Journal of Forecasting, Vol. 2, 211–223 (1983)

On the Usefulness of Macroeconomic Forecasts As Inputs to Forecasting Models

RICHARD ASHLEY
Department of Economics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, U.S.A.

ABSTRACT

A forecasting model for y_t based on its relationship to exogenous variables (e.g. x_t) must use $\hat{x}_t$, the forecast of x_t. An example is given where commercially available $\hat{x}_t$'s are sufficiently inaccurate that a univariate model for y_t appears preferable. For a variety of types of models inclusion of an exogenous variable x_t is shown to worsen the y_t forecasts whenever x_t must itself be forecast by $\hat{x}_t$ and $\mathrm{MSE}(\hat{x}_t) > \mathrm{Var}(x_t)$. Tests with forecasts from a variety of sources indicate that, with a few notable exceptions, $\mathrm{MSE}(\hat{x}_t) > \mathrm{Var}(x_t)$ is common for macroeconomic forecasts more than a quarter or two ahead.

Thus, either:

(a) available medium range forecasts for many macroeconomic variables (e.g. the GNP growth rate) are not an improvement over the sample mean (so that such variables are not useful explanatory variables in forecasting models), and/or

(b) the suboptimization involved in directly replacing x_t by $\hat{x}_t$ is a luxury that we cannot afford.

KEY WORDS Box–Jenkins Bivariate Commercial forecasting Transfer function

On average, a multivariate model will yield better conditional forecasts than a univariate model, provided that the chosen explanatory variables do matter and that both models are well specified and estimated. The multivariate model's superiority increases as the forecast horizon lengthens. This is because it is well known that the univariate model's forecasts are actually just extrapolations from the sample data that decay to the sample mean. Moreover, if sufficiently accurate forecasts are available for the explanatory variables, then the multivariate model retains its superiority in unconditional forecasting as well.

One might well ask, what constitutes sufficient accuracy in forecasting an explanatory variable? In particular, suppose that this explanatory variable is the growth rate in U.S. GNP. Considerable effort has been expended in constructing models to forecast this particular growth rate. Is it safe to

The Gunning Fog Index for this paper is less than 15.

0277–6693/83/030211–13$01.30

Received October 1981
Revised November 1982

assume that available forecasts of the growth rate in U.S. GNP are sufficiently accurate that a well-constructed bivariate model based on U.S. GNP will produce better short-to-medium range forecasts than a univariate (Box–Jenkins) model?

This last question is examined below. The answer is 'no'.

At least, the answer is 'no' for the particular postsample forecasting period used for model validation below. These results are then generalized to other national variables driving multivariate models over other forecasting periods, by proving a general theorem. The theorem gives a criterion with which to evaluate *ex ante* whether a given exogenous variable is forecast so poorly that its inclusion will only worsen even a well-specified model's forecasting performance. The criterion is a practical one since it depends only on the variance of the variable and the MSE of the forecasts of it.

A wide variety of available macroeconomic forecasts, commercial and otherwise, are then examined using this criterion. For many variables the criterion is typically not satisfied by forecasts more than a quarter or two ahead. These results confirm that the specific bivariate forecasting results reported below were not a fluke. They also yield insight into which kinds of macroeconomic variables can be forecast well enough to serve as useful explanatory variables in forecasting models and which cannot.

A BIVARIATE EXAMPLE

A recent amendment to the Texas constitution requires the Legislative Budget Board to forecast the ratio of average Texas personal income over the next biennium (1981.4–1983.3) to average Texas personal income over the current biennium (1979.4–1981.3). This ratio yields a growth rate, the forecast of which is to be a constitutional upper bound on the growth rate of appropriations from State tax revenues not constitutionally dedicated.

