Confidence intervals of forecasts from nonlinear econometric models

Bianchi, Carlo and Calzolari, Giorgio

IBM Scientific Center, Pisa, Italy.

5 June 1983

CONFIDENCE INTERVALS OF FORECASTS FROM NONLINEAR ECONOMETRIC MODELS

by Carlo BIANCHI and Giorgio CALZOLARI

Centro Scientifico IBM

Pisa

ABSTRACT

Several methods have been proposed in the last few years for evaluating uncertainty in forecasts produced by nonlinear econometric models. Some methods resort to Monte Carlo, while others resort to different simulation techniques. This work aims at comparing these methods by means of experiments on some econometric models of small, medium and large size, used in practice for forecasting purposes. In most cases of practical interest, direct simulation of confidence intervals allows to overcome the difficulties connected with the nonexistence of finite second order moments, often encountered by the authors when applying Monte Carlo methods to real world models.

CONTENTS

1. INTRODUCTION

Some papers, in the last few years, have dealt with the problem of evaluating the degree of uncertainty associated with forecasts produced by macro economic models. Alternative methods have been proposed, which estimate the variance of the forecast errors, when forecasts are produced with systems of simultaneous equations. These methods are of three types.

1) Full analytical methods: they were originally designed for linear systems (e.g. Goldberger et al., 1961, or Schmidt, 1974), but even in case of models containing nonlinearities, these methods can be applied to solve a good deal of the problem (Calzolari, 1981).
2) Mixed methods, partially analytical and partially based on numerical simulation procedures (analytic simulation): conceptually equivalent to the full analytical methods, they allow for a considerable reduction of computational complexity and are suitable for application even to medium-large size models (Bianchi and Calzolari, 1980).
3) Monte Carlo methods: estimates of the variances are computed from sample variances of replicated simulation experiments, after additive pseudo-random errors have been inserted into the structural equations of the model (Schink, 1971), or even into model's coefficients (Fair, 1980).

In two recent papers the authors have performed some experiments on a set of small, medium and large size real world models, both linear and nonlinear, comparing the results and performances of the different methods proposed in the literature for the computation of variances of forecasts. Comparisons were first confined to the case of forecasts one period ahead (Bianchi and Calzolari, 1982), then to the case of multiperiod forecasts produced with dynamic simulation (1983). Two main conclusions were drawn from the experimental comparison of the methods on models used, in practice, for forecasting purposes.

1) All the Monte Carlo methods proposed in the literature for this purpose involve, to some extent, a risk of nonconvergence. This is because at some stage of the process the Monte Carlo methods are faced with the problem of computing the moments of random variables whose distribution may have no finite moment of any order. Forecasts produced by a system of simultaneous equations are, in fact, strictly connected with the reduced form coefficients, and there is no warranty of existence of finite moments for the small sample distribution of these coefficients, when derived from structural form estimators like 2SLS or 3SLS (see, for example, the discussion in Mariano, 1982, p.520). Empirical experiments simply showed that the risk may be purely theoretical for some models, but actual for others.
2) When convergence of the Monte Carlo methods was ensured, it was not rare the case of diverging results for the second order moments of forecasts obtained with the different methods. The alternative Monte Carlo methods may lead to results substantially different from one another, and different from results obtained analytically or from analytic simulation. These differences were often encountered in case of multiperiod dynamic simulation forecasts, rarely in case of one-period (static) forecasts.

Of course, the case of coincident or at least similar results obtained with all the methods does not give particular troubles.

The model builder can use the resulting standard errors either as simple indicators of the degree of uncertainty associated with his forecasts, or to perform the test of hypotheses, or finally to build confidence intervals of forecasts.

The case of large differences, on the contrary, leaves the model builder doubtful about the amount of uncertainty involved in his forecasts. Can he overcome this difficulty? Not completely, unfortunately, but to some extent. This paper aims at showing, with empirical evidence on real world models, that if all Monte Carlo methods are used for direct construction of forecast intervals, without intermediate computation of the moments of forecast errors, the cases where large differences are encountered become much rarer.

2. NOTATIONS AND ASSUMPTIONS

The following notations will be adopted in this paper. Let the structural econometric model be represented as

$$(2.1) \qquad f(y_t, y_{t-1}, x_t, a) = u_t ; \qquad t=1,2,\dots,T$$

where $f = (f_1, f_2, \dots, f_m)'$ is a vector of functional operators, continuously differentiable with respect to the elements of current and lagged y, x and a; $y_t = (y_{1t}, y_{2t}, \dots, y_{mt})'$, $x_t = (x_{1t}, x_{2t}, \dots, x_{nt})'$ and y_{t-1} are the vectors of current endogenous, exogenous and lagged endogenous variables, respectively; $a = (a_1, a_2, \dots, a_s)'$ is the vector of the structural coefficients to be estimated (all the other known coefficients of the model are excluded from this vector and included in the functional operators); $u_t = (u_{1t}, u_{2t}, \dots, u_{mt})'$ is the vector of structural stochastic disturbances (or error terms) at time t, having zero mean and being independently and identically distributed over time, with finite contemporaneous covariance matrix, and independent of all the predetermined variables. In all the experiments described in this paper, the contemporaneous distribution of the error terms is assumed multivariate normal: $u_t \sim N(0, \Sigma)$.

