INTERNATIONAL ECONOMIC REVIEW
Vol. 21, No. 2, June, 1980

ESTIMATING THE EXPECTED PREDICTIVE ACCURACY OF ECONOMETRIC MODELS*

BY RAY C. FAIR[1]

1. INTRODUCTION

A method is proposed in this paper for estimating the uncertainty of a forecast from an econometric model. The method accounts for the four main sources of uncertainty: uncertainty due to (1) the error terms, (2) the coefficient estimates, (3) the exogenous-variable forecasts, and (4) the possible misspecification of the model. It also accounts for the fact that the variances of forecast errors are not constant across time. Because the method accounts for all four sources of uncertainty, it can be used to make accuracy comparisons across models.

The method has two advantages over the common procedure of computing root mean squared errors (RMSEs) to evaluate the accuracy of econometric models. The first is that the RMSE procedure does not account for the fact that the variances of the forecast errors vary across time. Although RMSEs are in some loose sense estimates of the averages of the variances across time, no rigorous statistical interpretation can be placed on them. The second advantage is that the RMSE procedure does not take into account the uncertainty from the exogenous-variable forecasts, and so it is not possible to use RMSEs to compare models with different degrees of exogeneity.

Estimating the uncertainty from the error terms and coefficient estimates is a straightforward exercise in stochastic simulation, for which there is now a fairly large literature.[2] The uncertainty from the exogenous variables can also be estimated by means of stochastic simulation, although, as will be discussed, before doing this some assumption about the uncertainty of the exogenous variables themselves must be made. Estimating the uncertainty from the possible misspecification of the model is the most difficult and costly part of the method, and it rests on a strong "constancy" assumption. Although, as will be seen, this assumption is quite restrictive, some assumption of this kind is needed if comparisons across models are to be made. An assumption like this is, for ex-

* Manuscript received October 10, 1978; revised May 9, 1979.

[1] The research described in this paper was financed by grant SOC77-03274 from the National Science Foundation.

[2] For examples of studies in which stochastic simulation with respect to the error terms only has been performed, see Nagar [1969]; Evans, Klein and Saito [1972]; Fromm, Klein and Schink [1972]; Green, Liebenberg and Hirsch [1972]; Sowey [1973]; Cooper and Fischer [1972]; Cooper [1974]; Garbade [1975]; Bianchi, Calzolari and Corsi [1976]; and Calzolari and Corsi [1977]. For examples of studies in which stochastic simulation with respect to the error terms and coefficient estimates has been performed, see Cooper and Fischer [1974]; Schink [1971], [1974]; Haitovsky and Wallace [1972]; and Muench, Rolnick, Wallace and Weiler [1974].

ample, implicit in any comparison of RMSEs across models.

The method is described in Section 2, and the results of applying the method to two models are presented and discussed in Section 3. Section 4 contains a brief discussion of the application of the method to subjectively-adjusted models. The exact stochastic-simulation procedures that were followed for the results in Section 3 are explained in the Appendix.

With respect to the misspecification part of the method, it should be noted at the outset that this paper is not concerned with *testing* for misspecification. The present approach is rather to *estimate* the effects of misspecification on the uncertainty of forecasts. In other words, the basic premise of this paper is that misspecification is likely to exist and so must be accounted for in some way. For good examples of the hypothesis-testing approach, see Brown, Durbin, and Evans [1975] for single equation models, and Muench, Rolnick, Wallace, and Weiler [1974] for simultaneous equations models.

2. THE METHOD

2.1. *The Notation.* The method can be applied to a model that is nonlinear in both variables and coefficients. Let G denote the total number of equations in the model, M the number of stochastic equations, and N the total number of predetermined (both exogenous and lagged endogenous) variables. Assume (for expositional convenience only) that the model is quarterly, and let the i-th equation of the model for quarter τ be written:

$$\phi_i[y_1(\tau),\ldots, y_G(\tau), x_1(\tau),\ldots, x_N(\tau), \beta_i] = e_i(\tau), \qquad i = 1,\ldots, G, \tag{1}$$

where the $y_i(\tau)$ are the endogenous variables, the $x_i(\tau)$ are the predetermined variables, β_i is the vector of unknown coefficients in equation i, and $e_i(\tau)$ is the error term corresponding to equation i. For identities, $e_i(\tau)$ is zero for all τ. Also, let $e(\tau)$ denote the M-component vector of the error terms of the stochastic equations for quarter τ. For simplicity it will be assumed that $e(\tau)\sim N(0, \Omega)$ for all τ, although the following discussion can be modified to incorporate different assumptions about the distribution of $e(\tau)$.

Some of the definitions that are needed in the following discussion are listed in Table 1. It will be important to keep track of what the various expected values and variances are conditional on, and the notation in Table 1 is designed with this in mind. In addition, it should be noted that all expected values and variances in this paper are conditional on the actual values of the endogenous and predetermined variables up to the beginning of the prediction period. Also, all expected values and variances are conditional on the actual values of the exogenous variables for the prediction period unless otherwise stated.[3]

[3] With respect to the definitions in Table 1, it should also be noted that it is implicitly assumed in this paper that the variances of the forecast errors exist. For some estimation tech-

(*Continued on next page*)

TABLE 1

DEFINITIONS

β =	vector of all the unknown coefficients.
Ω =	covariance matrix of the structural error terms ($M \times M$).
t_1 =	first quarter of the estimation period.
t_2 =	last quarter of the estimation period.
$\hat{\beta}(t_1, t_2)$ =	estimate of β from the $t_1 - t_2$ sample period.
$\hat{\Omega}(t_1, t_2)$ =	estimate of Ω from the $t_1 - t_2$ sample period.
$V(t_1, t_2)$ =	covariance matrix of $\hat{\beta}(t_1, t_2)$.
$\hat{V}(t_1, t_2)$ =	estimate of $V(t_1, t_2)$.
t =	any quarter of the prediction period.
$\tilde{y}_i(t, k)$ =	expected value of the k-quarter-ahead forecast of variable i for quarter t conditional on Ω and β.
$\sigma_i^2(t, k)$ =	variance of the forecast error for the k-quarter-ahead forecast of variable i for quarter t conditional on Ω and β.
$\tilde{y}_i(t, k, t_1, t_2)$ =	expected value of the k-quarter-ahead forecast of variable i for quarter t conditional on Ω, $\hat{\beta}(t_1, t_2)$, and $V(t_1, t_2)$.
$\sigma_i^2(t, k, t_1, t_2)$ =	variance of the forecast error for the k-quarter-ahead forecast of variable i for quarter t conditional on Ω, $\hat{\beta}(t_1, t_2)$, and $V(t_1, t_2)$.
$\tilde{\tilde{y}}_i(t, k)$ =	stochastic-simulation estimate of $\tilde{y}_i(t, k)$.
$\tilde{\sigma}_i^2(t, k)$ =	stochastic-simulation estimate of $\sigma_i^2(t, k)$.
$\tilde{\tilde{y}}_i(t, k, t_1, t_2)$ =	stochastic-simulation estimate of $\tilde{y}_i(t, k, t_1, t_2)$.
$\tilde{\sigma}_i^2(t, k, t_1, t_2)$ =	stochastic-simulation estimate of $\sigma_i^2(t, k, t_1, t_2)$.
$\tilde{y}_i(t, k, t_1, t_2, x)$ =	expected value of the k-quarter-ahead forecast of variable i for quarter t conditional on Ω, $\hat{\beta}(t_1, t_2)$, $V(t_1, t_2)$, and some assumption about the uncertainty of the exogenous-variable forecasts.
$\sigma_i^2(t, k, t_1, t_2, x)$ =	variance of the forecast error for the k-quarter-ahead forecast of variable i for quarter t conditional on Ω, $\hat{\beta}(t_1, t_2)$, $V(t_1, t_2)$, and some assumption about the uncertainty of the exogenous-variable forecasts.
$\tilde{\tilde{y}}_i(t, k, t_1, t_2, x)$ =	stochastic-simulation estimate of $\tilde{y}_i(t, k, t_1, t_2, x)$.
$\tilde{\sigma}_i^2(t, k, t_1, t_2, x)$ =	stochastic-simulation estimate of $\sigma_i^2(t, k, t_1, t_2, x)$.
$\tilde{\sigma}_i^2(t, k, t_1, t_2, x, d)$ =	estimate of the total variance of the forecast error for the k-quarter-ahead forecast of variable i for quarter t.

Although the method relies heavily on the use of stochastic simulation, it can be explained without going into the details of the simulation procedures. Because of this and because these details are in part model specific, no mention of particular simulation procedures is made in this section. As noted in the Introduction, the stochastic-simulation procedures that were followed for the results in Section 3

(Continued)

niques this is not always the case. If in a given application the variances do not exist, then one should estimate other measures of dispersion of the distribution, such as the interquartile range or mean absolute deviation.

are explained in the Appendix.

