International Journal of Forecasting 5 (1989) 417-426
North-Holland

Econometric GNP forecasts: Incremental information relative to naive extrapolation

Robert T. CLEMEN
University of Oregon, Eugene, OR 97403, USA

John B. GUERARD, Jr.
Drexel, Burnham, Lambert, Chicago, IL 60606, USA

Abstract: Recent studies of macroeconomic forecasts have focused primarily on the relative performance of individual forecasts and combinations thereof. We suggest that these forecasts be evaluated in terms of the incremental information that they provide relative to a simple extrapolation forecast. Using a Bayesian approach, we measure the incremental information contained in econometric forecasts of U.S. GNP relative to a random-walk-with-drift time series forecast. The results indicate that (1) substantial incremental gains can be obtained from econometric GNP forecasts for the current quarter, but that these gains decrease rapidly as the forecast horizon increases, and (2) after one econometric forecast has been consulted, subsequent such forecasts add little information.

Keywords: Forecast evaluation, Combination of forecasts, GNP forecasts, Econometric models.

Introduction

Forecast evaluation research has typically been concerned with the performance of forecasts relative to each other. This has been true of studies evaluating both individual and combined econometric forecasts (see, e.g., McNees, 1986; Zarnowitz, 1979, 1984; Clemen and Winkler, 1986). However, perhaps more pertinent to decision makers is the issue of whether a forecast provides additional information beyond what is already known. In the case of econometric forecasting, some information is available to forecast users at minimal cost; time series forecasting software for microcomputers is now inexpensive and readily available, permitting decision makers to generate extrapolation forecasts at the touch of a button. Do econometric forecasts provide information beyond that contained in a simple extrapolation forecast? This is the question that we seek to answer in this study.

We taken a Bayesian approach in addressing this question. Following Clemen and Winkler (1985), we derive a measure of incremental information content which we apply to econometric forecasts relative to an extrapolation forecast (random walk with drift). In essence, the incremental information measure compares the extrapolation forecast with the combination of the extrapolation and econometric forecasts. This idea is in the spirit of Granger and Newbold's (1973, 1977) concept of 'conditional efficiency'. Granger and Newbold suggest that "if the variance of the combined forecast error is not significantly less than that of the forecast of interest, then the competing forecast would appear to possess no additional useful information" (1977, p. 283). However, instead of asking whether a forecast adds information, our Bayesian approach asks how much it adds, and we show how to find a Bayesian posterior distribution for the amount of information contained in a set of forecasts.

A closely related classical approach to forecast evaluation is that of Nelson (1972) and Cooper and Nelson (1975). Their approach involves a regression-like combination of forecasts and examination of the statistical significance of the estimated combination weights. More recently, Hendry and Richard (1983) have proposed the similar concept of 'encompassing' as a basis for forecast model comparison; a forecast which encompasses another is one which subsumes the other in terms of information. The null hypothesis that one forecast encompasses another can be tested empirically using historical data in a regression framework.

The article is organized as follows. In section 1, we describe the Clemen–Winkler model for measuring aggregate information, and we extend it to be operational for our situation. On the basis of this model, we derive an estimate of the incremental information contained in a set of forecasts relative to another forecast. Section 2 describes the GNP forecast data that we analyzed. (Clemen and Winkler (1986) analyzed the same econometric forecasts. Our analysis differs substantially from theirs in considering incremental information relative to the extrapolation forecast.) The results of the analysis are presented in section 3. Section 4 concludes with a discussion of the results, including implications for users of econometric forecasts as well as for researchers studying combined forecasts.

1. Methodology

Suppose that, at time t, a decision maker (DM) is interested in forecasting the quantity θ_{t+h}. Given one forecast of θ_{t+h}, the incremental information from a second forecast depends on the extent to which the two forecasts are based on overlapping information; the more the information overlaps, the less the second forecast brings to the combination. Here we outline and extend Clemen and Winkler's (1985) model for measuring the information contained in a number of dependent forecasts.

Represent the forecasts by the vector $f_t = (f_{1,t}, \ldots, f_{k,t})^{\mathrm{T}}$, where T indicates transposition, and with

$$\theta_{t+h} = f_{i,t} + \varepsilon_{i,t+h}, \qquad i = 1, 2, \ldots, k. \tag{1}$$

Following Winkler (1981), the vector of error terms $\varepsilon_{t+h} = (\varepsilon_{1,t+h}, \ldots, \varepsilon_{k,t+h})^{\mathrm{T}}$ is assumed to be multinormally distributed, with mean vector $(0, \ldots, 0)^{\mathrm{T}}$ and covariance matrix Σ_h. Furthermore, ε_i and ε_j are assumed to be independent for all $i \neq j$. (These assumptions imply that the forecast error variances and covariances are independent of θ_{t+h}, and that the k series of errors are white noise series.) If the DM has a diffuse prior for θ_{t+h} and hears forecasts f_t, her posterior distribution for θ_{t+h} will be normal with mean

$$\mu = \left(e^{\mathrm{T}} \Sigma_h^{-1} f_t\right) \sigma^{-2} \tag{2}$$

and variance

$$\sigma^2 = \left(e^{\mathrm{T}} \Sigma_h^{-1} e\right)^{-1}, \tag{3}$$

where e is a conformable vector of ones (Winkler, 1981).