It is usually assumed that a simultaneous equation system like (2.1) implicitly defines a single inverse relationship (reduced form) for relevant values of the coefficients, the predetermined variables, and any values of the disturbance terms:

$$(2.2) \qquad y_t = g(y_{t-1}, x_t, a, u_t).$$

Of course, the vector of functions g implicitly defined is usually unknown, but can be assumed continuously differentiable, like f.

Processing the sample data by means of a suitable estimation method, we get an estimated vector of coefficients, $\hat{a}$, an estimated covariance matrix of the structural disturbance process, $\hat{\Sigma}$, and an estimate of the coefficients covariance matrix which will be indicated as $\hat{\Psi}/T$. Four remarks may be helpful at this point.

1) Few assumptions in addition to those listed above are usually enough to ensure consistency and asymptotic normality of $\hat{a}$, produced by suitable estimation methods, in case of linear dynamic or nonlinear static models. In these cases, it is usually ensured that, asymptotically as $T \to \infty$,

$$(2.3) \qquad \sqrt{T}\,(\hat{a} - a) \sim N(0, \Psi).$$

2) Unfortunately, to our knowledge, there are no general theoretical tools to prove that (2.3) holds when the nonlinear model includes lagged endogenous variables among the predetermined variables. It can, however, be assumed that (2.3) holds under heuristic considerations, as in Gallant (1977, pp.73-74). If (2.3) holds, then several results which will be derived are asymptotically exact; if (2.3) does not hold exactly, the results which will be derived are not asymptotically exact, but simply "reasonable" approximations.
3) If $\hat{\Psi}$ is a consistent estimate of Ψ, an estimate of the covariance matrix of a multinormal distribution which approximates the small sample distribution of $\hat{a}$ is obtained as $\hat{\Psi}/T$, that is dividing $\hat{\Psi}$ by the actual length of the sample period (see Schmidt, 1976, p.254).
4) $\hat{\Psi}/T$ is, together with $\hat{a}$, a standard outcome of system estimation methods. For example, in case of FIML estimation, $\hat{\Psi}/T$ is nothing but the inverse of the Hessian (with minus sign) of the concentrated log-likelihood, calculated at the point which maximizes the likelihood. When limited information estimation methods are applied, this matrix must be built block by block, after the estimated coefficients have been obtained, and the resulting matrix may be singular, in case of undersized samples.

Let us now introduce an additional simplification, which is crucial: we disregard from misspecification. Of course, the results which will be obtained could be used for tests of misspecification or, more simply, for empirical measurement of the misspecification effects (e.g. Fair, 1980), however we shall work, in the rest of the paper, as if equation (2.1) represents the "true" structure of the economic system and, therefore, the "true" process which generates endogenous variables is represented by equation (2.2). An assumption like this is, of course, considerably hazardous and unrealistic, but quite helpful and useful. We should not forget, in fact, that we are going to analyze and measure the degree of uncertainty associated with forecasts obtained from simulation experiments. Roughly speaking, if a measure of uncertainty is obtained under the assumption of correct specification, it should be interpreted as the "minimum" degree of uncertainty associated with a given simulation result. In other words, we get in this way a kind of empirical upper bound for the reliability of model's forecasts. Forecasts obtained with the usual simulation procedure, which is deterministic, are in some way related to the reduced form notation (2.2) where a is replaced by its estimate, $\hat{a}$, and the random error terms u_t are set to their expected value (zero). For example, if we are interested in one period forecasts at time h, given all available information on endogenous variables at time $h-1$, and conditional on the exact knowledge of all the exogenous variables at time h, we calculate the values of the endogenous variables $\hat{y}_h$ such that $f(\hat{y}_h, y_{h-1}, x_h, \hat{a})=0$, that is, resorting to the reduced form notation,

$$(2.4) \qquad \hat{y}_h = g(y_{h-1}, x_h, \hat{a}, 0).$$

Generalization to dynamic simulation is straightforward (see section 4).

Whichever use is made of the model, we must not forget that, while (2.2) is the process which generates the real data, what we

are using is the simulation process (2.4). The differences in the two processes are given by the presence, in the latter, of the random vector of estimates, $\hat{a}$, instead of the vector of constant "true" coefficients a, and by the presence of the random error terms u_t in the former. These two differences, which cause the uncertainty of simulated results, are the two error sources whose effects will be measured in this paper.

We shall try to isolate the effects of these two error sources, as far as possible. When separation of effects can be made, the two error sources can be treated with different computation procedures. Although both error sources have, in some sense, a common origin, it is extremely important to distinguish between their different behaviour.

3. THE TWO ERROR SOURCES

The structural disturbance vector, u_t, is a vector of zero mean random error terms embodied in the model by the very nature of the endogenous variables, which are random variables. Whichever sample period length we have at our disposal, and whatever estimation method we apply, we cannot reduce the size of the error caused by u_t. All we can get, from using more and more efficient estimation methods, is a more accurate estimate of the error process covariance matrix, $\hat{\Sigma}$, which however remains a nonzero matrix. Even if all the other variables and parameters of the model were known with certainty, the simulation process would differ from the real data generation process by a nonlinear function of u_t.