Thus, tpy_t, the quarterly growth rate in nominal Texas personal income, must be forecast through 1983.3. Three kinds of forecasting models were prepared:

(a) a naive model (based on a constant growth rate assumption),
(b) a univariate (Box–Jenkins) model (based on tpy_t's own past), and
(c) two bivariate transfer function models (based on tpy_t's own past plus current and past values of gnp_t (in one case) and $uspy_t$ (in the other), where these are the growth rates in U.S. GNP and personal income, respectively).

The naive model was designed to provide benchmark forecasts for comparison purposes. The univariate model was not expected to perform well, since many-step-ahead forecasts were required. It was almost costless to produce, however, since tpy_t was pre-whitened using a univariate model as part of the bivariate modelling process.

The data on Texas personal income were obtained from the Texas Legislative Budget Board. The data on gnp_t and $uspy_t$ were taken from NBER tapes. All of the data used are listed in Ashley (1980), a paper which also describes the modelling process in detail and gives intermediate results. Briefly, each series was pre-whitened using Box–Jenkins models. Cross-correlograms were then calculated. These cross-correlograms suggested specific models to estimate.

The resulting models are quoted in Appendix I, which is available from the author on request. The naive model merely regresses tpy_t against a constant. The univariate model is an AR(4) process with an R^2 of 0.372. The bivariate-gnp model is an AR(3) with a contemporaneous term in gnp_t. The coefficient on gnp_t is 0.48 with a t statistic of 4.3; R^2 is 0.49. The bivariate-uspy model is an AR(2) with a contemporaneous term in $uspy_t$. The coefficient on $uspy_t$ is 1.16 with a t statistic of

10.5; R^2 is 0.71. The univariate and bivariate models passed all the appropriate diagnostic checks on their residuals. In particular the residuals appear to be white noise on both Ljung–Box statistics and inspection of the correlogram. In addition, the residuals from the bivariate models appear to be uncorrelated with past values of the explanatory variable.

Since the univariate and bivariate models passed their diagnostic checks, there was no point in considering more complex models of these types. It seems worth while to emphasize at this point that the Legislative Budget Board specifically requested forecasting models of this simple nature to complement the other types of forecasts (judgemental and econometric) available to them.

These models were all specified using the sample period 1960.1–1977.4; previous data were not of good quality. All models were estimated over the period 1961.2–1977.4 to allow for lags and differencing. This left eight quarters of tpy_t data available for postsample forecasting. That represents a rather small postsample period. More postsample data would have been reserved, except that the sample size (N = 67) was already rather small for the use of time series methods. Re-estimation of the final models using a smaller sample (e.g. 1961.2–1970.4) was contemplated. This course was rejected on the grounds that the data for 1971.1–1977.4 were used in obtaining the model specifications. Hence, forecasts for these quarters would not really be postsample forecasts and one of the models might thereby obtain an unfair advantage in the forecasting comparison.

The postsample forecasting results are presented in Exhibit 1. Root mean square forecast errors (RMSE) are given there for each model. Note that the tpy_t forecast for 1978.1 is a one-step-ahead forecast, the forecast for 1978.2 is a two-step-ahead forecast, etc. Thus, what is computed is the RMSE of a sequence of one to eight-step-ahead forecasts rather than of a sequence of one-step-ahead forecasts. Ideally, one would compute a sequence of one-step-ahead errors (successively updating the parameter estimates), then a similar sequence of two-step-ahead errors, and so on. That was not done here for several reasons. First, the postsample period is so small that little information would be obtained thereby about the longer horizon forecasts. (For example, only one eight-step-ahead forecast could be computed.) Second, the GNP and personal income forecasts available to me were insufficiently complete to support such a program; the approach used here required only one set of forecasts from each source. Finally, this mix of one to eight-step-ahead forecasts is fairly close to what is actually needed in forecasting the growth in Texas personal

		Exogenous data source			
Model	None	Actual historical data	Box–Jenkins forecast[a]	Chase Econometrics forecast	DRI forecast
Naive	0.013				
Univariate	0.009				
Bivariate gnp		0.006	0.012	0.012	0.009
Bivariate uspy		0.005	0.012	0.012	0.010
Econometric (STEM)[b]		0.008	n.a.[c]	n.a.[c]	n.a.[c]

[a] The Box–Jenkins models used in forecasting uspy and gnp are quoted in Ashley (1980, Appendix 4). They were specified and estimated using data through 1977.1, the same data which were available to Chase Econometrics and DRI in July, 1977.