2.2. *Estimating* $\sigma_i^2(t, k)$. For a one-quarter-ahead forecast ($k=1$), $\sigma_i^2(t, k)$ is merely the variance of the reduced form error term for variable i and quarter t. For a linear model this variance is not a function of t (assuming that Ω is not a function of t), and an analytic expression for it can be obtained. For a nonlinear model, neither of these is in general true. It is, however, fairly straightforward to estimate $\sigma_i^2(t, k)$ by means of stochastic simulation. If β and Ω were known, one would merely draw for each trial a set of error terms from the $N(0, \Omega)$ distribution and solve the model for this set using the known value of β. In this case the stochastic-simulation estimate of $\sigma_i^2(t, k)$ would differ from the true value only because of sampling error (i.e., because of a finite number of draws). In practice, of course, only estimates of β and Ω are available, and these must be used for the stochastic simulation. This is another reason the stochastic-simulation estimate of $\sigma^2(t, k)$ will differ from the true value.

2.3. *Estimating* $\sigma_i^2(t, k, t_1, t_2)$. For $k=1$, $\sigma_i^2(t, k, t_1, t_2)$ is what is usually referred to as the variance of the forecast error. This variance can also be estimated in a straightforward way by stochastic simulation, using $\hat{\beta}(t_1, t_2)$, $\hat{\Omega}(t_1, t_2)$, and $\hat{V}(t_1, t_2)$. Note in this case that in practice $\hat{V}(t_1, t_2)$ is usually only an estimate of the asymptotic approximation of $V(t_1, t_2)$, not an estimate of $V(t_1, t_2)$ directly.[4] This is another source of simulation error. Each simulation trial for estimating this variance consists of draws of both error terms and coefficients.

2.4. *Estimating* $\sigma_i^2(t, k, t_1, t_2, x)$. Estimating this variance by stochastic simulation is less straightforward than estimating the previous two. There is no obvious estimate available of the degree of uncertainty of the exogenous-variable forecasts themselves, and so some assumption about this must first be made. There are two polar assumptions that can be made in this regard. One is, of course, that there is no exogenous-variable uncertainty. The other is that the exogenous-variable forecasts are in some way as uncertain as the endogenous-variable forecasts. With respect to this latter assumption, one could, for example, estimate an autoregressive equation for each exogenous variable and then add these equations to the model. This expanded model, which would have no exogenous variables, could then be used for the stochastic-simulation estimates of $\sigma_i^2(t, k)$ and $\sigma_i^2(t, k, t_1, t_2)$. This procedure is likely, however, to exaggerate the uncertainty of many exogenous variables. This is particularly true of fiscal-policy variables, where government-budget data are usually quite useful for purposes of forecasting up to at least about eight quarters ahead. The assumption of no uncertainty is also clearly unrealistic, and so the truth seems likely to lie somewhere between the two polar assumptions.

[4] Note that the use of estimated asymptotic distributions for the stochastic simulations may mask the problem of the possible nonexistence of variances mentioned in footnote 3. The variances may exist for the estimated asymptotic distributions but not for the true finite sample distributions.

The assumption that was made in this study for the results in Section 3 is probably closer to the second polar assumption than it is to the first. The procedure followed in this case, which is explained in detail in the Appendix, was to estimate an eighth-order autoregressive equation for each exogenous variable (including a constant and time in the equation) and then to take the estimated standard error from this regression as the estimate of the degree of uncertainty attached to forecasting the change in this variable for each quarter. This procedure ignores the uncertainty of the coefficient estimates in the autoregressive equations, which is one reason it is not as extreme as the procedure that would be followed under the second polar assumption.[5]

Each simulation trial for estimating this variance consists of draws of error terms, coefficients, and exogenous-variable errors.

2.5. *Estimating the Uncertainty from Misspecification — Computing* $\tilde{\sigma}_i^2(t, k, t_1, t_2, x, d)$. As noted in the Introduction, this part of the method is costly and rests on a strong assumption. It is based on a comparison of estimated variances computed by means of stochastic simulation with estimated variances computed from outside-sample forecast errors. For a correctly specified model the expected value of the difference between these two estimates for any given variable and length of forecast is, ignoring simulation error, zero. Misspecification has two effects on this difference. First, if the model is misspecified, the estimated covariance matrices that are used for the stochastic simulation will not in general be unbiased estimates of the true covariance matrices. The estimated variances computed by means of stochastic simulation will thus in general be biased. Second, the estimated variances computed from the forecast errors will also in general be biased estimates of the true variances. Since misspecification affects both estimates, the expected value of the difference between these estimates may be negative, positive, or even zero for a misspecified model.

The assumption upon which this part of the method is based is that the model is misspecified in such a way that for each variable and length of forecast, the expected value of the difference between the two estimates of the variance is constant across time. As will be seen, given this assumption, it is possible to estimate the total variance of the forecast error for each variable and length of forecast. This part of the method effectively accounts for misspecification effects (if any) that are not already reflected in the variances that are estimated by means of stochastic simulation. It requires successive reestimation and stochastic simulation of the model.

[5] In the stochastic-simulation study of Haitovsky and Wallace [1972], third-order autoregressive equations were estimated for the exogenous variables, and these equations were then added to the model. This procedure is thus consistent with the second polar assumptions above *except* that for purposes of the stochastic simulations Haitovsky and Wallace took the variances of the error terms to be one-half of the estimated variances. They defend this procedure (pp. 267–268) on the grounds that the uncertainty from the exogenous-variable forecasts is likely to be less than is reflected in the autoregressive equations, a view that is consistent with the above discussion.

It is easiest to describe this part of the method by means of an example. Consider first the case for $k=1$ (a one-quarter-ahead forecast), and assume that data are available from quarters 1 through 100. Assume also that the lags in the model are such that the estimation period can begin with quarter 11. All stochastic simulations described in this subsection are based on actual values of the exogenous variables (no exogenous-variable uncertainty).

Consider now the case in which the model is estimated for the 11–70 period. Given this set of estimates (i.e., given $\hat{\beta}(11, 70)$, $\hat{\Omega}(11, 70)$, and $\hat{V}(11, 70)$), one can estimate by stochastic simulation the variance of the forecast error for each variable i for quarter 71 (i.e., one can compute $\tilde{\sigma}_{i}^{2}(71, 1, 11, 70)$). In the process of doing this, one also obtains an estimate of the expected value of the forecast for each variable i for quarter 71, $\tilde{\bar{y}}_i(71, 1, 11, 70)$. The difference between this value and the actual value, $y_i(71)$, is the mean forecast error for quarter 71:

$$(2) \qquad \hat{\varepsilon}_i(71, 1, 11, 70) = y_i(71) - \tilde{\bar{y}}_i(71, 1, 11, 70).$$

If it is assumed that the stochastic-simulation estimate of $\bar{y}_i(71, 1, 11, 70)$ exactly equals the true expected value, then $\hat{\varepsilon}_i(71, 1, 11, 70)$ is a sample draw from a distribution with a known mean of zero and variance $\sigma_i^2(71, 1, 11, 70)$. $\hat{\varepsilon}_i^2(71, 1, 11, 70)$ is thus under this assumption an unbiased estimate of $\sigma_i^2(71, 1, 11, 70)$. One thus has two estimates of this variance, one computed from the mean forecast error and one computed by stochastic simulation. Let $d_i(71, 1, 11, 70)$ denote the difference between these two estimates:

$$(3) \qquad d_i(71, 1, 11, 70) = \hat{\varepsilon}_i^2(71, 1, 11, 70) - \tilde{\sigma}_{i}^{2}(71, 1, 11, 70).$$

If it is further assumed that the stochastic-simulation estimate of $\sigma_i^2(71, 1, 11, 70)$ exactly equals the true value, then $d_i(71, 1, 11, 70)$ is the difference between the estimated variance based on the mean forecast error and the true variance. Therefore, under the above two assumptions of no error in the stochastic-simulation estimates, the expected value of $d_i(71, 1, 11, 70)$ is zero.

Given that data are available through quarter 100, the above procedure can be repeated for quarters 72 through 100. The model can, for example, be reestimated through quarter 71 and the above calculations performed for quarter 72. This will yield a value of $d_i(72, 1, 11, 71)$ for each variable i. Similarly, a value of $d_i(73, 1, 11, 72)$ can be computed by reestimating the model through quarter 72 and performing the above calculations for quarter 73, and so on through quarter 100. This procedure will yield 30 values of $d_i(t, 1, 11, t-1)$ $(t=71, 72, \ldots, 100)$ for each variable i, each of the 30 values being based on a different set of coefficient estimates of the model and a different stochastic simulation. If the above two assumptions of no simulation error hold for all t, then the expected value of $d_i(t, 1, 11, t-1)$ is zero for all t.

The discussion of this example has so far been based on the assumption that the model is correctly specified. As noted at the beginning of this subsection, misspecification will in general affect both estimates of the variance, and so the sign of the effect of misspecification on the difference between the two estimates is ambiguous. It is clearly possible for misspecification to affect the two estimates in the same way and thus leave the expected value of the difference between them equal to zero. In general, however, this does not seem likely.[6] In other words, one would not generally expect the mean of the distribution of $d_i(t, 1, 11, t-1)$ to be zero for a misspecified model. There is also no particular reason to expect the mean of this distribution to be constant across time, but the method of this paper does rest on a constancy assumption of this kind. In particular, the following assumption is made. For variables that have no trend, it is assumed that the mean of $d_i(t, 1, 11, t-1)$ is constant across time (i.e., is not a function of t). For variables that have a trend, it is assumed that the mean of $d_i(t, 1, 11, t-1)/\tilde{\bar{y}}_i^2(t, 1, 11, t-1)$ is constant across time, i.e., that the mean of $d_i(t, 1, 11, t-1)$ is proportional to the square of the estimated mean of the variable (remember that $d_i(t, 1, 11, t-1)$ is in units of the variable squared).