The simplest (and perhaps the only possible, in case of medium-large models) way of treating nonlinear transformations of random variables with finite variances and covariances is surely stochastic simulation (Monte Carlo). We may proceed as follows.

First of all a vector of pseudo-random error terms must be generated for each period of time to which simulation results are related (one only, for instance, if we are dealing with one-period forecasts). The distribution of these pseudo-random vectors must be as close as possible to the distribution of u_t; a suitable choice may be the multinormal distribution with zero mean and covariance matrix equal to the available estimate $\hat{\Sigma}$. When we are dealing with a case of undersized samples, as for most models used in practice for forecasting purposes, a suitable generation method has been proposed in McCarthy (1972). The generated error terms are inserted into the model, and the usual simulation procedures are applied to produce forecasts. The whole procedure is repeated from the beginning a given (possibly large) number of times, in such a way as to produce a sample of outcomes, from which we calculate sample means and variances. A number of replications more and more increasing is, usually, expected to produce more accurate values of means and variances. Analytical approximations, based on model's linearization, are possible, but they are not exact; we shall always resort to stochastic simulation when dealing with this source of uncertainty.

The nature of the error involved in the estimated coefficients, $\hat{a}$, is rather different. If we are able to isolate its effects from those due to the structural disturbances, we can process and measure such effects in a completely different way. In many cases, in fact,

we could even apply full analytical methods or, if simpler in practice, some simulation procedures which are straightforward numerical applications of the analytical methods, therefore called "analytic simulation".

We notice that the error involved in the estimated coefficients is only due to the availability of data for a short number of time periods. If an infinitely large sample could be available, the consistency of the estimation method would produce a coefficient vector without errors. For finite sample lengths, we now observe that, conditional on the exact knowledge of predetermined variables, the simulation process is nothing but a function of the estimated coefficients (2.4) and that this function, even if usually unknown, is continuously differentiable. We can, therefore, apply a well known theorem on the limiting distribution of functions of sample statistics, which states as follows (see Rao, 1973, p.388).

Theorem. Let $\hat{a}$ be an s-dimensional statistic, such that the asymptotic distribution of $\sqrt{T}(\hat{a}-a)$ is s-variate normal with mean zero and covariance matrix Ψ. Let g be an m-dimensional vector of functions of the s variables and each g be totally differentiable. Then the asymptotic distribution of $\sqrt{T}[g(\hat{a})-g(a)]$ is m-variate normal with zero means and covariance matrix $G\Psi G'$, where $G=\partial g/\partial a'$.

A practically equivalent statement, which is suitable for our case (see again Rao, 1973, p.388), is that, if $\hat{a}$ is distributed approximately as s-variate normal with mean a and covariance matrix $\hat{\Psi}/T$, then $g(\hat{a})$ is distributed approximately as m-variate normal with mean g(a) and covariance matrix $G\hat{\Psi}G'/T$. The accuracy of the approximation, of course, increases as the sample size increases.

We can resort to the above theorem to measure the degree of uncertainty due to errors in estimated coefficients. When we evaluate the effects of coefficients errors on forecasts, then $g(\hat{a})$ is the simulation process (2.4), and the matrix of its derivatives, G, can be computed analytically even if the functional operators g are unknown. For the implicit functions theorem, in fact, we have the equality

$$(3.1) \qquad \partial g/\partial a' = -(\partial f/\partial y')^{-1}\,\partial f/\partial a'$$

which involves only the functional operators of the structural form of the model. However, we have preferred to calculate derivatives only numerically, as ratios of finite differences. This method (analytic simulation) seems to be computationally simpler and has proved to be sufficiently accurate in most cases.

4. FORECASTS AND FORECAST ERRORS

Let h be a time period not belonging to the sample estimation period 1,2,...,T, and let the model be used to forecast at times h+1, h+2, ..., h+r. Given the values of the endogenous variables at time h, y_h, and the values of the exogenous variables in the forecast periods, x_{h+1}, x_{h+2}, ..., x_{h+r}, then the values of the endogenous variables in the forecast periods can be obtained recursively as:

$$
\begin{aligned}
(4.1)\qquad y_{h+1} &= g(y_h, x_{h+1}, a, u_{h+1});\\
y_{h+2} &= g(y_{h+1}, x_{h+2}, a, u_{h+2}) = g(g(y_h, x_{h+1}, a, u_{h+1}), x_{h+2}, a, u_{h+2})\\
&= g_2(y_h, x_{h+1}, x_{h+2}, a, u_{h+1}, u_{h+2});\\
&\ \ \vdots\\
y_{h+r} &= g(y_{h+r-1}, x_{h+r}, a, u_{h+r}) = g(g(\cdots), x_{h+r}, a, u_{h+r})\\
&= g_r(y_h, x_{h+1}, \cdots, x_{h+r}, a, u_{h+1}, \cdots, u_{h+r}).
\end{aligned}
$$