In Clemen and Winkler's framework, the information contained in the k dependent forecasts is measured in terms of their collective impact on the posterior variance of θ_{t+h}. To accomplish this, the k dependent forecasts are compared to a number of hypothetical independent forecasts. If, instead of the dependent forecasts, the DM could obtain n independent and unbiased forecasts, each with error variance τ^2, her posterior variance for θ_{t+h} would be

$$\sigma_*^2 = \tau^2 / n. \tag{4}$$

To obtain the the number of independent forecasts that would be equivalent to the k dependent forecasts, set $\sigma_*^2 = \sigma^2$ and solve for n:

$$n(\Sigma_h, \tau) = \tau^2 \left(e^{\mathrm{T}} \Sigma_h^{-1} e\right). \tag{5}$$

The number of equivalent independent forecasts is a function of τ as well as all of the elements of Σ_h; hence, the indication of these as arguments of $n(\cdot, \cdot)$.

To consider incremental information, we can find the number of equivalent independent forecasts $[n(\Sigma_h^*, \tau)]$ for a subset of the forecasts as well as for the entire set. The difference between $n(\Sigma_h, \tau)$ and $n(\Sigma_h^*, \tau)$ represents the amount of incremental information contained in the additional forecasts, given that the DM already has consulted those forecasts contained in the subset. In particular, we are often interested in how much information $k-1$ forecasts add, given that the DM has already obtained one forecast. If the first forecast has error variance σ^{*2}, then we can calculate that the incremental information added by the $k-1$ forecasts is equivalent to $\sigma^{*2}(e^{T}\Sigma_h^{-1}e)-1$ independent forecasts, each having error variance σ^{*2}.

As it stands, the model requires that the covariance matrix be known and that the forecasters be unbiased. Biased forecasters present no major problem, as the forecast can be adjusted for a known or estimated bias. Having to estimate the covariance matrix poses a problem, however, especially when the forecasts are highly dependent. Small changes in the data can result in substantial changes in the estimated covariance matrix, combining weights and estimated posterior variance (Clemen and Winkler, 1986; Kang, 1986).

To operationalize the model, assume that the DM has observed n realizations of the random vector ε. She can use these data to update her prior beliefs regarding Σ_h before combining forecasts f_t. Let $X=(\varepsilon^1,\ldots,\varepsilon^n)$ denote the $k\times n$ data matrix. In this case, it is helpful to work in terms of the precision matrix Σ_h^{-1}. We assume that the DM has a diffuse prior distribution for Σ_h^{-1}:

$$f(\Sigma_h^{-1}) \propto \left|\Sigma_h^{-1}\right|^{-(k+1)/2}. \tag{6}$$

Discussion of this prior distribution and details of the following derivation are available from Press (1982). Although we do the analysis assuming a diffuse prior for Σ_h^{-1}, it is straightforward to incorporate informative prior beliefs if those beliefs can be modeled adequately with a Wishart distribution for Σ_h^{-1}.

Observation of the data matrix X results in the posterior distribution

$$f(\Sigma_h^{-1} \mid X) \propto \left|\Sigma_h^{-1}\right|^{(n-k-1)/2} \times \exp\left\{-\operatorname{tr} XX^{T}\Sigma_h^{-1}/2\right\}, \tag{7}$$

which is a Wishart distribution with n degrees of freedom and scale matrix XX^{T}. For the equivalent number of independent forecasts, we are interested in the distribution of $e^{T}\Sigma_h^{-1}e$. For convenience, let $z=e^{T}\Sigma_h^{-1}e$, and $z_o=e^{T}(XX^{T})^{-1}e$. Standard results regarding transformations of Wishart matrices give

$$f(z \mid X)=\left[z^{(n-2)/2}\exp\{-z/2z_o\}\right] / 2z_0^{n/2}\Gamma(n/2), \tag{8}$$

which is a chi-square distribution for z with n degrees of freedom and scale parameter z_0. An estimate for the number of equivalent independent forecasts is then given by the expected value of $n(\Sigma_h, \tau)$:

$$E\left[n(\Sigma_h, \tau)\right]=\tau^2 z_0 \mathrm{n}. \tag{9}$$

On the basis of the distribution for z, we can also calculate Bayesian credible intervals for the unknown equivalent number of independent forecasts $n(\Sigma_h, \tau)=\tau^2 z$ for various values of τ.