[6] The following example may help in understanding the effect of misspecification on the two estimates. Assume that the model is a single equation and that the true equation is $y_\tau = \beta_1 x_{1\tau} + \beta_2 x_{2\tau} + e_\tau$, where e_τ; obeys all the assumptions of the classical regression model. Assume also that $x_{2\tau}$ is (incorrectly) excluded from the estimated equation, and let $\hat{\beta}_1$ be the least squares estimate of β_1 from the regression of y_τ on $x_{1\tau}$ (for say, $\tau = 1, 2, \ldots, T$). $\hat{\beta}_1$ is a biased estimate of β_1, with bias $p\beta_2$, where p is the coefficient estimate in the regression of $x_{2\tau}$ on $x_{1\tau}$. Consider now the forecast of y_t, $\hat{\beta}_1 x_{1t}$, where t is, say, $T+1$. The forecast error, $\hat{\varepsilon}_t$, is $\hat{\beta}_1 x_{1t} - (\beta_1 x_{1t} + \beta_2 x_{2t} + e_t)$, and the expected value of $\hat{\varepsilon}_t^2$ can be easily seen to be:

(i) $$\sigma^2\left(1 + \frac{x_{1t}^2}{\sum x_{1\tau}^2}\right) + \beta_2^2(x_{2t} - px_{1t})^2,$$

where σ^2 is the variance of e_τ and $\sum$ denotes summation from 1 to T. If $\beta_2 = 0$, then the estimated equation is not misspecified, and the expected value of $\hat{\varepsilon}_t^2$ is merely the first term in (i), a well known result. Now, the estimated variance of the regression of y_τ on $x_{1\tau}$ is the sum of squared residuals divided by $T-1$, and this is the variance of the error term that would be used in the stochastic simulation. Its expectation is

(ii) $$\sigma^2 + \beta_2^2 \frac{SSR_{21}}{T-1},$$

where SSR_{21} is the sum of squared residuals in the regression of $x_{2\tau}$ on $x_{1\tau}$. Ignoring simulation error, the variance of the forecast error for period t that would be computed from stochastic simulation is

(iii) $$\left(\sigma^2 + \beta_2^2 \frac{SSR_{21}}{T-1}\right)\left(1 + \frac{x_{1t}^2}{\sum x_{1\tau}^2}\right).$$

The difference between (i) and (iii) is the expected value of the difference between the two estimates of the variance:

(iv) $$\beta_2^2\left[(x_{2t} - px_{1t})^2 - \frac{SSR_{21}}{T-1}\left(1 + \frac{x_{1t}^2}{\sum x_{1\tau}^2}\right)\right].$$

Unless $\beta_2 = 0$, this difference is not in general zero, although it can obviously be either positive or negative. This example thus shows clearly that misspecification affects both estimates of the variance and so has an ambiguous effect on the sign of the difference between the two. This example also shows that the assumption that the expected value of the difference is constant across time is quite strong and is at best likely to be only a rough approximation to the truth. I am indebted to an anonymous referee for this example.

For variables that have no trend, let $\bar{d}_i(1)$ denote the sample mean of the $d_i(t, 1, 11, t-1)$ values, and for variables that have a trend, let $\bar{d}_i(1)$ denote the sample mean of the $d_i(t, 1, 11, t-1)/\tilde{\bar{y}}_i^2(t, 1, 11, t-1)$ values. In the above example, $\bar{d}_i(1)$ would be based on a sample of size 30.

It should be stressed that all the stochastic simulations and outside-sample forecast errors that are involved in computing $\bar{d}_i(1)$ are based on actual values of the exogenous variables. Unless actual exogenous-variable values are used, the expected value of $d_i(\cdots)$ is not in general zero even under the assumptions of no simulation error and correct specification. Once $\bar{d}_i(1)$ has been obtained, however, one can then account for exogenous-variable uncertainty. From Section 2.4 one has a stochastic-simulation estimate of $\sigma_i^2(t, 1, t_1, t_2, x)$, and given this and $\bar{d}_i(1)$, one can compute the total variance of the forecast error. In other words, one can compute $\tilde{\sigma}_i^2(t, 1, t_1, t_2, x, d)$ in Table 1. For variables that have no trend, this is merely:

$$\tilde{\sigma}_i^2(t, 1, t_1, t_2, x, d) = \tilde{\sigma}_i^2(t, 1, t_1, t_2, x) + \bar{d}_i(1). \tag{4}$$

For variables that have a trend, the correct formula is:

$$\tilde{\sigma}_i^2(t, 1, t_1, t_2, x, d) = \tilde{\sigma}_i^2(t, 1, t_1, t_2, x) + \bar{d}_i(1)\cdot\tilde{\bar{y}}^2(t, 1, t_1, t_2, x). \tag{5}$$

The generalization of the above procedure to k-quarter-ahead forecasts is straightforward. Just substitute k for 1 in the discussion. Each length of forecast will have its own $\bar{d}_i(k)$ value, and these values will in general be different for different lengths. Also, note that one observation is lost for each one quarter increase in the length of the forecast. In the above example, given the beginning quarter of 71 and the ending quarter of 100, 29 values for $d_i(t, 2, 11, t-2)$ could be computed, 28 values of $d_i(t, 3, 11, t-3)$ could be computed, and so on.

There are at least two options available in computing $\bar{d}_i(k)$, and these should be mentioned. First, in computing the individual $d_i(\cdots)$ values, one can vary the beginning quarter of the estimation period (t_1). In the above discussion t_1 was always taken to be 11. Second, the distance between the last quarter of the estimation period and the first quarter of the prediction period need not be one quarter, as assumed above. In fact, as will be discussed, for the results in Section 3 the distance was taken to be two quarters. The criterion that one should use in choosing the first option is to choose the one that seems likely to correspond to the constancy assumption about the mean of $d_i(\cdots)$ being the best approximation.

The assumption that the mean of $d_i(\cdots)$ is constant across time is clearly the strongest of the above assumptions, and it is an open question how good an approximation it is likely to be. The other two assumptions, of no simulation error, are not nearly as important, since they are really only needed to prove that the expected value of $d_i(\cdots)$ is zero under the null hypothesis of no misspecification. Given that some kind of constancy assumption has to be made for comparisons across models, the present assumption seems the most obvious one to make, and this is the main defense for making it. As noted in the Introduction, an assumption of this kind is also implicit in comparisons of RMSEs across models.

Another constancy assumption that could be made is that the mean of $d_t(\cdots)$ follows a linear time trend. This trend could be estimated by regressing the 30 or so values of $d_t(\cdots)$ for each variable and length of forecast on a constant and time. The above formulas can be easily modified to incorporate this assumption. It is also possible, of course, to plot the individual $d_t(\cdots)$ values over time and look for systematic patterns. This information could then be used to help formulate a constancy assumption. In short, while the method proposed in this paper does require a constancy assumption, there are a number of choices, and in future work it will be of interest to examine the sensitivity of results like those in Section 3 to alternative choices.

Since for a correctly specified model the mean of $d_t(\cdots)$ is zero, examining the individual $d_t(\cdots)$ values may also reveal information about the strengths and weaknesses of the model that is useful in future work on the model. In other words, the individual $d_t(\cdots)$ values may be of interest in their own right aside from their use in comparisons across models.

One other point about the constancy assumption should be noted, which concerns the question of data mining. If in the traditional sense one has mined the data within some sample period, then one would expect that variances estimated from outside-sample forecast errors would on average be larger than variances estimated by means of stochastic simulation. Aside from possible reservations about the constancy assumption, the present method does penalize a model for this kind of data mining. There is, however, a subtler form of data mining that the method does not account for. If, say, a model were specified in quarter 100, estimated through quarter 90, and tested with respect to its outside-sample forecasting accuracy for the period 91–100, then it is clear that this is not a strict outside-sample test. Information on what happened between quarters 91 and 100 may have been used in the specification of the model, and so one cannot be sure that the model's "outside-sample" accuracy estimated for quarters 91–100 will hold for, say, quarters 101–110. In the present context this means that the expected value of the difference between the two estimates of the variance may be larger for the period 101–110 than it is for the period 91–100, which, of course, violates the constancy assumption. The present method thus does not take into account this subtler form of data mining.