Beyond the possible misspecification of the model and the two error sources discussed in the previous section, at least one more source of uncertainty should be considered (see, for example, Klein and Young, 1980, p.81): the unknown values of the exogenous variables in the forecast periods x_{h+1}, x_{h+2}, ..., x_{h+r}. If the model is used for ex-ante forecasting, that is h+1, h+2, ..., h+r are in the future, the exogenous variables must themselves be forecasted, thus involving uncertainty. If ex-post forecasting is performed, some uncertainty still exists about the values of the exogenous variables (and of the initial values of the endogenous variables, y_h) especially if h is a sufficiently recent time period and preliminary data are used. This problem is not considered in this paper, as well as the uncertainty deriving from the possible misspecification of the model; the present analysis and comparison of methods performed are conditional on exact knowledge of the initial values of the endogenous variables in the forecast period, y_h, and on the values of the exogenous variables x_{h+1}, x_{h+2}, ..., x_{h+r}. As discussed in section 2, this implies that the measure of uncertainty is a "minimum" degree of uncertainty associated with forecasts.

Let us now have a look at the usual way in which modelers produce their forecasts. The model builder must choose a starting point for the simulation experiment. Such a point (h) is usually the last time period for which "sure" information is available; in many cases it is h=T, the last point of the sample estimation period. For the purposes of this paper, it is simpler to start from a period h not belonging to 1,2,...,T (for example, h=T+1); when this will not be possible for lack of data, we shall choose h inside the sample period and a slight approximation will occur.

The model builder next introduces values for y_h and x_{h+1}, x_{h+2}, ..., x_{h+r}, assumed exact, sets the random error terms u_{h+1}, u_{h+2}, ..., u_{h+r} to their expected value (zero) and solves simultaneously the dynamic system (4.1) at time h+1, h+2, ..., h+r. Forecasts are so obtained as

$$
\begin{aligned}
(4.2)\qquad \hat{y}_{h+1} &= g(y_h, x_{h+1}, \hat{a}, 0);\\
\hat{y}_{h+2} &= g(\hat{y}_{h+1}, x_{h+2}, \hat{a}, 0) = g(g(y_h, x_{h+1}, \hat{a}, 0), x_{h+2}, \hat{a}, 0)\\
&= g_2(y_h, x_{h+1}, x_{h+2}, \hat{a}, 0, 0);\\
&\ \ \vdots\\
\hat{y}_{h+r} &= g(\hat{y}_{h+r-1}, x_{h+r}, \hat{a}, 0) = g(g(\cdots), x_{h+r}, \hat{a}, 0)\\
&= g_r(y_h, x_{h+1}, \cdots, x_{h+r}, \hat{a}, 0, \cdots, 0).
\end{aligned}
$$

Forecasts $\hat{y}_{h+1}$, $\hat{y}_{h+2}$, ..., $\hat{y}_{h+r}$ differ from the values of the endogenous variables in the forecast period y_{h+1}, y_{h+2}, ..., y_{h+r}, because the estimated $\hat{a}$ is used instead of the unknown coefficients vector a, and due to the existence of the random error terms u_{h+1}, u_{h+2}, ..., u_{h+r}.

In order to assign forecasts an estimate of their degree of uncertainty, it would be useful to get an estimate of the first two moments of the forecast errors $\hat{y}_{h+1} - y_{h+1}$, $\hat{y}_{h+2} - y_{h+2}$, ..., $\hat{y}_{h+r} - y_{h+r}$, or

some confidence intervals of forecasts (briefly, forecast intervals, Intriligator, 1978, p.518). In most practical applications, however, the estimation of the first moment of the forecast errors does not appear particularly interesting, due to the mild nonlinearity of the econometric models used, which causes it to be very small, at least for for static or short dynamic simulation periods (see, for example, Calzolari, 1979; Fair, 1980; Mariano and Brown, 1980).

Much more interesting, in practice, is the information which can be derived from an estimate of the second order moments.

Alternative methods have been proposed in the literature for the evaluation of the second order moments of forecast errors in nonlinear econometric models. Apart from the technical differences and the computational approximations which they involve, these methods are conceptually equivalent in case of single equation models and for nonsimultaneous equation models. This equivalence does not extend to simultaneous equation systems (even linear). Experimental comparisons of results on real world models used for practical purposes aim at investigating to what extent the methods can still be considered as "practically" equivalent. The methods can briefly be described as follows.

- Stochastic simulation in the sample period and in the forecast period, with re-estimation of the structural coefficients.

With this method, a certain number of stochastic simulation runs over the sample period must be performed. In each run, T vectors (one for each period) of pseudo-random additive error terms, with zero mean and covariance matrix equal to the available $\hat{\Sigma}$, are inserted into the stochastic equations of the system, whose simultaneous solution provides pseudo-random values of all the endogenous variables over the sample period.

Each set of pseudo-random values is used, like a new set of data, to re-estimate the model's vector of structural coefficients. Each vector is used for a new stochastic simulation run, this time not in the sample period, but in the forecast period, also introducing new pseudo-random additive error terms into the stochastic equations.