We can find $E[n(\Sigma_h^*, \tau)]$ for a subset of the forecasts, and the difference $E[n(\Sigma_h, \tau)]-E[n(\Sigma_h^*, \tau)]$ is an estimate of the incremental information added by the forecasts not contained in the subset. When the subset contains one forecast, the estimate of the equivalent number of independent experts is simply τ^2/s^2, where s^2 is the sample error variance for the one forecast. Thus, the incremental information added by the additional $k-1$ forecasts would be estimated as $\tau^2(z_0\mathrm{n}-1/s^2)$.

2. Data

Wharton Econometrics (Wharton), Chase Econometrics (Chase), Data Resources, Inc. (DRI), and the Bureau of Economic Analysis (BEA) make quarterly forecasts of many economic variables. We used their level forecasts of nominal GNP (1970–1982) (obtained directly from Wharton and BEA and from the *Statistical Bulletin* published by the Conference Board for Chase and DRI) to construct growth rate forecasts in percentage terms, and we calculated forecast errors (actual − forecast) as deviations from actual growth as determined from GNP reported in *Business Conditions Digest*. Forecasts with four different horizons ($h=1$, 2, 3, and 4 quarters) were analyzed. For example, the one-quarter GNP forecast predicts the percentage change for the current three-month period, and the four-quarter forecast predicts the percentage change for the three-month period four quarters in the future.

We constructed a simple extrapolation forecast (EXT) of the GNP growth rate using a random walk with drift model (an ARMA(0, 1) process). Using values of GNP growth from 1959 through

Exhibit 1
Data sets for GNP forecasts. The labeling indicates year and quarter: 1971.3 refers to the third quarter of 1971.

Horizon	Estimation sample		Evaluation sample	
	Period	n	Period	n
1 quarter	1971.1–1979.4	36	1980.1–1982.2	9 [a]
2 quarters	1971.1–1979.4	36	1980.1–1982.3	10 [a]
3 quarters	1971.2–1979.4	35	1980.1–1982.4	11 [a]
4 quarters	1971.3–1979.4	33 [a]	1980.1–1983.1	12 [a]

[a] Data from 1977.1 and 1981.4 were missing.

1970, we initially identified and estimated the following model over 48 quarterly observations:

$$\underset{\text{[standard error]}}{\text{GNP growth}} = \underset{[0.006]}{0.064} + (1 + \underset{[0.151]}{0.064}B)a_t, \quad (10)$$

Ljung–Box statistic, $\chi^2_{22} = 10.33$.

Although our estimated coefficient for the first-order moving average operator was not significantly different from zero, the random walk with drift model is readily accepted in economics (Granger and Newbold, 1977). We re-estimated the model for each additional quarter and then calculated forecasts over the relevant forecast horizons.

The GNP forecast data were split into estimation and evaluation samples. All model fitting and estimation were done on pre-1980 data, while data from 1980 on were used to evaluate individual forecasts as well as various forecast combination techniques. Exhibit 1 gives the details for the four data sets, including sample sizes for both estimation and evaluation samples.