3. THE APPLICATION OF THE METHOD TO TWO MODELS

For purposes of this study the method was applied to the model in Fair [1976, 1978a], which will be called Model I, and to a "naive" model, which will be called Model II. Model I consists of 97 equations, 29 of which are stochastic, and has 183 unknown coefficients to estimate (including 13 serial correlation coefficients). There are 60 exogenous variables (not counting the constant, time, and various dummy variables). The model is nonlinear in variables and coefficients, the latter because of the serial correlation coefficients, which are treated for present purposes

TABLE 2
ESTIMATED STANDARD ERRORS OF FORECASTS

	1978			1979				1980				1981			
t:	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV
k:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Model I. Real GNP															
a	0.65	0.88	1.03	1.15	1.25	1.30	1.35	1.34	1.36	1.40	1.43	1.44	1.47	1.46	1.43
b	0.67	0.95	1.19	1.38	1.49	1.59	1.66	1.69	1.77	1.81	1.82	1.84	1.88	1.88	1.94
c	0.74	1.09	1.37	1.63	1.76	1.94	2.04	2.08	2.15	2.18	2.22	2.30	2.34	2.36	2.43
d	0.80	1.23	1.54	1.96	2.27	2.51	2.48	2.27							
Model II. Real GNP															
a	0.61	1.02	1.34	1.64	1.84	1.94	2.01	2.03	2.04	2.03	2.04	2.04	2.03	2.03	2.03
b, c	0.67	1.13	1.53	1.90	2.20	2.38	2.50	2.59	2.64	2.68	2.73	2.77	2.81	2.84	2.87
d	1.09	1.93	2.72	3.45	4.01	4.32	4.58	4.74							
Model I. GNP Deflator															
a	0.28	0.35	0.42	0.47	0.51	0.55	0.59	0.61	0.64	0.65	0.65	0.65	0.66	0.67	0.68
b	0.31	0.47	0.58	0.71	0.83	0.93	1.02	1.10	1.19	1.28	1.37	1.44	1.50	1.57	1.63
c	0.44	0.67	0.84	1.04	1.21	1.36	1.49	1.62	1.75	1.88	1.98	2.09	2.23	2.35	2.43
d	0.53	0.93	1.37	1.87	2.33	2.74	3.15	3.48							
Model II. GNP Deflator															
a	0.20	0.36	0.53	0.71	0.90	1.08	1.24	1.37	1.49	1.58	1.65	1.71	1.76	1.80	1.83
b, c	0.24	0.45	0.70	1.00	1.36	1.73	2.10	2.48	2.84	3.18	3.52	3.85	4.17	4.48	4.80
d	0.45	0.94	1.53	2.25	3.12	4.05	5.10	6.20							
Model I. Unemployment Rate (units of percentage points)															
a	0.27	0.45	0.57	0.64	0.71	0.77	0.80	0.82	0.82	0.85	0.90	0.92	0.93	0.96	0.97
b	0.36	0.58	0.76	0.92	1.03	1.12	1.16	1.23	1.28	1.34	1.38	1.42	1.50	1.56	1.62
c	0.36	0.60	0.80	0.95	1.08	1.17	1.24	1.31	1.35	1.41	1.47	1.50	1.55	1.59	1.64
d	0.35	0.60	0.77	0.82	0.85	0.83	0.77	0.71							
Model II. Unemployment Rate (units of percentage points)															
a	0.28	0.55	0.77	0.94	1.02	1.08	1.12	1.14	1.15	1.16	1.16	1.16	1.15	1.15	1.16
b, c	0.29	0.58	0.84	1.04	1.17	1.27	1.34	1.40	1.44	1.48	1.52	1.55	1.59	1.63	1.66
d	0.36	0.74	1.12	1.48	1.73	1.91	2.07	2.19							
Model I. Wage Rate															
a	0.60	0.77	0.88	0.89	0.96	1.01	1.03	1.05	1.07	1.10	1.08	1.07	1.08	1.04	1.05
b	0.70	0.93	1.12	1.34	1.52	1.65	1.76	1.82	1.94	2.04	2.15	2.27	2.35	2.45	2.51
c	0.67	0.95	1.16	1.35	1.53	1.66	1.80	1.94	2.08	2.20	2.32	2.40	2.52	2.61	2.69
d	0.65	1.06	1.45	2.01	2.53	3.07	3.59	4.16							
Model II. Wage Rate															
a	0.30	0.40	0.48	0.53	0.59	0.61	0.67	0.72	0.76	0.81	0.85	0.88	0.91	0.95	0.98
b, c	0.36	0.48	0.59	0.75	0.86	0.97	1.15	1.29	1.46	1.64	1.81	1.99	2.19	2.39	2.59
d	0.63	0.84	1.04	1.26	1.41	1.56	1.81	2.04							

TABLE 2 (CONTINUED)

	1978			1979				1980				1981			
t:	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV
k:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Model I. Bill Rate (units of percentage points)															
a	0.45	0.67	0.78	0.84	0.91	0.93	0.97	0.98	0.97	0.98	0.98	0.97	0.97	1.01	1.03
b	0.48	0.71	0.86	1.01	1.08	1.14	1.21	1.25	1.28	1.32	1.32	1.35	1.37	1.37	1.42
c	0.49	0.72	0.92	1.06	1.16	1.25	1.31	1.37	1.44	1.51	1.53	1.54	1.56	1.58	1.61
d	0.61	0.96	1.08	1.17	1.31	1.47	1.56	1.72							
Model II. Bill Rate (units of percentage points)															
a	0.46	0.72	0.80	0.85	0.90	0.93	0.94	0.94	0.94	0.94	0.96	0.99	1.02	1.03	1.04
b, c	0.47	0.77	0.93	1.05	1.14	1.20	1.22	1.19	1.16	1.14	1.11	1.14	1.15	1.16	1.18
d	0.69	1.12	1.33	1.51	1.63	1.71	1.77	1.83							
Model I. Money Supply															
a	0.83	1.09	1.29	1.47	1.62	1.76	1.84	1.92	1.98	2.03	2.11	2.13	2.17	2.19	2.23
b	0.91	1.31	1.63	1.87	2.13	2.36	2.56	2.79	2.96	3.16	3.35	3.55	3.72	3.94	4.15
c	0.91	1.33	1.69	1.98	2.34	2.68	3.06	3.45	3.79	4.14	4.51	4.88	5.28	5.63	5.97
d	1.39	2.16	2.95	3.75	4.62	5.50	6.49	7.50							
Model II. Money Supply															
a	0.54	0.67	0.81	0.88	0.97	1.02	1.10	1.15	1.22	1.28	1.33	1.40	1.45	1.49	1.51
b, c	0.62	0.81	1.03	1.21	1.36	1.53	1.72	1.93	2.12	2.34	2.51	2.73	2.96	3.19	3.42
d	1.38	1.64	1.95	2.32	2.62	2.91	3.30	3.70							

a = uncertainty due to error terms — $\tilde{\sigma}_{it}(t, k)$.
b = uncertainty due to error terms and coefficient estimates — $\tilde{\sigma}_{it}(t, k, t_1, t_2)$.
c = uncertainty due to error terms, coefficient estimates, and exogenous-variable forecasts — $\tilde{\sigma}_{it}(t, k, t_1, t_2, x)$.
d = uncertainty due to error terms, coefficient estimates, exogenous-variable forecasts, and possible misspecification of the model — $\tilde{\sigma}_{it}(t, k, t_1, t_2, x, d)$.
Basic Estimation Period: t_1=1954I, t_2=1977IV.
Forecast Period: t=1978II,..., 1981IV
Model I = model in Fair [1978a].
Model II = naive model. For Model II there are no exogenous variables, so $c=b$ for this model.
For the unemployment rate and the bill rate, the errors are in the natural units of the variables. For the other variables, the errors are expressed as percentages of the forecast means (in percentage points).

as structural coefficients. Model II is a system of completely separate equations. Each variable in the model is simply regressed on a constant, time, and its first eight lagged values.

In order to apply the method one must first choose a forecast period. For present purposes this was taken to be 1978II–1981IV (15 quarters). Model I was estimated for the 1954I–1977IV period (96 observations) and then used to forecast