The values of the endogenous variables obtained as solutions of the model in the forecast period are affected either by errors in the estimated coefficients, which have been randomly generated by the stochastic simulation in the sample period and re-estimation, or by the structural error terms, which have been introduced by stochastic simulation in the forecast period. A convenient number of replications of this process allows one to compute either a variance-covariance matrix of forecasts at times h+1, h+2, ..., h+r, or forecast intervals for each endogenous variable, or joint intervals for groups of variables.

This method is often used to analyze the small sample behaviour of estimation methods, when the analytical investigation is difficult or impossible (see, for example, the discussion in Mariano, 1982, p.503). Its use for investigating the forecast errors of nonlinear econometric models is proposed and described in detail in Schink (1971).

- Monte Carlo on coefficients and stochastic simulation in the forecast period.

Instead of obtaining pseudo-random vectors of coefficients by means

of a stochastic simulation and re-estimation process, this method gets them directly by sampling from the distribution of estimated coefficients. The small sample distribution of structural coefficients in a simultaneous equations system is usually very complicated, or even unknown. Therefore, recalling that, under sufficiently wide conditions, the asymptotic distribution is multivariate normal, it is much easier for the experimenter to perform the generation of pseudo-random coefficients from a multinormal distribution (see, for example, Cooper and Fischer, 1974; Fair, 1980)

For this method, the covariance matrix of the structural coefficients is also required. An estimate of this matrix is standard output from system estimation methods, like three stage least squares or full information maximum likelihood; otherwise, if single equation methods are used, this matrix must be built block by block, as in Brundy and Jorgenson (1971,p.215) for limited information instrumental variables, or Theil (1971,p.500) for two stage least squares estimates. Pseudo-random forecasts can then be produced by performing stochastic simulation of the model in the forecast period. For each simulation run, a pseudo-random vector of coefficients must be generated from the multinormal distribution just mentioned, and r independent vectors of pseudo-random error terms (one for each period of forecast) must also be generated. Both error sources are, therefore, taken into account. As for the previous method, a convenient number of replications allows the computation of a variance-covariance matrix of forecasts at time h+1, h+2, ..., h+r, or the construction of forecast intervals.

- Mixed methods: analytical method, analytic simulation on coefficients and stochastic simulation in the forecast period.

These methods extend that proposed by Bianchi and Calzolari (1980), for static simulation of nonlinear models, to dynamic simulation. In turn, the earlier proposal extended the fully analytical methods developed, for linear models, in Goldberger et al. (1961). The case of dynamic simulation of linear models is treated in Schmidt (1974).

The random error terms have been assumed serially independent; therefore, assuming exact knowledge of all the predetermined variables $(y_h, x_{h+1}, x_{h+2}, \ldots, x_{h+r})$, the vector of coefficients, $\hat{a}$, which is obtained from an estimation procedure applied to the data of the sample period, is independent of the random error terms in the forecast period (which is outside the sample estimation period). We can now decompose the vector of forecast errors, in the generic forecast period h+k, as follows:

$$(4.3)\quad \hat{y}_{h+k} - y_{h+k} = g_k(y_h, x_{h+1}, \ldots, x_{h+k}, \hat{a}, 0, \ldots, 0) - g_k(y_h, x_{h+1}, \ldots, x_{h+k}, a, u_{h+1}, \ldots, u_{h+k})$$

$$= [g_k(y_h, x_{h+1}, \ldots, x_{h+k}, \hat{a}, 0, \ldots, 0) - g_k(y_h, x_{h+1}, \ldots, x_{h+k}, a, 0, \ldots, 0)] + [g_k(y_h, x_{h+1}, \ldots, x_{h+k}, a, 0, \ldots, 0) - g_k(y_h, x_{h+1}, \ldots, x_{h+k}, a, u_{h+1}, \ldots, u_{h+k})] .$$

Having assumed exact knowledge of all the predetermined variables involved in our forecast, the two components of the forecast error vector are independent, since the former depends on $\hat{a}$, while the latter depends on $u_{h+1}, \ldots, u_{h+k}$. We can , therefore, calculate the variances or the covariance matrices of the two components separately, and sum them to get the final results. Forecast intervals for endogenous variables can be obtained from the

estimated standard errors, under an approximate hypothesis of normality.

As far as the first component is concerned, we can apply the Theorem of section 3. If we assume that, as T (sample period length) increases, asymptotically

$$(4.4) \qquad \sqrt{T}\,(\hat{a} - a) \sim N(0, \Psi),$$

and define G_{h+k} as the (m x s) matrix of first order partial derivatives of the vector of functions g_k with respect to the elements of a, computed at the point $(y_h, x_{h+1}, \ldots, x_{h+k}, a, 0, \ldots, 0)$, then, asymptotically,

$$(4.5) \qquad \sqrt{T}\,[g_k(y_h, x_{h+1}, \ldots, x_{h+k}, \hat{a}, 0, \ldots, 0) - g_k(y_h, x_{h+1}, \ldots, x_{h+k}, a, 0, \ldots, 0)] \sim N(0, G_{h+k}\Psi G'_{h+k})$$

An estimated covariance matrix of a multinormal distribution, which approximates the small sample distribution of the first component of the forecast errors, is obtained by calculating G_{h+k} at the point $(y_h, x_{h+1}, \ldots, x_{h+k}, \hat{a}, 0, \ldots, 0)$, replacing Ψ with the available estimate $\hat{\Psi}$, and dividing $\hat{G}_{h+k}\hat{\Psi}\hat{G}'_{h+k}$ by the actual length of the sample period, T. This approximation is asymptotically exact if the functions of the vector g_k are continuously differentiable and if the estimated structural coefficients are consistent and asymptotically normally distributed.