3. Results

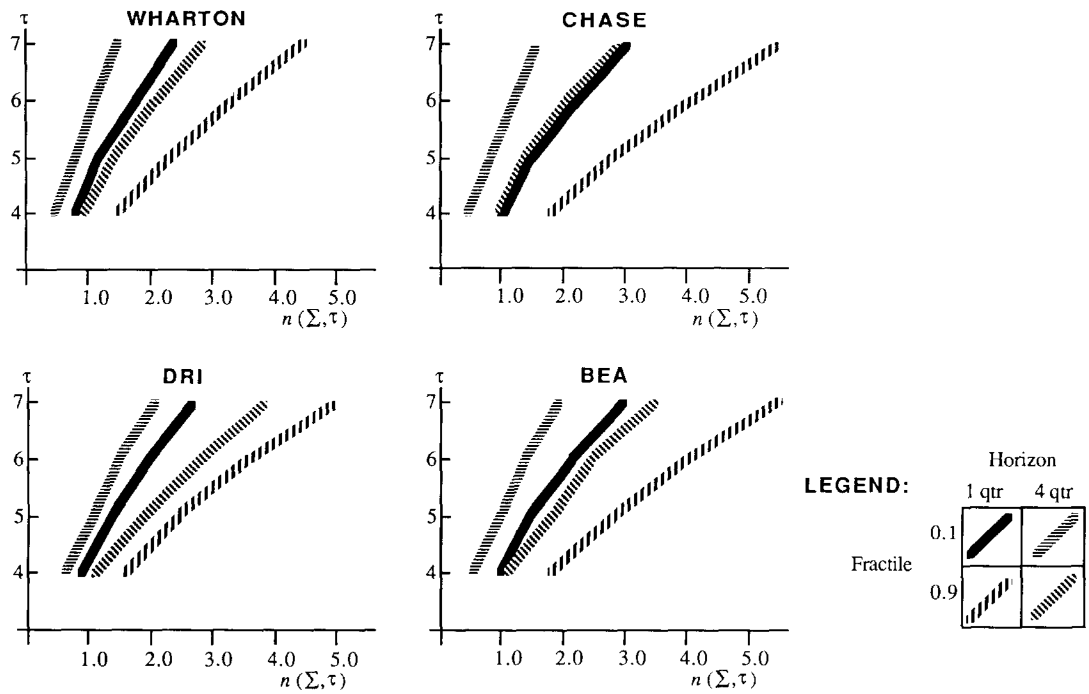
Exhibit 2 shows a number of summary statistics for Wharton, Chase, DRI, BEA, and EXT during the estimation period as well as information relating to the value of the forecasts. The mean errors of the forecasts are of interest since the model developed in section 2 assumes that the forecast are unbiased. The mean errors are all positive except for Wharton's current-quarter forecast, indicating that in general the forecasts tended to underestimate nominal GNP for the estimation period. However, none of the means are significantly different from zero, and Clemen and Winkler (1986) reported that adjusting for estimated forecast bias resulted in poorer performance in their post-sample evaluation. In our analysis, EXT also performs better if it remains unadjusted. Thus, the assumption of unbiasedness appears to be reasonable.

The standard deviations and correlations with EXT reported in exhibit 2, calculated under the assumption that the forecasts are unbiased, provide the necessary information for the analysis from section 1. For example, the posterior expected value of $n(\Sigma, \tau)$ shown in exhibit 2 estimates the amount of information in the combination of EXT with each of the econometric forecasts when τ equals the estimated standard deviation of EXT. For example, the current-quarter forecast of GNP made by Wharton is worth approximately 92% of an independent forecast with error standard deviation 5.24. The incremental information available from each econometric forecast decreases dramatically as the horizon increases. For Wharton and Chase, the pattern is the same; the econometric forecast, when com-

Exhibit 2
Summary statistics for individual forecasts based on the estimation samples. The standard deviations and correlations have been calculated under the assumption that the forecasts are unbiased.

	Wharton	Chase	DRI	BEA	EXT
1-quarter horizon					
Mean error	−0.17	0.99	0.87	0.71	3.00
Error standard deviation	3.80	3.44	3.54	3.41	5.24
Error correlation with EXT	0.66	0.74	0.74	0.69	–
Expected $n(\Sigma, \tau = 5.24)$	1.92	2.36	2.14	2.33	–
2-quarter horizon					
Mean error	0.44	1.19	0.52	0.52	3.31
Error standard deviation	4.40	4.57	3.81	3.80	5.37
Error correlation with EXT	0.78	0.80	0.87	0.81	–
Expected $n(\Sigma, \tau = 5.37)$	1.50	1.39	2.19	2.06	–
3-quarter horizon					
Mean error	0.76	1.12	0.81	0.76	3.11
Error standard deviation	4.24	4.92	4.25	4.06	5.15
Error correlation with EXT	0.84	0.80	0.85	0.85	–
Expected $n(\Sigma, \tau = 5.15)$	1.48	1.17	1.48	1.63	–
4-quarter horizon					
Mean	1.29	1.19	0.75	0.70	3.04
Error standard deviation	5.72	5.07	4.29	4.53	5.15
Error correlation with EXT	0.72	0.84	0.82	0.93	–
Expected $n(\Sigma, \tau = 5.15)$	1.07	1.11	1.44	1.31	–

Exhibit 3. Fractiles of the posterior distribution for $n(\Sigma, \tau)$ for the combination of each econometric forecast with EXT. Comparison of the fractiles for the current-quarter and four-quarter forecasts shows the extent to which the incremental information contained in the econometric forecast decreases with the longer forecast horizon.



bined with EXT, adds almost no information for the four-quarter forecast. DRI and BEA fare slightly better in the longer horizons, adding information approximately equivalent to 44% and 31% of an independent forecast, respectively, for the fourquarter forecast.

For the one- and four-quarter horizons, exhibit 3 shows 0.1 and 0.9 fractiles of the posterior distribution for $n(\Sigma, \tau)$, with values of τ from 4 to 7. Obviously, the smaller the value of τ, the lower $n(\Sigma, \tau)$ for the combined forecasts. Exhibit 3 demonstrates graphically the extent to which the information from the combined forecasts deteriorates from the one-quarter horizon to the four-quarter horizon, and reinforces the conclusion that the decay appears to be more substantial for Wharton and Chase than for DRI and BEA. Since the standard deviation of EXT is roughly the same for every horizon, this also indicates that the incremental information from the econometric forecasts is much less for the longer horizon than for the shorter one.