[b] STEM is the State of Texas Econometric Model, a simultaneous equation econometric model developed by the Texas Comptroller's Office.

[c] Not available. STEM was re-estimated by the Comptroller's office at my request using only pre-1978 data. Forecasts for 1978.1–1979.4 were prepared by them using historical values for the variety of exogenous variables required by STEM.

Exhibit 1. Root mean square error in postsample forecast

income over the next biennium. In fact this type of forecasting is quite common—e.g. Glickman (1977), pp. 156–159). Hence, these forecast errors were the most appropriate set for the task at hand.

In the column marked 'historical data', the actual (final) values of the relevant explanatory variable were used in preparing the forecasts. Note that the RMSE is substantially smaller for the two bivariate models than for the naive and univariate models.

This RMSE difference is statistically significant at the 5 per cent level for the bivariate-gnp model and at the 15 per cent level for the bivariate-uspy model using the asymptotic test developed in Ashley *et al.* (1980). The two bivariate models also provide apparently better tpy_t forecasts than the State of Texas Econometric Model (STEM), a 30 equation econometric model developed by the Comptroller's office and re-estimated by them over pre-1978 data. It can be concluded that the bivariate models for tpy_t were reasonably well specified and forecast quite successfully in the postsample period when the actual values of the exogenous variables are made available.

In practice, of course, the actual values of these exogenous variables are not available—they must be forecast. The Legislative Budget Board must begin producing provisional forecasts in the July preceding the new biennium. Consequently, the three sources of gnp_t and $uspy_t$ forecasts over the period 1978.1–1979.4 considered here were:

1. Box–Jenkins univariate models based on pre-1977.2 data,
2. Forecasts released by Chase Econometrics in July 1977, and
3. Forecasts released by Data Resources, Inc. in July 1977.

The postsample forecasting results for the two bivariate models using exogeneous variable forecasts from each of these sources are also given in Exhibit 2. In each case the root mean square forecast error rises dramatically when the actual values of an exogeneous variable are replaced by forecasts. Using mean square error instead of root square error would only accentuate this rise. In fact, the tpy_t forecasts from the univariate Box–Jenkins model look quite attractive by comparison. This univariate model would no doubt look even better for short-term forecasts.

Of course, one can always claim that 1978–79 was an unusually turbulent period, but this begs the question. The main reason for including a variable like gnp_t in the model is to help track tpy_t through periods where gnp_t varies substantially. Therefore turbulent periods are the bivariate model's natural opportunity to shine.

A GENERAL THEOREM

One might well ask, how bad do forecasts of the exogenous variable need to be in order to obtain results like those reported above? The following theorem gives a remarkably simple and general answer to this question:

Theorem

Suppose that x_t is an explanatory variable in a forecasting model for y_t. But x_t is itself unknown and must be as forecast; x_t is replaced by the forecast $\hat{x}_t$ where

$$x_t - \hat{x}_t = \eta_t$$

and η_t is uncorrelated with y_t. Then the MSE of a naive forecast for y_t, ignoring the relationship between y_t and x_t, will fail to exceed the MSE of a forecast based on the model whenever

$$\mathrm{MSE}(\hat{x}_t) > \mathrm{Var}(x_t)$$

Let P^t_{t+h} denote the forecast of A_{t+h} made in quarter t. Then the h-step-ahead forecast of the annualized compound growth rate of A_t is

$$100((P^t_{t+h}/P^t_{t+h-1})^4 - 1).$$

P^t_t is equivalent to what McNees calls A^*_t, the current estimate of A_t available to the forecaster.