As already observed in section 3, the first order derivatives might be computed analytically. The computation would be rather simple in case of one-period forecasts with models linear in the coefficients (even if nonlinear in the variables, see Calzolari, 1981), but becomes rather difficult (even if still possible, in principle) in all the other cases. It is simpler to abandon the fully analytical method, and compute the first order derivatives by means of numerical differentiation, using finite differences between a control solution and disturbed solutions (with increments on the structural coefficients) in the forecast period.

As far as the second component of the forecast errors is concerned, that is the component which is due to the random error terms $u_{h+1}, \ldots, u_{h+k}$, the variances of its elements can be computed as in the previous methods, by means of replicated stochastic simulations of the model in the forecast period, introducing additive random error terms into the stochastic equations, but with fixed coefficients. Forecast intervals can be constructed from standard errors of forecasts.

5. FINITE MOMENTS AND CONVERGENCE OF MONTE CARLO METHODS

The case of linear systems of simultaneous equations well exemplifies the problems which may be encountered when calculating variances of forecasts with Monte Carlo methods. Let the model be

$$(5.1) \qquad Ay_t + Bx_t = u_t,$$

where A and B are matrices of structural coefficients (A is square and nonsingular), some of which are fixed a-priori (zero and one restrictions, for example), while others must be estimated.

Defining $\Pi_1 = -A^{-1}B$ and $v_t = A^{-1}u_t$, the restricted reduced form can be represented as

$$(5.2) \qquad y_t = \Pi_1 x_t + v_t .$$

If $\hat{A}$, $\hat{B}$ and $\hat{\Pi}_1 = -\hat{A}^{-1}\hat{B}$ ($\hat{A}$ nonsingular) are the available estimates of the coefficients, the model can be used to forecast at time h+1, h+2, ..., h+r. For example, at time h+r, it will be

$$(5.3) \qquad \hat{y}_{h+r} = \hat{\Pi}_1 x_{h+r} .$$

The vector of forecast errors at time h+r is

$$(5.4) \qquad \hat{y}_{h+r} - y_{h+r} = (\hat{\Pi}_1 - \Pi_1) x_{h+r} - v_{h+r} .$$

Equation (5.4) shows the strict relation between forecast errors and errors in estimated coefficients of the restricted reduced form. When performing stochastic simulation in the sample period, with re-estimation of structural coefficients, we reproduce experimentally the small sample distribution of either the structural form or the reduced form coefficients. When calculating the sample variances of the random forecasts, we could be sure that these variances converge only if the distribution of $\hat{\Pi}_1$ has finite moments, at least to the second order.

The existence of finite moments of reduced form coefficients is not a general property of simultaneous equations systems, but it strictly depends on the model and on the estimation method which is adopted. Detailed proofs and references can be found in the recent article by Mariano (1982); a simple intuitive explanation is given in McCarthy (1981). Some estimation methods, like full information maximum likelihood, directly estimate $\hat{\Pi}_1$; they make use of any overidentifying restrictions of the structural form, but do not need inversion of the matrix $\hat{A}$. However, for other estimation methods, like two or three stage least squares, the reduced form coefficients are derived from the structural form coefficients through the inversion of the matrix $\hat{A}$, and "the model builder has no control on the determinant" of such a matrix. The determinant of $\hat{A}$ (appearing in the denominator of each element of the inverse) "can take a zero value with positive probability density associated with all points in the neighborhood of zero". The same, of course, may happen when Monte Carlo is applied by adding random errors to the structural coefficients. We see, in this way, that in the general case both Monte Carlo methods are not equivalent to the analytic simulation method; the latter, in fact, calculates the covariance matrix of a distribution which asymptotically approximates the small sample distribution of our reduced form coefficients or forecasts; since this asymptotic distribution is normal, it always has finite moments.