Exhibit 4 shows the estimated correlation matrices for all five forecasters. By combining this information with the standard deviations provided in exhibit 2, we can calculate covariance matrices, and from those the estimated incremental information of the four econometric forecasters combined with EXT. When the econometric forecasts are at their best, in forecasting for the current quarter, the econometric forecasts are collectively worth approximately 1.63 independent forecasts with the same standard deviation as EXT, while in the fourth quarter, they are estimated as being collectively worth only around one-half of an independent forecast. For each of the four forecast horizons, exhibit 5 shows the 0.1 and 0.9 fractiles of the posterior distribution for $n(\Sigma, \tau)$ for values of τ from 4 to 7. Again, the information in the econometric forecasts appears to deteriorate with the increased forecast horizon.

Comparing results in exhibits 2 and 4 also allows us to examine the incremental information of additional econometric forecasts, given that the

Exhibit 4
Error correlation matrices and expected number of equivalent independent forecasts from combining all five individual forecasts.

	Wharton	Chase	DRI	BEA
1-quarter horizon				
Chase	0.84			
DRI	0.85	0.86		
BEA	0.82	0.85	0.83	
EXT	0.66	0.74	0.74	0.69
Expected $n(\Sigma,\ \tau = 5.24) = 2.63$				
2-quarter horizon				
Chase	0.84			
DRI	0.86	0.90		
BEA	0.81	0.85	0.93	
EXT	0.67	0.77	0.76	0.72
Expected $n(\Sigma,\ \tau = 5.37) = 2.13$				
3-quarter horizon				
Chase	0.88			
DRI	0.92	0.92		
BEA	0.91	0.92	0.94	
EXT	0.77	0.76	0.78	0.78
Expected $n(\Sigma,\ \tau = 5.15) = 1.72$				
4-quarter horizon				
Chase	0.79			
DRI	0.82	0.92		
BEA	0.82	0.92	0.91	
EXT	0.70	0.80	0.75	0.83
Expected $n(\Sigma,\ \tau = 5.15) = 1.50$				

DM already has EXT and one econometric forecast. The potential gains are small indeed, particularly if the DM had chosen the best of the econometric forecasts as the first one to consult. The greatest potential appears to be in the current-quarter forecast, where adding Wharton, DRI, and BEA to the combination already consisting of Chase and EXT results in a net increase of approximately 0.27 independent forecast. For the two-quarter forecast, it would appear that the DM would be better off with only DRI and EXT, ignoring the remaining three econometric forecasts. In the three- and four-quarter forecasts, consulting more than one econometric forecast apparently adds very little information.

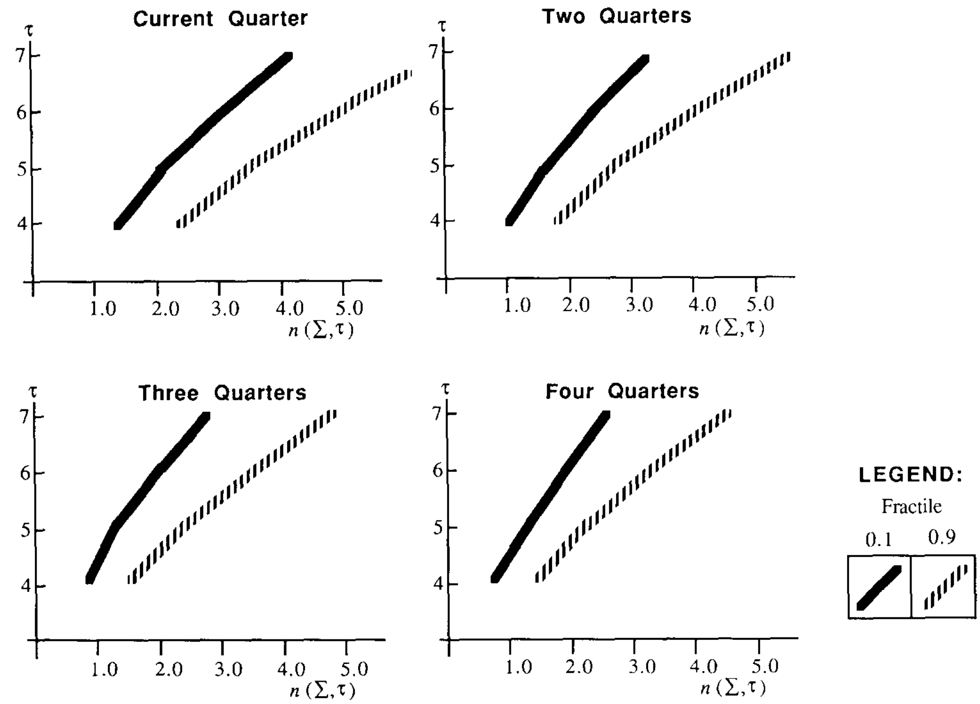
The results discussed above provide an a priori analysis of the forecasts. That is, on the basis of the estimation data, we have calculated credible intervals and expected $n(\Sigma,\ \tau)$ that describe an hypothetical DM's beliefs relating to the *anticipated* effect of combining one or more econometric forecasts with EXT. We can use our evaluation sample to perform an ex post analysis and compare anticipated with realized forecasting performance. Exhibit 6 gives root mean squared errors (RMSE) for the individual forecasts as well as for various combinations of each econometric forecast with EXT. The primary combination technique is the posterior mean from (2), which amounts to a weighted combination of the forecasts with the weights adding to one. For convenience, we will refer to this combination as the 'model'. Although this combination is consistent with our a priori measurement of incremental value, we also include the arithmetic average of the two forecasts as well as an OLS regression combination. The average has been shown to be a robust combination technique for a wide variety of situations (see, e.g., Makridakis and Winkler, 1983), while the regression approach provides a way to account for bias present in the forecasts (Granger and Ramanathan, 1984).