Thus, the h-step-ahead forecast errors for the variables expressed in growth rates can be computed from

$$\varepsilon\%(t,h) = 100[(P^t_{t+h}/P^t_{t+h-1})^4 - (A_{t+h}/A_{t+h-1})^4]$$

where t runs from 1976I to h quarters prior to 1980III. Thus, there will ordinarily be a sample of 19 one-step-ahead forecast errors, 18 two-step-ahead forecast errors, etc. However, some forecasters (RSQE, BMARK, MHT, TG, UCLA, GE) chose not to produce forecasts P^t_{t+h} for some values of t and h. A listing of precisely which forecasts were omitted by each forecaster is available from McNees; a table listing the sample size for each forecaster and forecast horizon, h, is provided in Appendix III, which is available from the author on request. The statistics on GE forecasts eight quarters ahead were computed using a sample of only eight forecasts; otherwise the smallest sample size used was ten. In each case, the sample of A_t's used in calculating the Var (x_t) was appropriately adjusted; the A_t's were provided by McNees and fully described in McNees (1981, p. 8). The data used incorporated revisions made by McNees in March, 1982.

One can also consider the cumulative error in the h-step-ahead forecast of the annualized compound growth rate of A_t. These errors are computed from

$${}^{\text{cum}}\varepsilon\%(t,h) = 100[(P^t_{t+h}/A^*_t)^{4/h} - (A_{t+h}/A_t)^{4/h}].$$

These are the errors analysed in McNees' published tables. Aside from a minor notational difference, this formula differs from the one McNees quotes only because of a typographical error in his paper.

The level variables (the unemployment rate, the 90 day Treasury bill rate and the change in business inventories) were analysed in cumulative form only. Following McNees' procedure, the errors for these variables are calculated from

$$\begin{aligned}\varepsilon(t,h) &= [P^t_{t+h} - A^*_t] - [A_{t+h} - A_t]\\ &= [P^t_{t+h} + (A_t - A^*_t)] - A_{t+h}.\end{aligned}$$

Thus, the forecast actually considered is P^t_{t+h} adjusted for the data revision, $A_t - A^*_t$, and the variance used in calculating MSE$(\hat{x}_t)$/Var$(\hat{x}_t)$ is the sample variance of A_{t+h}.

There is good reason for considering the growth rate and the cumulative growth rate in, say, GNP but not the level of GNP. The reason is that the Var(x_t) grows over time for variables like the level of GNP. In contrast, the variance of the growth rate in GNP, the variance of the cumulative growth rate in GNP over any fixed number of periods, and the variance of variables like the unemployment rate are roughly constant over time. The theorem is equally applicable regardless of whether the Var(x_t) is constant over time or not. The empirical estimation of Var(x_t), however, is vastly simplified by a sample over which this variance is constant. In any case, the results on the cumulative growth rate of x_t are a nearly perfect substitute for results on the level of x_t. This is because the h-period-ahead forecasts of the level of x_t can be decomposed into the sum of the current value (which is known) plus the cumulative growth rate of over the next h periods.

Exhibits 3 and 4 describe the observations on the ratio MSE$(\hat{x}_t)$/Var(x_t) for the various variables and forecasters over the various horizons. Following McNees, the forecasters are grouped (early, mid or late) depending on whether they issue their forecasts in the early, middle or late part of the quarter. Thus, a one-quarter-ahead forecast from a 'late' forecaster may in actuality be issued only a month before the quarter in question begins.

At this point it must be emphasized that the theorem proved above does *not* indicate that substitution of $\hat{x}_t$ for x_t will improve forecasts if MSE $(\hat{x}_t)$/Var (x_t) is less than one. It merely shows that such substitution will certainly worsen forecasts if this ratio is greater than one. In fact, the proof itself indicates that the 'breakeven point' in general will be strictly less than one due to sampling error in the coefficient on x_t and due to proxying for x_t by the remaining variables in the model. Thus, the usefulness of x_t forecasts with MSE $(\hat{x}_t)$/Var (x_t) greater than, say, 0.7 is clearly in doubt.