6. LARGE DIFFERENCES EVEN WHEN MONTE CARLO METHODS CONVERGE

In the large set of experiments on one period forecasts performed in Bianchi and Calzolari (1982), an approximate equivalence of results was found in all cases in which Monte Carlo methods converged; in other words, no Monte Carlo experiment ever converged

to values of the second order moments largely different from those produced by analytic simulation. The only cases in which large differences occurred were those cases in which Monte Carlo did not converge anywhere. It seemed, therefore, that the nonconvergence of Monte Carlo, due to some determinant close to zero (or for some other reason, presumably similar, but not so explicit for nonlinear models), could be the only source of large differences in the variances computed with the different methods for models actually used for forecasting purposes. This, however, is not true for some models, when we pass from one period (static) forecasts to forecasts produced by multiperiod dynamic simulations. Then, even if the matrix which must be inverted is the same as in the static case, the dynamic simulation mechanism increases the risk of generating values of the denominator close to zero, as will soon be clear. Let

$$(6.1) \qquad Ay_t + Bx_t + Cy_{t-1} = u_t$$

be the structural form of the model. Defining $\Pi_0 = -A^{-1}C$, the reduced form of the model is

$$(6.2) \qquad y_t = \Pi_0 y_{t-1} + \Pi_1 x_t + v_t .$$

The solution of the model at time h+r, conditional on the value of y at time h, assumed known, is

$$(6.3) \qquad y_{h+r} = \Pi_0^r y_h + \sum_{k=1}^{r} \Pi_0^{r-k}\Pi_1 x_{h+k} + \sum_{k=1}^{r} \Pi_0^{r-k} v_{h+k} .$$

Using the model with estimated coefficients to forecast at times h+1, h+2, ..., h+r, the forecast error at time h+r is

$$(6.4) \qquad \hat{y}_{h+r} - y_{h+r} = (\hat{\Pi}_0^r - \Pi_0^r) y_h + \sum_{k=1}^{r} (\hat{\Pi}_0^{r-k}\hat{\Pi}_1 - \Pi_0^{r-k}\Pi_1) x_{h+k} - \sum_{k=1}^{r} \Pi_0^{r-k} v_{h+k} .$$

The matrix A which must be inverted to compute Π_0 and Π_1 is still the same as in the static case. Therefore, the risk of nonconvergence of Monte Carlo should be the same as in the static case. However, in a long dynamic simulation run the behaviour of the model crucially depends on the powers of Π_0 and, therefore, on the powers of A^{-1}. Values of the determinant of A not too close to zero, but smaller than one, may become sufficiently close to zero as soon as they are raised to a moderate power, as in dynamic simulation over a few periods; the probability density of the determinant raised to a power is higher, near zero, than for the determinant raised to one. In this case, Monte Carlo experiments would still converge after a sufficient number of replications, but the results would crucially depend on the form of the distribution of the determinant near zero, and not just on the first two moments of the determinant itself. Slight differences between the probability densities of the determinant, in the region between zero and one, may cause sufficiently large differences of results in the computation of variances of forecasts after several periods of dynamic simulation.

7. EMPIRICAL RESULTS: STANDARD ERRORS AND FORECAST INTERVALS

When estimating the variance of forecasts in the dynamic simulation of a large nonlinear model, we cannot have a clear analytical insight of the mechanisms which cause similarity or differences in moments obtained with the different methods. It seems, however, reasonable to believe that the mechanisms are similar to those which act in the linear case.

As far as the comparison of the different methods in producing variances of forecasts is concerned, the experimental results obtained in Bianchi and Calzolari (1983) showed cases in which:

- when computing the variances of forecasts, all Monte Carlo methods converge and their results are close to each other and to analytic simulation results;
- some methods do not converge, while other methods converge to values of the variances of forecasts similar to those produced by analytic simulation;
- some or all methods converge, but produce variances of forecasts which are close to those produced by analytic simulation in the first period of forecast, but diverge from each other and from analytic simulation after a few periods.

For completeness sake, some of the results in Bianchi and Calzolari (1983) are reproduced in the first part of each table. We have repeated all the Monte Carlo experiments without computing second order moments, but simply collecting all the simulation results and building for each variable a forecast interval at 90% by dropping 5% of each tail of the distribution. The extreme points of the intervals, for each endogenous variable, are displayed in the second part of each table. Values are given as percentages of the deterministic simulation value (that is the actual forecast). For example, an interval (-4% : +5%) means that 90% of the simulation results fall in the interval between -4% and +5% of the forecasted value. In this way, the outliers which are responsible of the nonexistence of finite moments, or of large differences when finite moments exist, are dropped in most cases of practical interest.

Even if differences still remain, as obviously expected, the results become comparable in most cases.

The extreme points of the forecast intervals related to analytic simulation have been obtained from the standard errors (plus or minus 1.65 times each standard error, using a normal approximation).

Results indicated with I.V. in the tables are related to stochastic simulation in the sample period and in the forecast period, with re-estimation of the structural coefficients; re-estimation has always been performed by means of limited information instrumental variables methods (I.V.), obtained with one iteration of Brundy and Jorgenson's method after a first estimation with ordinary least squares. For discussions on the problem of efficiency of instrumental variables in case of nonlinear models, reference can be made to Amemiya (1977,p.966).

Results indicated with M.C. in the tables have been obtained with Monte Carlo on coefficients and stochastic simulation in the forecast period. Analytic simulation results are indicated with A.S. For one model only (Klein-I) we also present results related to stochastic simulation with re-estimation performed by means of full information maximum likelihood (FIML).

The first model experimented with is the well known model of the U.S. economy 1921-1941, described in Klein (1950), usually referred

Table 1

Klein-I model. Initial estimates: FIML 1921-1941
Dynamic simulation forecasts 1931-1941.
Forecasts and standard errors at 1931.

