." New England Economic Review, [Federal Reserve Bank of Boston] (Nov./Dec. 1975):2-39.

[10] Nelson, C. R. "The Prediction Performance of the FRB-MIT-PENN Model of the U.S. Economy. American Economic Review, Dec. 1972, pp. 902-16.

[11] Newbold, P., and Granger, C. W. J. "Experience with Forecasting Univariate Time Series and the Combination of Forecasts." Journal of the Royal Statistical Society, A., 137 (1974): 131-46.

[12] von Neumann, J. "Distribution of the Ratio of the Mean Square Successive Differences to the Variance." Annals of Mathematical Statistics 12 (1941): 367-95.