	Forecast horizon							
Early	1	2	3	4	5	6	7	8
ASA	0.8	1.0	1.0	1.0	0.9	—	—	—
BEA	1.0	1.0	0.8	1.1	1.0	1.0	—	—
BMARK	1.4	1.3	1.3	1.3	—	—	—	—
CHASE	1.1	1.6	1.5	3.5	1.4	1.7	1.7	1.6
DRI	0.9	1.1	1.1	1.0	0.9	1.0	1.5	1.3
GSU	0.7	1.1	1.1	1.2	1.1	0.9	0.8	0.7
RSQE	0.9	0.9	1.0	0.9	1.5	1.1	0.9	—
WEFA	1.0	0.9	1.0	1.1	0.9	0.8	0.8	0.9
Mid								
KEDI	0.7	1.4	1.2	1.0	0.7	0.8	1.9	1.1
MGT	0.8	1.5	1.2	1.1	1.0	—	—	—
TG	0.6	1.0	1.0	1.4	0.7	0.8	0.9	—
UCLA	0.6	1.4	1.0	1.2	0.9	0.9	—	—
WEFA	0.5	1.0	0.9	1.1	0.8	0.8	0.8	0.8
Late								
CHASE	0.4	1.7	1.8	1.7	1.4	1.7	1.7	1.6
DRI	0.3	1.3	1.0	0.9	1.1	1.2	1.2	1.3
GE	0.6	1.1	0.8	1.4	1.0	1.8	1.3	1.1

Exhibit 3. GNP growth rate MSE $(\hat{x}_t)$/Var (x_t)

The first variable to look at is the growth rate of GNP. As Exhibit 3 shows, the value of MSE $(\hat{x}_t)$/Var (x_t) is nearly always close to or greater than one, the only exceptions being 'late' one-step-ahead forecasts. Thus, it can be concluded that the Texas personal income model results were not a fluke—the general theorem shows that GNP is forecast so poorly that direct inclusion of its growth rate as a driving variable in a forecasting model is in fact likely to worsen the model's forecasts.

The results on the other variables are given in Appendices IV–XV which are available from the author. Owing to space limitations these appendices are summarized here by Exhibit 4. Exhibit 4 gives a four digit integer characterizing the results on each forecaster's ability to forecast a given variable over this period. The four digit integer, symbolized as *abcd*, is to be interpreted as follows. The ratio MSE $(\hat{x}_t)$/Var (x_t) was always less than 0.5 only over horizons no longer than *a* quarters, less than 0.8 only over horizons no longer than *b* quarters, less than 1.0 only over horizons no longer than *c* quarters, and less than 2.0 only over horizons no longer than *d* quarters. Thus, '1235' means that the ratio is below 0.5 for one-quarter-ahead forecasts, below 0.8 for one- and two-quarters-ahead forecasts, and below 1.0 for one-, two- and three-quarters-ahead forecasts. It can be inferred that the ratio is greater than or equal to 1.0 for four-quarters-ahead forecasts and greater than or equal to 2.0 for six-quarters-ahead forecasts. The interpretation of '1235' is identical except that the underlining indicates this forecaster did not produce enough forecasts with horizons greater than five to enable McNees to reliably calculate the RMSE. Therefore, the ratio is merely not available for larger horizons and the inference that it is greater than or equal to 2.0 is not intended.

Exhibit 4 is not very easy to read; its only virtue is that it condenses 13 tables into one. Inevitably the condensation produces some distortions. With one exception, these distortions are fairly minor. Nevertheless, the reader is urged to request the full tables before drawing any detailed conclusions about the relative effectiveness in this period of any one forecaster over another. The one exception is the table on the cumulative growth rate in real GNP. The entries in this row of Exhibit 4 are discussed below.