For the current-quarter forecast, the econometric forecasts clearly contained substantial incremental information relative to EXT, a result consistent with exhibit 2. In all cases, the regression combination performed the best, providing substantial improvement over EXT by itself and slight improvement over the econometric forecast. Our model and the simple average performed about the same.

For the two-quarter forecast, the results in exhibit 6 are somewhat mixed. The Wharton and DRI forecasts evidently contained some incremental information, as reflected by the performance of the individual forecasts compared to the regression and simple average combinations. For these forecasts, our model provided no improvement over EXT used alone. The Chase and BEA forecast combinations showed no performance improvement over EXT regardless of the combination technique. For DRI and Wharton, these results agree with the results in exhibit 2. However, for Chase and BEA we might have expected each of the econometric forecasts to add some information, with Chase adding the least.

The results for the three- and four-quarter forecasts are similar. In every case, the econometric forecast does not help. In fact, for both horizons, the DM would be better off ignoring the econometric forecasts and using only EXT. No con-

Exhibit 5. Fractiles of the posterior distribution for $n(\Sigma, \tau)$ for the combination of all four econometric forecasts with EXT.



sistent pattern appears evident when comparing the performance of the individual forecasts and the combinations. These results are also inconsistent with exhibit 2. We might have expected Wharton, DRI, and BEA to add information in the three-quarter forecast, and DRI and BEA to do so in the fourth quarter.

Exhibit 7 shows performance measures when the three combination techniques are applied to all five forecasts. For the current-quarter forecast, combining all of the forecasts, regardless of the approach, provides substantial improvement over EXT. This is consistent with exhibit 4. On the other hand, the results for the two-, three- and four-quarter horizons show that EXT is a strong performer when compared with the combinations These results are inconsistent with exhibit 4, on the basis of which we anticipated some improvement from the combinations. For all three horizons, the regression approach performs best among the combining techniques, but it provides only a slight improvement over EXT and only in the two-quarter horizon. In the three- and four-quarter horizons, EXT outperforms all of the combinations.

Why are the results from the evaluation sample inconsistent with the results from the estimation sample? It appears that the difference is due primarily to an inauspicious choice of estimation and evaluation samples. Both the relative accuracy and possibly the bias of the forecasting methods change from one period to the next. While the econometric forecasts consistently outperformed EXT (in terms of error standard deviation) during the estimation period, this is not the case during the evaluation period. It is evident from the RMSE's and error standard deviations in exhibit 6 that EXT was more accurate than any of the econometric forecasts during the evaluation period for the three- and four-quarter horizons. The forecasts' mean errors (also reported in exhibit 6) indicate that during the evaluation period the econometric forecasts tended to overestimate, whereas during the estimation period they tended to underestimate.

In light of these results, one may wonder why

Exhibit 6
Performance measures for the individual forecasts and for combinations of individual econometric forecasts with EXT. For the individual forecasts, the mean errors (with standard deviation in parentheses) are shown. Root mean squared error (RMSE) is also given for the individual forecasts and for combinations of each econometric forecast with EXT. These error statistics are calculated for the evaluation period only.