TABLE 3

ESTIMATED FORECAST MEANS

	1978			1979				1980				1981			
t:	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV
k:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Model I. Real GNP (billions of 1972 *dollars)*															
0	1374.8	1390.5	1404.8	1419.2	1432.2	1445.1	1458.9	1472.4	1485.4	1498.5	1511.7	1524.7	1537.4	1549.9	1562.8
a	1374.5	1390.1	1403.9	1418.4	1431.1	1444.2	1458.0	1472.1	1485.2	1497.4	1510.2	1523.4	1535.8	1548.0	1560.9
b	1375.1	1390.3	1404.1	1418.2	1430.8	1443.2	1457.5	1471.2	1483.6	1496.6	1510.0	1522.9	1535.6	1548.1	1561.3
c	1374.3	1389.4	1402.6	1416.4	1428.8	1441.1	1454.9	1468.6	1481.8	1495.8	1508.8	1521.2	1534.2	1546.6	1559.3
Model II. Real GNP (billions of 1972 *dollars)*															
0	1356.7	1358.6	1362.1	1368.0	1378.9	1389.9	1399.2	1408.6	1417.5	1425.0	1432.5	1440.8	1449.2	1457.5	1466.3
a	1356.7	1358.3	1361.7	1367.7	1378.8	1390.0	1399.5	1409.1	1418.0	1425.5	1432.8	1440.8	1449.1	1457.4	1466.4
b	1356.6	1357.9	1361.3	1367.3	1378.3	1389.7	1399.7	1409.5	1418.7	1426.4	1434.2	1442.5	1451.1	1459.1	1468.5
Model I. GNP Deflator (1972=1.0)															
0	1.4859	1.5044	1.5259	1.5452	1.5649	1.5848	1.6077	1.6284	1.6494	1.6708	1.6951	1.7172	1.7397	1.7625	1.7884
a	1.4858	1.5043	1.5261	1.5454	1.5650	1.5848	1.6078	1.6286	1.6495	1.6713	1.6958	1.7183	1.7412	1.7639	1.7903
b	1.4859	1.5045	1.5262	1.5456	1.5653	1.5853	1.6085	1.6293	1.6504	1.6719	1.6967	1.7190	1.7424	1.7657	1.7922
c	1.4861	1.5051	1.5268	1.5465	1.5661	1.5856	1.6088	1.6297	1.6507	1.6723	1.6972	1.7195	1.7423	1.7655	1.7919
Model II. GNP Deflator (1972=1.0)															
0	1.4922	1.5183	1.5466	1.5759	1.6060	1.6369	1.6683	1.7001	1.7323	1.7647	1.7976	1.8309	1.8648	1.8993	1.9345
a	1.4922	1.5182	1.5465	1.5755	1.6056	1.6364	1.6677	1.6995	1.7318	1.7644	1.7974	1.8310	1.8650	1.8998	1.9353
b	1.4922	1.5185	1.5469	1.5762	1.6065	1.6376	1.6693	1.7014	1.7337	1.7663	1.7993	1.8326	1.8663	1.9008	1.9360

Model I. Unemployment Rate (percentage points)

0	6.38	6.41	6.43	6.46	6.48	6.51	6.51	6.50	6.48	6.46	6.43	6.39	6.36	6.33	6.30
a	6.40	6.43	6.46	6.48	6.50	6.51	6.50	6.48	6.44	6.40	6.38	6.34	6.31	6.29	6.27
b	6.39	6.42	6.45	6.49	6.49	6.51	6.51	6.50	6.49	6.46	6.44	6.43	6.41	6.39	6.36
c	6.41	6.45	6.48	6.53	6.54	6.55	6.56	6.54	6.53	6.48	6.44	6.41	6.38	6.34	6.32

Model II. Unemployment Rate (percentage points)

0	5.98	5.90	5.98	6.18	6.35	6.45	6.53	6.59	6.64	6.69	6.73	6.74	6.76	6.78	6.80
a	5.96	5.87	5.95	6.15	6.32	6.42	6.50	6.55	6.60	6.66	6.70	6.73	6.76	6.79	6.82
b	5.98	5.91	6.00	6.19	6.35	6.45	6.53	6.96	6.66	6.71	6.75	6.77	6.80	6.82	6.84

Model I. Wage Rate (current dollars per hour)

0	6.47	6.59	6.71	6.83	6.95	7.08	7.21	7.34	7.48	7.62	7.76	7.90	8.05	8.20	8.35
a	6.47	6.59	6.70	6.83	6.95	7.08	7.21	7.34	7.48	7.62	7.76	7.91	8.06	8.21	8.36
b	6.47	6.59	6.71	6.83	6.95	7.08	7.21	7.35	7.49	7.63	7.77	7.91	8.06	8.22	8.37
c	6.47	6.59	6.71	6.83	6.96	7.09	7.22	7.36	7.49	7.64	7.78	7.93	8.08	8.23	8.38

Model II. Wage Rate (current dollars per hour)

0	6.49	6.64	6.76	6.90	7.05	7.20	7.35	7.51	7.67	7.83	8.00	8.17	8.35	8.53	8.72
a	6.49	6.64	6.77	6.90	7.05	7.20	7.35	7.51	7.67	7.83	8.00	8.17	8.35	8.53	8.72
b	6.49	6.64	6.77	6.90	7.05	7.19	7.35	7.51	7.67	7.83	8.00	8.17	8.35	8.53	8.72

Model I. Bill Rate (percentage points)

0	6.68	6.85	6.95	7.04	7.10	7.15	7.21	7.28	7.35	7.41	7.48	7.55	7.62	7.68	7.74
a	6.68	6.85	6.97	7.05	7.09	7.16	7.21	7.28	7.35	7.41	7.47	7.55	7.62	7.67	7.71
b	6.69	6.85	6.93	7.02	7.08	7.14	7.24	7.32	7.35	7.40	7.49	7.59	7.67	7.73	7.81
c	6.67	6.83	6.93	7.01	7.10	7.16	7.22	7.28	7.33	7.39	7.47	7.56	7.63	7.72	7.78

Model II. Bill Rate (percentage points)

0	6.87	7.43	7.81	7.95	8.00	8.02	7.92	7.72	7.51	7.34	7.20	7.09	7.06	7.10	7.19
a	6.87	7.42	7.82	7.96	8.01	8.05	7.95	7.74	7.54	7.36	7.21	7.11	7.06	7.10	7.20
b	6.88	7.43	7.79	7.92	7.98	7.99	7.88	7.68	7.48	7.29	7.14	7.02	6.98	7.02	7.12

TABLE 3 (CONTINUED)

	1978			1979				1980				1981			
	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV
k:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Model I. Money Supply (billions of current dollars)															
0	373.1	380.8	388.6	396.4	404.3	412.3	420.5	428.8	437.1	445.6	454.2	463.0	471.9	480.8	490.0
a	373.2	380.9	388.7	396.5	404.5	412.4	420.6	429.0	437.5	446.0	454.6	463.4	472.2	481.8	490.4
b	373.2	381.0	388.8	396.9	404.9	413.0	421.2	429.5	437.9	446.7	455.6	464.3	473.3	482.2	491.5
c	373.1	380.9	388.9	396.9	404.8	412.9	421.0	429.4	437.8	446.5	455.3	464.1	473.2	482.2	491.7
Model II. Money Supply (billions of current dollars)															
0	371.6	377.2	383.0	388.5	394.7	401.2	407.3	413.6	420.0	426.4	432.9	439.6	446.3	453.2	460.1
a	371.6	377.1	383.1	388.5	394.7	401.1	407.2	413.5	419.9	426.3	432.7	439.5	446.2	453.0	459.9
b	371.7	377.3	383.3	388.7	395.0	401.5	407.8	414.2	420.9	427.3	434.2	440.9	448.0	454.8	462.0

0 = error terms set equal to zero (no stochastic simulation).
a = stochastic simulation with respect to error terms only — $\bar{\tilde{y}}_i(t, k)$.
b = stochastic simulation with respect to error terms and coefficient estimates — $\bar{\tilde{y}}_i(t, k, t_1, t_2)$.
c = stochastic simulation with resepct to error terms, coefficient estimates, and exogenous variables — $\bar{\tilde{y}}_i(t, k, t_1, t_2, x)$.
See notes to Table 2.

1978II–1981IV; and Model II was estimated for the 1954II–1977IV period (95 observations) and then used to forecast 1978II–1981IV.[7] These were actual *ex ante* forecasts, and so guessed values of the exogenous variables were used for Model I. (Model II has no non-trivial exogenous variables.) Data for 1978I were used as initial conditions for the forecast, but because they were only preliminary at the time (April 1978), they were not used for the estimation.

Given the coefficient estimates and forecast period, one can compute by stochastic simulation $\tilde{\sigma}_i^2(t, k)$, $\tilde{\sigma}_i^2(t, k, t_1, t_2)$, and $\tilde{\sigma}_i^2(t, k, t_1, t_2, x)$. If one is interested only in the estimates of total uncertainty, then the first two of these need not be computed. For present purposes, however, all three were computed, which means that three 15-quarter stochastic simulations were performed per model. For each simulation 1000 trials were made for Model I and 2000 trials were made for Model II. The results of these computations for six variables and the 15 quarters are presented in Tables 2 and 3. The square roots of the estimated variances are presented in Table 2, and the estimated forecast means are presented in Table 3.

The misspecification part of the method requires successive reestimation and stochastic simulation. For Model I the first estimation period was taken to be 1954I–1968IV, the second to be 1954I–1969I, and so on through the last, 1954I–1977II. These same periods were used for Model II except that the beginning quarter was always 1954II rather than 1954I. This is a total of 35 sets of estimates, and for each set an 8-quarter stochastic simulation (of 100 trials for Model I and 500 trials for Model II) was performed per model.[8] This allowed 35 values of $d_i(\cdots)$ to be computed for each variable for the one-quarter-ahead forecast, 34 values for the two-quarter-ahead forecast, and so on. This then allowed the means $\bar{d}_i(k)$, $k=1,\ldots, 8$, to be computed. Given $\tilde{\sigma}_i^2(t, k, t_1, t_2, x)$ from the simulation for the basic forecast period, this in turn allowed $\tilde{\sigma}_i^2(t, k, t_1, t_2, x, d)$ to be computed (see equations (4) and (5)). The square roots of these latter estimates are presented in the *d* rows in Table 2. These are the estimates of the total uncertainty of the forecast. For these results the formula for variables with a trend (equation (5)) was used for real GNP, the GNP deflator, the wage rate, and the money supply; and the formula for variables without a trend (equation (4)) was used for the unemployment rate and the bill rate.