The results in Appendix IV show that the forecasts of the growth rate in non-residential fixed investment are likely to be even less useful than those of the growth rate in GNP. Appendix V indicates that the growth rates in the CPI and the implicit GNP deflator can be usefully forecast one quarter ahead, but no further. Similarly, Appendix VI indicates that the growth in the implicit GNP deflator can be usefully forecast only one or two quarters ahead. So inclusion of these variables in a forecasting model will be helpful only if forecasts beyond these limits are not required.

Interestingly enough, forecasts of the growth rate of real GNP (Appendix VII) are more useful than those of nominal GNP. Still, looking at these figures, it hardly seems likely that inclusion of this variable in a forecasting model will be very helpful if more than one-step-ahead forecasts are required.

Appendices VIII, IX, and X give the results on the forecasts of the change in business inventories, the 90-day Treasury bill rate, and the unemployment rate. The change in business inventories appears to be usefully forecast only a couple of quarters ahead. The Treasury bill rate forecasts appear to be useful up to three to five quarters ahead, depending on the forecaster. The quality of the unemployment rate forecast is even more forecaster-dependent, with the useful forecast horizon varying from one or two quarters ahead (Chase) to seven quarters ahead (TG).

Finally, the forecasts of the five growth rate variables (GNP, CPI, GNP deflator, real GNP, and non-residential fixed investment) were re-examined in cumulative form. In Appendices XI to XV the mean square of ${}^{\mathrm{cum}}\varepsilon\%$ (defined above) is compared to the variance of the cumulative annualized compound growth rate of the actual series:

$$100[(A_{t+h}/A_t)^{4/h} - 1].$$

For GNP, non-residential fixed investment, and the CPI the cumulative growth rate forecasts appear to be only slightly better than the analogous quarter-to-quarter growth rate forecasts analysed above. In contrast, the cumulative growth rate forecasts for the GNP deflator and for real GNP are both noticeably superior to the quarter-to-quarter forecasts.

The relative superiority of the cumulative real GNP growth rate forecasts is particularly striking. Appendix XV was summarized in Exhibit 4 differently from the others because the observed ratios in this appendix show a distinct tendency to decrease as the horizon increases. Therefore, the entries for this row of Exhibit 4 give the horizon interval over which the ratio is less than 0.5. Thus, whereas BMARK and CHASE never had ratios less than 0.5, the ratio for KEDI was less than 0.5 for horizons of three quarters to seven quarters ahead, inclusive. The underlining has the same significance as before; e.g. RSQE's forecasts had a ratio less than 0.5 for horizons from two to six quarters but the ratio was not computed for seven- and eight-quarters-ahead

Variable	Forecaster															
	Early								Mid					Late		
	ASA	BEA	BMARK	CHASE	DRI	GSU	RSQE	WEFA	KEDI	MHT	TG	UCLA	WEFA	CHASE	DRI	GE
Growth rates																
GNP	0015	0006	0004	0003	0018	0118	0027	0008	0118	0015	0117	0116	0118	1118	1118	0118
Fixed investment*	—	—	—	0008	0008	0288	0026	0118	—	0025	1237	0006	0118	0018	0118	0357
CPI	—	0125	—	1114	0124	1225	—	1226	0112	1124	1113	1224	1236	1224	1225	1236
GNP deflator	0234	1236	0113	1224	1225	1235	1126	1224	0112	1124	1125	1235	1126	1224	1226	0457
Real GNP	0115	0066	0004	0018	0158	0188	0466	0088	0168	0015	1377	0116	0188	1118	1118	0135
Levels																
Δ business inventories	0125	—	—	0126	0225	0118	0346	0228	—	0225	1226	0126	0258	1457	0558	0008
90 day *T*-bill note	—	—	—	2335	3447	—	2446	3458	4568	—	—	—	3458	1345	—	—
Unemployment rate	3555	3666	—	1223	3778	—	3446	7778	2558	2234	4777	3446	7888	2223	3888	3557
Cumulative growth rates																
GNP	0055	0006	0002	0003	0038	0138	0026	0008	0117	0134	0777	0166	0488	1113	1368	0118
Fixed investment*	—	—	—	0004	0008	0678	0046	0238	—	0034	1347	0006	0348	0012	0148	0568
CPI	—	0337	—	1336	0236	3458	—	3458	0125	1345	1236	2346	3458	3457	2457	3568
GNP deflator	0555	3566	0234	2346	1347	3568	3566	1458	0168	3555	3567	3446	2458	3457	3568	0888
Real GNP	3–5	2–6	0–0	0–0	2–7	3–8	2–6	2–8	3–7	3–5	1–7	3–6	3–8	1–1	3–8	4–8