	Wharton	Chase	DRI	BEA	EXT
1-quarter horizon					
Mean error	−0.33	1.41	0.36	0.37	−1.07
(Standard deviation)	(5.28)	(5.62)	(4.94)	(5.29)	(6.22)
RMSE					
Individual forecast	4.99	5.48	4.67	5.00	5.96
Combined with EXT:					
Simple average	5.36	5.51	5.14	5.01	–
OLS regression	4.72	5.31	4.58	4.68	–
Model	5.03	5.54	4.63	5.04	–
2-quarter horizon					
Mean error	−1.90	−0.08	−0.15	−2.08	−0.14
(Standard deviation)	(7.00)	(8.43)	(6.86)	(7.56)	(7.00)
RMSE					
Individual forecast	6.91	7.99	6.51	7.47	6.64
Combined with EXT:					
Simple average	6.59	7.10	6.42	6.85	–
OLS regression	6.38	7.14	6.34	6.97	–
Model	6.83	7.75	6.82	7.81	–
3-quarter horizon					
Mean error	−4.22	−3.11	−2.91	−3.42	−0.76
(Standard deviation)	(7.50)	(8.82)	(7.74)	(7.77)	(6.96)
RMSE					
Individual forecast	8.31	8.90	7.86	8.08	6.46
Combined with EXT:					
Simple average	7.20	7.49	7.00	7.08	–
OLS regression	7.64	8.44	7.91	9.11	–
Model	8.43	7.91	8.02	8.49	–
4-quarter horizon					
Mean error	−3.85	−3.47	−3.53	−3.09	0.09
(Standard deviation)	(7.60)	(7.90)	(7.68)	(7.78)	(6.85)
RMSE					
Individual forecast	8.20	8.36	8.28	8.28	6.32
Combined with EXT:					
Simple average	7.08	7.10	7.06	7.03	–
OLS regression	6.86	6.59	7.41	6.56	–
Model	6.76	7.21	8.19	9.26	–

we chose the estimation and evaluation periods were chosen as we did. Our model and the regression combination were fit using 33–36 observations. Regression analysis generally should be performed with at least 30 observations (Makridakis and Wheelwright, 1979); given the use of four econometric forecast variables, the combined models are estimated with approximately 30 degrees of freedom. It is doubtful that the research design could be significantly altered without creating regression estimations plagued by relatively few degrees of freedom.

Exhibit 7
Performance measures (RMSE) for combined forecasts. In each case all five individual forecasts are combined using the indicated technique.

Horizon (quarters)	Combining method		
	Average	Regression	Model
1	4.89	4.85	4.84
2	6.88	6.14	6.60
3	7.76	6.60	7.68
4	7.68	6.49	7.83

4. Conclusion

Our analysis of econometric forecasts of GNP relative to a simple extrapolation forecast provides some interesting insights. Perhaps most interesting are our results, from both estimation and evaluation samples, demonstrating clearly that the econometric forecasts add information for the current-quarter forecast. Also, we have shown that the incremental information contained in the econometric forecasts decreases substantially with an increased horizon. Our results in this regard are somewhat mixed. The a priori analysis would indicate that the econometric forecasts should have some information to add, albeit slight, for the longer-term forecasts. On the other hand, the results from our evaluation sample seem to indicate that, if EXT is available, the econometric forecasts could be ignored in forecasting GNP over several quarters. Finally, our results regarding combinations of all of the forecasts demonstrate that, if a DM already has access to an extrapolation forecast and one econometric forecast of GNP, consulting more econometric forecasts will provide little incremental information.

Further research might include the use of alternative extrapolation or other 'naive' forecasts. We chose the Box–Jenkins forecast simply as a good representative of extrapolation forecasts, and because of its acceptance by economists. Also, it would be interesting to evaluate the econometric forecasts relative to a combination of several automatic extrapolation forecasts. Results of Makridakis and Winkler (1983) on averaging many time series forecasts suggest that with many such forecasts included, the incremental value of the econometric forecasts may be very small indeed.

Our original premise was that a decision maker could easily generate an extrapolation forecast of GNP and, having done so, might be concerned about the additional information provided by an econometric forecast. A perceptive reviewer pointed out that the reverse may also hold. That is, a company may subscribe to an econometric forecasting service for any of a number of reasons. If this is the case, our decision maker may be able to obtain the econometric forecast at no marginal cost and then wonder about the incremental information contained in an extrapolation forecast. Statistics from exhibit 2 can be used to estimate the incremental information contained in EXT, given that the decision maker has one of the econometric forecast in hand. The incremental value of EXT in this situation is indeed low, as one might expect.

Our analysis has demonstrated how a DM can use information concerning the variances and covariances of forecast errors to make inferences about the incremental information contained in a number of forecasts. The model provides a satisfactory way for the DM to incorporate data regarding the covariance matrix in order to make an a priori assessment of the information she may obtain by consulting additional forecasts. This kind of analysis is particularly appealing in forecast combination situations, where decision makers may have access to relatively little data.