The following is a discussion of the results in Tables 2 and 3. The results for Model I will be discussed first, and then they will be compared to the results for Model II.

[7] Because of data requirements due to lags, the beginning quarter for Model II had to be one quarter later than the beginning quarter for Model I.

[8] Since the data ended in 1977IV, the simulations for the last 7 sets were shorter than 8 quarters. The first quarter of each simulation period was taken to be the second quarter after the end of the estimation period. This was done to be consistent with the procedure followed for the basic forecast period. Partly because of cost considerations and partly because of a relative small number of outside-sample observations, the length of the simulation periods for this part of the method was taken to be 8 rather than 15 quarters.

The results in Table 3 should provide some encouragement to model builders. They show that the forecast values computed by setting the error terms equal to zero and solving once are quite close to the forecast values computed by means of stochastic simulation. Although it is well known (see, for example, Howrey and Kelejian [1971]) that the common practice of setting the error terms to zero and solving once produces biased estimates of the true means of the endogenous variables for nonlinear models, this bias does not appear to be very large, at least for Model I.[9]

Consider now Table 2. The results in the *a*, *b*, and *c* rows are self explanatory. As might be expected, the sensitivity of the standard errors of the forecasts to exogenous-variable uncertainty (rows *c* versus *b*) is greater for some variables than for others. This sensitivity is small for the unemployment rate, the wage rate, and the bill rate and fairly large for the money supply. Although in most cases these sensitivity differences can be explained, given a knowledge of the structure of the model, these explanations are unnecessary for purposes of this paper.

The numbers in the *d* rows are the estimates of the total uncertainty of the forecasts. A brief summary of them is as follows. For the four-quarter-ahead forecasts, the estimated standard errors are 1.96 percent (27.8 billion dollars)[10] for real GNP, 1.87 percent for the GNP deflator, 0.82 percentage points for the unemployment rate, 2.01 percent for the wage rate, 1.17 percentage points for the bill rate, and 3.75 percent (14.9 billion dollars) for the money supply. For the eight-quarter-ahead forecasts, the estimated standard errors are 2.27 percent (33.3 billion dollars) for real GNP, 3.48 percent for the GNP deflator, 0.71 percentage points for the unemployment rate, 4.16 percent for the wage rate, 1.72 percentage points for the bill rate, and 7.50 percent (32.2 billion dollars) for the money supply.

Consider now Model I versus Model II. With respect to the estimates of the total uncertainty of the forecasts in the *d* rows, Model II is less accurate than Model I for real GNP, the GNP deflator, the unemployment rate, and the bill rate, and it is more accurate for the wage rate and the money supply. For the eight-quarter-ahead forecasts, the differences in the estimated standard errors are 2.47 percent for real GNP, 2.78 percent for the GNP deflator, 1.48 percentage points for the unemployment rate, −2.12 percent for the wage rate, 0.11 percentage points for the bill rate, and −3.80 percent for the money supply.

[9] Remember, however, that the stochastic-simulation estimates themselves are not quite right in that they are based on a limited number of trials and on only estimated coefficients and covariance matrices. Also, it is not even the case that the true expected values, $\bar{y}_i(t, k)$, $\bar{y}_i(t, k, t_1, t_2)$ and $\bar{y}_i(t, k, t_1, t_2, x)$, are necessarily the same. The results in Table 3 thus do not provide a completely accurate estimate of the bias that results from setting the error terms equal to zero. The conclusion reached here that the bias is small has also been reached by Nagar [1969], Sowey [1973], Cooper [1974], Bianchi, Calzolari and Corsi [1976], and Calzolari and Corsi [1977] for their stochastic simulations with respect to the error terms only.

[10] Any dollar figure used in this section has been obtained by multiplying the particular percent figure in Table 2 by the relevant number in the *c* rows in Table 3.

Given my wage rate, I would conclude from the results in Table 2 that Model I is enough of an improvement over Model II to justify the time that I have so far spent developing and working on it. The differences in the standard errors for real GNP, the GNP deflator, and the unemployment rate are substantial. It is, of course, somewhat embarrassing that Model I is less accurate with respect to the forecasts of the wage rate and the money supply than Model II. There is not too much that can be said about this except that I was aware before, and even more so now, that the wage-rate equation and one of the demand-for-money equations are two of the weakest equations in Model I, weakest in the sense that the coefficient estimates of these two equations tend to change more as the model is reestimated on the basis of new data than do the coefficient estimates of most of the other equations. There are clearly grounds for further work on these two equations.

One further point about the negative results for the money supply for Model I should be noted. There is some evidence that indicates that the demand-for-money equations in Model I are more accurate than other demand-for-money

TABLE 4

ROOT MEAN SQUARED ERRORS OF OUTSIDE-SAMPLE FORECASTS

		Number of Quarters Ahead							
Variable	Model	1	2	3	4	5	6	7	8
Real GNP	I	0.79	1.26	1.63	2.12	2.59	2.97	3.24	3.52
(percent)	II	1.11	1.96	2.76	3.51	4.09	4.42	4.70	4.91
GNP Deflator	I	0.50	0.93	1.43	1.97	2.49	2.95	3.43	3.83
(percent)	II	0.47	0.98	1.59	2.36	3.26	4.23	5.35	6.52
Unemployment Rate	I	0.36	0.60	0.75	0.80	0.79	0.79	0.77	0.77
(percentage points)	II	0.36	0.75	1.13	1.49	1.73	1.89	2.03	2.14
Wage Rate	I	0.78	1.25	1.71	2.31	2.89	3.49	4.06	4.65
(percent)	II	0.67	0.94	1.21	1.47	1.71	2.01	2.40	2.83
Bill Rate	I	0.61	1.01	1.16	1.31	1.50	1.70	1.86	2.06
(percentage points)	II	0.70	1.15	1.37	1.52	1.64	1.71	1.76	1.84
Money Supply	I	1.34	2.12	2.93	3.81	4.74	5.72	6.82	7.97
(percent)	II	1.41	1.71	2.10	2.52	2.89	3.23	3.64	4.09

Notes: i) These results are based on 35 sets of estimates of each model.

ii) Each eight-quarter outside-sample forecast began two quarters after the end of the estimation period. The first estimation period ended in 1968IV, and the last (the thirty-fifth) ended in 1977II. Data through 1977IV were used, which allowed 35 one-quarter-ahead errors to be computed for each variable, 34 two-quarter-ahead errors, and so on. The actual values of the exogenous variables were used for these calculations.

iii) The predicted values used were the mean values from the 35 stochastic simulations that were performed for the *d*-row results in Table 2.

equations. As discussed in Fair [1978b, fn. 6, p. 1169], the demand-for-money equation in Model I appear to be considerably more accurate for the 1973I–1976I period than the demand-for-money equation in the MPS model. The problems noted here regarding the demand-for-money equation in Model I are thus probably not unique to Model I.

This completes the discussion of the results in Table 2. For comparison with these results, root mean squared errors of the outside-sample forecasts have also been computed, and these errors are presented in Table 4. These RMSEs are based on the predicted values from the same 35 stochastic simulations that were used for the d-row results in Table 2. Comparing the numbers in the d rows in Table 2 to the respective numbers in Table 4, it can be seen that there are some sizeable differences. This is, of course, as expected, since the RMSE procedure ignores exogenous-variable uncertainty and the fact that forecast-error variances vary across time. It is true, however, that the overall ranking of the accuracy of the two models by variable is the same in the two tables.

4. ESTIMATING THE ACCURACY OF SUBJECTIVELY-ADJUSTED MODELS

Although the method described above is not relevant for models that are subjectively adjusted, it can be modified for such models. In particular, the following procedure could be followed for subjectively-adjusted models. (1) Treat the model mechanically and perform the calculations necessary for results like those in Table 2. (2) Over a period of a few years compile an *ex ante* forecasting record for both the model used mechanically and the model used subjectively. Let $e_i^m(t, k)$ denote the error of the k-quarter-ahead forecast of variable i for quarter t from the model used mechanically (the forecast starting at the beginning of quarter $t-k$) and let $e_i^s(t, k)$ denote the similar error for the model used subjectively. Let $\delta_i(t, k)$ be the difference in the errors squared: $\delta_i(t, k) = (e_i^m(t, k))^2 - (e_i^s(t, k))^2$. After, say, 12 values of $\delta_i(t, k)$ have been compiled, take the average of these values. Denote this average as $\bar{\delta}_i(k)$. $\bar{\delta}_i(k)$ will be positive if subjectively adjusting the model has on average improved its forecasting accuracy. (3) If it is assumed that the degree to which subjectively adjusting a model improves its forecasting accuracy with respect to a given variable and length of forecast is constant across time, then the $\bar{\delta}_i(k)$ values can be subtracted in the appropriate way from the numbers in the d rows in Table 2 to get a final estimate of the uncertainty of the forecasts from the subjectively-adjusted model.[11]

The constancy assumption in (3) is, of course, much stronger than the constancy assumption needed for the results in Table 2, but the $\bar{\delta}_i(k)$ values would be at least rough approximations of the degree to which subjectively adjusting a model improves its forecasting accuracy. Of more scientific interest, of course,

[11] The appropriate way would be to subtract $\bar{\delta}_i(k)$ from the square of the respective number in row d of Table 2 and then to take the square root of this difference. This number would then be the final estimate of the standard error of the forecast from the subjectively-adjusted model.

would be the mechanical results themselves, for only by observing results like those in Table 2 for models used mechanically can one hope to learn about the models in ways that are useful for further scientific research.