* Non-residential.

Exhibit 4. $\mathrm{MSE}(\hat{x}_t)/\mathrm{Var}(x_t)$ results (see text)

forecasts because this forecaster omitted so many forecasts that the sample size became small at large horizons.

These results indicate that many of the forecasters produced quite useful forecasts of the cumulative growth rate in real GNP over horizons from two–three to seven–eight quarters despite the fact that in most cases their predictions of the quarter-to-quarter growth rate appear to be useful only for one quarter ahead. This seems to indicate that the real GNP forecasts are inaccurate primarily with regard to the quarter-to-quarter *timing* of future changes. Thus, an over-estimate of the four-quarter-ahead growth rate will tend to be corrected in the cumulative forecast by an under-estimate of the five-quarters-ahead growth rate.

CONCLUSIONS

The results reported above demonstrate that considerable caution must be exercised in choosing explanatory variables for a forecasting model. An explanatory variable may appear to be quite significant, and be highly useful even outside the sample when actual values of it are available. Yet the errors in forecasting it may nevertheless overshadow its explanatory power.

Thus, these results seem to point toward errors in forecasting exogenous variables as an important source of endogenous variable forecasting errors. How can this be reconciled with Armstrong (1978, p. 219)'s puzzling finding that the conditional forecasts (based on actual exogenous variable values) were more accurate than the corresponding unconditional forecasts in only two out of eleven studies (with one tie)?

One possibility is that the exogenous variable forecasts used in those studies were systematically biased toward the trend or mean value. The bias would naturally arise in judgemental forecasts as part of an effort to produce 'conservative' forecasts. Armstrong (1978, p. 216) suggests a somewhat *ad hoc* procedure, which he calls 'mitigation', to systematically build in such a bias. And Ashley (1982) shows that a particular form of mitigation is optimal when information on the accuracy of the raw exogenous variable forecasts is available. So such a bias is not implausible. What would be its effect? If the estimated coefficient on a given exogenous variable is correct, one would still seem to be better off using the actual movements in the variable. But if the coefficient is not correct—due to sampling error or to econometric misspecification of some type—then a mitigated exogenous variable forecast might well be preferable to the use of the actual values.

What is surprising about the results reported above on the models for Texas personal income is that the two commercially available sources of forecasts for U.S. GNP and personal income used were of such low quality. These forecasts were so poor that bivariate models which reduced the postsample RMSE in the growth rate of Texas personal income by 28–39 per cent (compared to univariate models) when actual values of GNP and personal income were used as explanatory variables were, if anything, worse than the univariate models when the actual values were replaced by these commercially available forecasts.

Of course, the forecast origin used in that example (July, 1977) could in principle have been an unusually difficult one. However, the general theorem proved above provides a framework within which this objection can be tested. The theorem states that an exogenous variable which must be replaced by a forecast will only worsen the MSE of the endogenous variable forecasts if the MSE of the exogenous variable forecast exceeds the variance of the exogenous variable. This result is not so surprising once one notes that

$$1-(\mathrm{MSE}(\hat{x}_t)/\mathrm{Var}(x_t))$$

can be viewed as a sort of postsample R^2 value. The theorem then merely states that exogenous variable forecasts with negative postsample R^2's are of dubious value as inputs to a forecasting model.[1]