Should our approach replace the use of common error summary statistics (viz., mean squared error, mean absolute error) in forecast evaluation? We view our model as complementary rather than as a replacement for other forecast performance measures. In cases where insufficient data are available to first fit a combination model and then calculate error statistics in an independent sample, our approach would be worthwhile and could be used alongside performance measures derived from the fitting data. Given that our approach is Bayesian, it would appear to be more appropriate for decision makers choosing among forecasters.

Our model may be extended in a number of ways. First, as mentioned above, it is a straightforward matter to incorporate prior beliefs regarding Σ_h into the analysis providing those beliefs can be modeled adequately with a Wishart density for Σ_h^{-1}. Second, a straightforward extension would allow consideration of biased forecasts of the form $f_i = a_i + b_i\theta + \varepsilon_i$, for (possibly unknown) constants a_i and b_i. Finally, the model could be developed in a classical framework; such an approach would permit a test of the null hypothesis that a given forecast adds no incremental information. However, given current interest of forecast evaluation and selection, we find the Bayesian approach with its measure of information content more compelling than the classical approach.

Acknowledgements

We thank George Jaszi of the Bureau of Economic Analysis and Donald Straszheim of Wharton Econometrics who provided the forecasts of their respective econometric models. Frank Diebold, Allan Murphy, Robert Winkler and anonymous referees provided helpful comments on a previous draft of the paper. This research was supported in part by the National Science Foundation under Grant No. IST 8600788. An earlier version of this paper was presented at the ORSA/TIMS meeting in New Orleans (May, 1987).

References

Clemen, R.T. and R.L. Winkler, 1985, "Limits for the precision and value of information from dependent sources", *Operations Research*, 33, 427–442.

Clemen, R.T. and R.L. Winkler, 1986, "Combining economic forecasts", *Journal of Business and Economic Statistics*, 4, 39–46.

Cooper, J.P. and C.R. Nelson, 1975, "The ex ante prediction performance of the St. Louis and F.R.B.-M.I.T.-Penn. econometric models and some results on composite predictors", *Journal of Money, Credit, and Banking*, 7, 1–32.

Figlewski, S. and T. Urich, 1983, "Optimal aggregation of money supply forecasts: Accuracy, profitability and market efficiency", *Journal of Finance*, 28, 695–710.

Granger, C.W.J. and P. Newbold, 1973, "Some comments on the evaluation of forecasts", *Applied Economics*, 5, 35–47.

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

Granger, C.W.J. and R. Ramanathan, 1984, "Improved methods of combining forecasts", *Journal of Forecasting*, 3, 197–204.

Hendry, D.F. and J.-F. Richard, 1983, "The econometric analysis of economic time series", *International Statistical Review*, 51, 111–148.

Kang, H., 1986, "Unstable weights in the combination of forecasts", *Management Science*, 32, 683–695.

Makridakis, S. and S.C. Wheelwright, 1979, "Forecasting: Framework and overview", in: S. Makridakis and S.C. Wheelwright, Eds., *Forecasting* (North-Holland, Amsterdam).

Makridakis, S. and R.L. Winkler, 1983, "Averages of forecasts: Some empirical results", *Management Science*, 29, 987–996.

McNees, S.K., 1986, "Forecasting accuracy of alternative techniques: A comparison of U.S. macroeconomic forecasts", *Journal of Business and Economic Statistics*, 4, 5–15.

Nelson, C.R., 1972. "The prediction performance of the F.R.B.-M.I.T.-PENN model of the U.S. economy", *American Economic Review*, 62, 902–917.

Press, S.J., 1982. *Applied Multivariate Analysis*, 2nd ed. (Krieger, Malabar, FL).

Winkler, R.L., 1981, "Combining probability distributions from dependent information sources", *Management Science*, 27, 479–488.

Zarnowitz, V., 1979. "An analysis of annual and multiperiod quarterly forecasts of aggregate income, output and the price level", *Journal of Business*, 52, 1–33.

Zarnowitz, V., 1984. "The accuracy of individual and group forecasts from business outlook surveys", *Journal of Forecasting*, 3, 11–26.

Biographies: Robert T. CLEMEN holds a Ph.D. from Indiana University and is currently Associate Professor of Decision Sciences at the University of Oregon. His research interests are primarily in the areas of decision theory and decision analysis, especially the use of expert information, consensus, and combining forecasts.

John B. GUERARD, Jr., is a graduate of Duke University and earned his Ph.D. in Finance from the University of Texas, Austin. He is the co-author of the *Handbook of Financial Decision Making* (Probus Publishing, 1989). He has published articles on composite modeling, R&D management, strategic planning, robust regression, and mergers and acquisitions.