Cowles Foundation, Yale University, U.S.A.

APPENDIX

THE STOCHASTIC-SIMULATION PROCEDURES FOR MODELS I AND II

The exact procedures that were followed for the stochastic simulations are discussed in this Appendix. There are a number of ways in which stochastic simulation can be carried out, and it is important to note that the method in Section 2 does not require any one particular way. In fact, as will be discussed, some of the following assumptions are restrictive, and in future work it will be of interest to try to relax them. The simulations for Model I will be discussed first.

1. *Estimating* $\sigma_i^2(t, k)$. For the estimation period of 96 observations (1954I–1977IV), consistent estimates of the 29 error terms are available (from consistent two-stage least squares (TSLS) coefficient estimates). The covariance matrix of the 29 error terms (Ω) was estimated as $(1/96)EE'$, where E is the 29×96 matrix of values of the estimated error terms. In conformity with the notation in Table 1, this estimate of $\hat{\Omega}$ will be denoted $\hat{\Omega}(9, 104)$, where 9 is quarter 1954I and 104 is quarter 1977IV.

Let $e(t)$ denote the 29×1 vector of values of the error terms for quarter t: $e(t) = (e_1(t), \ldots, e_{29}(t))'$. Values of $e(t)$ were drawn from a multivariate normal distribution with mean zero and covariance matrix $\hat{\Omega}(9, 104)$ for the stochastic simulation. Since the prediction period is 15 quarters, each "trial" corresponds to drawing 15 values of $e(t)$ (29×15 numbers in all) and computing the forecast using these values.[1] In this first case, where only the uncertainty from the error terms is being estimated, the coefficient estimates and exogenous-variable values are kept the same for all the trials.

For each trial, one obtains a forecast of each endogenous variable for each quarter. Let $\tilde{y}_i^j(t, k)$ denote the value of the k-quarter-ahead forecast of variable i for quarter t on the j-th trial. For J trials, the estimate of the expected value of the forecast, $\tilde{\bar{y}}_i(t, k)$, is $(1/J) \sum_{j=1}^{J} \tilde{y}_i^j(t, k)$, and the estimate of the variance of the forecast error, $\tilde{\sigma}_i^2(t, k)$, is $(1/J) \sum_{j=1}^{J} [\tilde{y}_i^j(t, k) - \tilde{\bar{y}}_i(t, k)]^2$. The number of

[1] The draws were performed as follows. First, a matrix P was computed such that $PP' = \hat{\Omega}(9,104)$. This was done using the LUDECP subroutine in the IMSL library. Then for each of the 15 quarters, 29 values of a standard normal random variable with mean 0 and variance 1 were drawn. This was done using the function RNOR, which is part of the SUPER DUPER random number generator package at Yale. Let $u(t)$ denote the 29×1 vector of these draws for quarter t. Then $e(t)$ was computed as $Pu(t)$. Since $Eu(t)u'(t) = I$, then $Ee(t)e'(t) = EPu(t)u'(t)P' = \hat{\Omega}(9,104)$, which is as desired for the distribution of $e(t)$.

trials used for these estimates was 1000.[2]

2. *Estimating* $\sigma_i^2(t, k, t_1, t_2)$. For each of the 29 stochastic equations an estimate of the covariance matrix of the coefficient estimates is available. Let $\hat{\beta}_i(9, 104)$ denote the vector of TSLS coefficient estimates for equation i, let $\hat{V}_i(9, 104)$ denote the estimated covariance matrix of these estimates, and let β_i^* denote the vector of coefficient values for equation i actually used in a given trial. Values of β_i^* were drawn from a multivariate normal distribution with mean $\hat{\beta}_i(9, 104)$ and covariance matrix $\hat{V}_i(9, 104)$ for the stochastic simulation. In this case, where the uncertainty from the error terms and coefficient estimates is being estimated, each trial corresponds to drawing 15 values of $e(t)$ (29×15 numbers) and a value of β_i^* for each of the 29 equations (183 numbers).[3] Aside from drawing 183 extra numbers for each trial, this case is the same as the first case. The number of trials used in this case was also 1000.

3. *Estimating* $\sigma_i^2(t, k, t_1, t_2, x)$. Not counting variables like the constant, time, and various dummy variables, there are 60 exogenous variables in Model I. Each of these variables was regressed on a constant, time, and its first eight lagged values for the 1954II–1977IV sample period, and for each equation the variance of the error term was estimated as the sum of squared residuals divided by the number of observations. Let this estimated variance for variable i be denoted s_i^2. Also, let $u_i(t)$ be a normally distributed random variable with mean zero and variance s_i^2:

$$u_i(t) \sim N(0, s_i^2), \qquad \text{all } t. \tag{1}$$

Let $\hat{x}_i(t)$ be the value of exogenous variable i for quarter t that was used for the actual forecast, and let $x_i^*(t)$ be the value used in a given trial. Also, let r denote the first quarter of the prediction period (1978II). For the stochastic simulation, 60×15 values of $u(t)$ were drawn ($i = 1, \ldots, 60$; $t = r, r+1, \ldots, r+14$), and these values were taken to be the errors in forecasting the changes in the exogenous variables. In particular, the values of $x_i^*(t)$ are:

$$x_i^*(r) = \hat{x}_i(r) + u_i(r) \tag{2}$$

[2] I could see no obvious way to use any of the tricks in, for example, Hammersley and Handscomb [1974] to increase the efficiency of the stochastic simulation, and so each trial was merely an independent random draw. Each trial, which consists of solving the model once for 15 quarters, takes about 2.0 seconds on the IBM 370–158 at Yale, so the total time for 1000 trials is about 33.3 minutes.

[3] The draws for the β_i^* vectors were performed as follows. First, for each $\hat{V}_i(9,104)$, a matrix P_i was computed such that $P_i P_i' = \hat{V}_i(9,104)$. Then for each i, n_i values of a standard normal random variable with mean 0 and variance 1 were drawn, where n_i is the number of coefficients in equation i. Let u_i denote the $n_i \times 1$ vector of these draws. Then β_i^* was computed as $\hat{\beta}_i(9,104) + P_i u_i$. Since $E u_i u_i' = I$, then $E(\beta_i^* - \hat{\beta}_i(9,104))(\beta_i^* - \hat{\beta}_i(9,104))' = E P_i u_i u_i' P_i' = \hat{V}_i(9,104)$, which is as desired for the distribution of β_i^*. Subroutine LUDECP and function RNOR were also used for these calculations.

$$x_i^*(r+1) = \hat{x}_i(r+1) + u_i(r) + u_i(r+1)$$
$$\vdots$$
$$x_i^*(r+14) = \hat{x}_i(r+14) + u_i(r) + u_i(r+1) + \cdots + u_i(r+14).$$

Because of the assumption that the errors pertain to the changes in the exogenous variables, the error term $u_i(r)$ is carried along from quarter r on. Similarly, $u_i(r+1)$ is carried along from quarter $r+1$ on, and so on.

In this case each trial corresponds to drawing 29×15 numbers for the error terms, 183 numbers for the coefficient estimates, and 60×15 values of $u_i(t)$ for the exogenous variables.[4] Aside from drawing 60×15 extra numbers, this case is the same as the previous case. The number of trials used in this case was also 1000.

The assumption that the $u_i(t)$ errors pertain to forecasting the change in the exogenous variables perhaps requires some discussion. Given the way that many exogenous variables are forecast, by extrapolating past trends or taking variables to be unchanged from their last observed values, it seems likely that any error made in forecasting the level of a variable in, say, the first quarter will persist through the forecast period. If this is true, then the present assumption seems better than the assumption that the $u_i(t)$ errors pertain to forecasting the level of the exogenous variables.

4. *Computing* $\bar{d}_i(k)$. 35 sets of estimates of the model were obtained. For all sets the first quarter of the sample period was 1954I ($t_1=9$). The last quarter of the sample period was 1968IV ($t_2=68$) for the first set, 1969I ($t_2=69$) for the second set, and so on through 1977II ($t_2=102$) for the 35-th set.