From the results of empirical testing using this general theorem, it appears that the miserable forecasting results obtained using these bivariate models for Texas personal income are typical of what we should expect from *any* forecasting model specified and estimated using GNP growth rates and forecast using available GNP projections, commercial or otherwise. Moreover, unless attention is restricted to only one- or two-quarter-ahead forecasts, similarly miserable results would ordinarily be obtained using real GNP, or the CPI, or the implicit GNP deflator, or the change in business inventories, or non-residential fixed investment as explanatory variables instead of GNP. The results indicate that variables like the 90 day treasury bill rate and the unemployment rate may be substantially more useful, depending on which forecaster is considered. But the only variable tested which most forecasters are able to usefully forecast more than a couple of quarters ahead is the cumulative growth rate in real GNP.

It is widely and implicitly believed that available forecasts of variables like the growth rate of GNP are in general sufficiently accurate that it makes sense to use these forecasts to replace actual GNP growth rates in forecasting with models that require such exogenous information. (For example, see Glickman (1977, pp. 156–158).) The results reported here clearly show that complacent acceptance of this implicit belief is no longer tenable.

The criterion developed here provides the means to easily check *ex ante* whether or not the available forecasts of a given exogenous variable are sufficiently accurate over the relevant horizon to justify its inclusion as an explanatory variable in a forecasting model. Such an *ex ante* consideration of the forecastability of explanatory variables would seem to be an essential step in the construction of a successful forecasting model.

It is shown in Ashley (1982) that direct substitution of a forecast $\hat{x}_t$ for x_t itself is in general a suboptimal use of $\hat{x}_t$ when sufficient information is available about its accuracy. In addition, a useful expression is derived there for the optimal function of $\hat{x}_t$ to substitute for x_t when sufficient information about $\hat{x}_t$'s accuracy is available. Thus, another interpretation of the results reported above is that the quality of the available forecasts of variables like GNP is simply not high enough to afford us the luxury of tolerating such suboptimization.

ACKNOWLEDGEMENTS

The author thanks J. Strange for help with the computations, C. Holt, J. Olson, W. Weber, C. W. J. Granger, and S. McNees for stimulating criticism, and S. McNees for sharing his unpublished results. Part of this work was supported by the Bureau of Business Research of the University of Texas at Austin through a grant provided by the Texas Legislative Budget Board.

REFERENCES

Armstrong, J. Scott, *Long Range Forecasting*, New York: Wiley–Interscience, 1978.

Ashley, Richard, 'Texas forecasting studies: time series analysis', Bureau of Business Research Working Paper Series, University of Texas at Austin, BP 80–3, P.O. Box 7459, Austin, Texas 78712, 1980.

Ashley, Richard, 'On the optimal use of non-optimal forecasts' (in preparation).

[1] It should be explicitly noted that MSE is not the only possible criterion for judging the accuracy of an endogenous variable forecast. In some contexts for example, one might be more interested in turning point prediction accuracy.

Ashley, Richard, Granger, C. W. J. and Schmalensee, Richard. 'Advertising and aggregate consumption: an analysis of causality', *Econometrica*, **48** (1980), 1149–1167.
Glickman, Norman J., *Econometric Analysis of Regional Systems*, New York: Academic Press, 1977.
Granger, C. W. J., 'On the properties of forecasts used in optimal economic policy decisions', *J. Public Economics*, **2** (1973), 347–356.
McNees, Stephen K., 'The recent record of thirteen forecasters', *New England Economic Review*, Sept./Oct. (1981), 5–21.

Author's biography:
Richard Ashley received his B.Sc. in biology at Caltech in 1971 and his Ph.D. in economics at the University of California, San Diego in 1976.

He has taught at the University of Texas at Austin and is now Assistant Professor of Economics at Virginia Polytechnic Institute and State University.

Author's address:
Richard Ashley, Department of Economics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, U.S.A.