For each set of estimates the model was stochastically simulated for 8 quarters beginning with the second quarter after the end of the estimation period. The 35 stochastic simulations were performed in the same way as described above for the stochastic simulation with respect to the error terms and coefficient estimates. The first simulation used $\hat{\beta}_i(9, 68)$, $\hat{\Omega}_i(9, 68)$, and $\hat{V}_i(9, 68)$; the second simulation used $\hat{\beta}_i(9, 69)$, $\hat{\Omega}_i(9, 69)$, and $\hat{V}_i(9, 69)$; and so on. Values of $\tilde{\bar{y}}_i(69+k, k, 9, 68)$ and $\tilde{\sigma}_i^2(69+k, k, 9, 68)$ for $k=1,\ldots, 8$ were obtained from the first simulation; values of $\tilde{\bar{y}}_i(70+k, k, 9, 69)$ and $\tilde{\sigma}_i^2(70+k, k, 9, 69)$ for $k=1,\ldots, 8$ were obtained from the second simulation; and so on. The only difference between these 35 simulations and the simulation described above is that each of the 35 simulations was for 8 quarters rather than 15 and the number of trials was 100 rather than 1000.[5]

Data through 1977IV ($t=104$) were used for this work (the preliminary data for 1978I were not used), which meant that 35 values of $d_i(t, 1, 9, t-2)$ could be computed ($t=70,\ldots, 104$); 34 values of $d_i(t, 2, 9, t-3)$ could be computed ($t=$

[4] Drawing the 60×15 values of $u_i(t)$ is straightforward. For each i, 15 values of a standard normal random variable with mean 0 and variance 1 are drawn, and then each of these values is multiplied by s_i.

[5] For the eight-quarter simulations each trial takes about 1.1 seconds of computer time, so the total time for the 3500 trials was about 65 minutes.

71,..., 104); and so on. The estimates $\bar{d}_i(1)$ were thus based on 35 observations; the estimates $\bar{d}_i(2)$ on 34 observations; and so on.

5. *Model* II. The procedure followed for Model II is quite similar to the procedure followed for Model I, and so the discussion here can be brief. Each of the six equations of Model II (corresponding to the six variables in Table 2) was treated individually: no attempt was made to estimate and account for the possible correlation of the error terms across equations. The variance of the error term in each equation was estimated as the sum of squared residuals divided by the number of observations. The error term in each equation was then assumed to be normally distributed with mean zero and variance as estimated for purposes of the stochastic simulations. The coefficient estimates were treated in the same way as they were for Model I. The number of trials for both the simulation with respect to the error terms and the simulation with respect to the error terms and coefficient estimates was 2000. There are no non-trivial exogenous variables in Model II, and so no stochastic simulation regarding the exogenous variables was needed. For the calculations of $\bar{d}_i(k)$, Model II was also estimated 35 times, and the same periods were used here as were used for Model I except that the beginning quarter was 1954II rather than 1954I. The number of trials for each of the 35 stochastic simulations was 500.

6. *General Remarks.* A few more points about the procedures in this Appendix should be mentioned. Note first that the normality assumption has been used for all the simulations. Although this is the most convenient assumption to make and can be justified by an appeal to large sample properties, in some cases it may not be very realistic. It may be of interest in future work to examine the sensitivity of results like those in Section 3 to departures from normality. Note also that the estimated covariance matrix of the TSLS coefficient estimates was taken to be block diagonal. This matrix, which is 183×183, is in fact not block diagonal, and so an additional source of simulation error has been introduced by the present treatment. It is somewhat more expensive to deal with the full covariance matrix, but in future work it may be possible to do this. It should finally be noted in this regard that the correlation between the estimate of the serial correlation coefficient in an equation and the other coefficient estimates in the equation has been taken into account in computing the covariance matrices.[6]

No degrees-of-freedom corrections were made in this study. For Model II, which was estimated by ordinary least squares, there are well defined degrees-of-freedom corrections that could have been made, but this is not the case for Model

[6] The technique that was used to estimate Model I is described in Fair [1970]. It was suggested in this paper (p. 514) that the covariance matrix of the coefficient estimates inclusive of the estimate of the serial correlation coefficient be estimated by ignoring the correlation between the latter estimate and the other coefficient estimates. Fisher, Cootner and Baily [1972, fn. 6, p. 575], however, have pointed out that one need not ignore this correlation. In computing the estimates of the covariance matrices for use in this study, I have followed the Fisher, Cootner and Baily advice.

I. It thus seemed best to put both models on a comparable basis by not adjusting for degrees of freedom in either model. With at most 13 estimated coefficients per equation, the results in Section 3 would not have been much different had some adjustment for degrees of freedom been made. This is, of course, not necessarily true for models with many coefficients per equation, and for models of this type one may want to adjust for degrees of freedom.

It should finally be stressed that the particular treatment of exogenous variable uncertainty in this Appendix is only one of many that might be tried. In particular, it may be of interest in future work to experiment with the polar assumption that the exogenous-variable forecasts are as uncertain as the endogenous-variable forecasts (as reflected in, say, a set of autoregressive equations). Another possibility would be to make this assumption for all but exogenous fiscal-policy variables.

REFERENCES

BIANCHI, C., G. CALZOLARI AND P. CORSI, "Divergences in the Results of Stochastic and Deterministic Simulation of an Italian Non Linear Econometric Model," in, L. Dekker, ed., *Simulation of Systems* (Amsterdam: North-Holland Publishing Co., 1976).

BROWN, R. L., J. DURBIN AND J. M. EVANS, "Techniques for Testing the Constancy of Regression Relationships over Time," *Journal of the Royal Statistical Society*, Series B, 37, No. 2 (1975), 149–163.

CALZOLARI, G. AND P. CORSI, "Stochastic Simulation as a Validation Tool for Econometric Models," paper presented at IIASA Seminar, Laxenburg, Vienna, September 13–15, 1977.

COOPER, J. P., *Development of the Monetary Sector, Prediction and Policy Analysis in the FRB-MIT-Penn Model* (Lexington: D. C. Heath & Co., 1974).

——— AND S. FISCHER, "Stochastic Simulation of Monetary Rules in Two Macroeconometric Models," *Journal of the American Statistical Association*, 67 (December, 1972), 750–760.

———, "Monetary and Fiscal Policy in the Fully Stochastic St. Louis Econometric Model," *Journal of Money, Credit and Banking*, 6 (February, 1974), 1–22.

EVANS, M. K., L. R. KLEIN AND M. SAITO, "Short-Run Prediction and Long-Run Simulation of the Wharton Model," in, B. G. Hickman, ed., *Econometric Models of Cyclical Behavior* (New York: Columbia University Press, 1972).

FAIR, R. C., "The Estimation of Simultaneous Equation Models with Lagged Endogenous Variables and First Order Serially Correlated Errors," *Econometrica*, 38 (May, 1970), 507–516.

———, *A Model of Macroeconomic Activity. Volume II: The Empirical Model* (Cambridge: Ballinger Publishing Co., 1976).

———, "The Fair Model as of April 15, 1978," mimeo (1978a).

———, "The Sensitivity of Fiscal-Policy Effects to Assumptions about the Behavior of the Federal Reserve," *Econometrica*, 46 (September, 1978b), 1165–1179.

FISHER, F. M., P. H. COOTNER AND M. N. BAILY, "An Econometric Model of the World Copper Industry," *The Bell Journal*, 3 (Autumn, 1972), 568–609.

FROMM, G., L. R. KLEIN AND G. R. SCHINK, "Short- and Long-Term Simulations with the Brookings Model," in, B. G. Hickman, ed., *Econometric Models of Cyclical Behavior* (New York: Columbia University Press, 1972).

GARBADE, K. D., *Discretionary Control of Aggregate Economic Activity* (Lexington: D. C. Heath & Co., 1975).

GREEN, G. R., M. LIEBENBERG AND A. A. HIRSCH, "Short- and Long-Term Simulations with the OBE Econometric Model," in, B. G. Hickman, ed., *Econometric Models of Cyclical Behavior* (New York: Columbia University Press, 1972).

HAITOVSKY, Y. AND N. WALLACE, "A Study of Discretionary and Nondiscretionary Monetary and Fiscal Policies in the Context of Stochastic Macroeconometric Models," in, V. Zarnowitz, ed., *The Business Cycle Today* (New York: Columbia University Press, 1972).

HAMMERSLEY, J. M. AND D. C. HANDSCOMB, *Monte Carlo Methods* (London: Methuen & Co., 1974).

HOWREY, E. P. AND H. H. KELEJIAN, "Simulation versus Analytical Solutions: The Case of Econometric Models," in, T. H. Naylor, ed., *Computer Simulation Experiments with Models of Economic Systems* (New York: John Wiley, 1971).

MUENCH, T., A. ROLNICK, N. WALLACE AND W. WEILER, "Tests for Structural Change and Prediction Intervals for the Reduced Forms of Two Structural Models of the U.S.: The FRB-MIT and Michigan Quarterly Models," *Annals of Economic and Social Measurement*, 3 (July, 1974), 491–519.

NAGAR, A. L., "Stochastic Simulation of the Brookings Econometric Model," in, J. S. Duesenberry, G. Fromm, L. R. Klein and E. Kuh eds., *The Brookings Model: Some Further Results* (Chicago: Rand McNally & Co., 1969).

SCHINK, G. R., Ph.D Dissertation, University of Pennsylvania (1971).

———, "Estimation of Small Sample Forecast Error for Nonlinear Dynamic Models: A Stochastic Simulation Approach," (May, 1974).

SOWEY, E. R., "Stochastic Simulation for Macroeconomic Models: Methodology and Interpretation," in, A. A. Powell and R. A. Williams, eds., *Econometric Studies of Macro and Monetary Relations* (Amsterdam: North-Holland Publishing Co., 1973).