	Obsrv.	Forec.	A.S.	FIML	I.V.	M.C.
C	50.9	54.5	2.4	2.3	6.5	9.6
I	-3.4	-.557	2.0	2.0	2.4	7.8
W1	34.5	37.2	2.3	2.3	8.5	6.8
Y	50.7	57.2	4.3	4.3	9.1	18.
P	11.4	15.2	2.3	2.2	2.7	10.
K	213.	216.	2.0	2.0	2.4	7.8

Forecasts and standard errors at 1941
(... indicate values greater than 1000).

	Obsrv.	Forec.	A.S.	FIML	I.V.	M.C.
C	69.7	63.5	3.8	3.8	...	...
I	4.9	-1.58	2.6	2.6	...	...
W1	53.3	45.4	3.6	3.6	...	...
Y	85.3	72.7	6.2	6.2	...	...
P	23.5	18.7	3.0	2.9	...	...
K	209.	208.	10.	9.7	...	...

Forecast intervals at 90%. Extreme points of the intervals are given as deviations from the corresponding forecasts

(1931)	A.S.	FIML	I.V.	M.C.
C	(-3.9 : +3.9)	(-3.5 : +3.9)	(-3.1 : +8.1)	(-4.8 : +4.3)
I	(-3.3 : +3.3)	(-3.2 : +3.5)	(-3.0 : +5.6)	(-4.0 : +3.6)
W1	(-3.8 : +3.8)	(-3.5 : +4.0)	(-3.9 : +7.3)	(-4.4 : +3.9)
Y	(-7.1 : +7.1)	(-6.9 : +7.5)	(-6.5 : +13.)	(-8.6 : +7.7)
P	(-3.8 : +3.8)	(-3.5 : +3.9)	(-3.0 : +6.9)	(-5.0 : +4.4)
K	(-3.3 : +3.3)	(-3.1 : +3.6)	(-3.0 : +5.6)	(-4.0 : +3.6)

(1941)	A.S.	FIML	I.V.	M.C.
C	(-6.3 : +6.3)	(-6.0 : +6.8)	(-6.8 : +9.8)	(-6.8 : +8.4)
I	(-4.3 : +4.3)	(-4.0 : +4.7)	(-4.4 : +7.0)	(-4.4 : +5.9)
W1	(-5.9 : +5.9)	(-5.5 : +6.2)	(-6.5 : +8.8)	(-6.1 : +6.8)
Y	(-10. : +10.)	(-9.1 : +10.)	(-11. : +12.)	(-11. : +12.)
P	(-4.9 : +4.9)	(-4.1 : +5.0)	(-4.8 : +7.2)	(-4.8 : +7.2)
K	(-16. : +16.)	(-15. : +16.)	(-19. : +21.)	(-19. : +16.)

to as the "Klein-I model". It consists of 3 stochastic equations plus 3 identities; it includes 4 exogenous variables and 3 lagged endogenous variables. The 6 endogenous variables are: C= consumption; I= net investment; W1= private wage bill; Y= national income; P= profits; K= end-of-year capital stock. Table 1 displays forecasts, standard errors of forecasts and forecast intervals at 90% for the 6 endogenous variables of this model in the first and last year of a dynamic simulation run from 1931 to 1941.

The results displayed in Table 2 are related to the annual model of the Italian economy developed by ISPE (Istituto di Studi per la

Table 2

ISPE model of Italian economy. Dynamic simulation forecasts 1960-1977. Initial estimates: 2SLS with principal components. Standard errors of forecasts

	at 1960 (percentage);			at 1977 (percentage)		
	A.S.	I.V.	M.C.	A.S.	I.V.	M.C.
CPNCF	1.7	1.7	1.8	5.8	6.7	8.2
DXML	3.5	3.2	3.5	6.7	7.6	9.9
IFIT	10.	10.	10.	11.	13.	17.
LI	2.3	2.3	2.3	3.4	4.0	4.7
MT	6.0	5.9	6.0	13.	15.	18.
PCL	1.7	1.6	1.7	10.	12.	15.
VAP	1.9	1.9	2.0	3.7	4.5	5.9
XT	3.5	3.2	3.5	8.8	10.	13.

Forecast intervals at 90%. Extreme points of the intervals are given as percentages of the corresponding forecasts

(1960)	A.S.	I.V.	M.C.
CPNCF	(-3% : +3%)	(-3% : +2%)	(-3. : +3%)
DXML	(-6% : +6%)	(-3% : +5%)	(-5% : +5%)
IFIT	(-16% : +16%)	(-16 : +13%)	(-17% : +14%)
LI	(-4% : +4%)	(-3% : +3%)	(-3% : +4%)
MT	(-10% : +10%)	(-9% : +8%)	(-10% : +8%)
PCL	(-3% : +3%)	(-2% : +2%)	(-3% : +3%)
VAP	(-3% : +3%)	(-3% : +2%)	(-3% : +3%)
XT	(-6% : +6%)	(-3% : +5%)	(-5% : +7%)

(1977)	A.S.	I.V.	M.C.
CPNCF	(-9% : +9%)	(-11% : +10%)	(-10% : +11%)
DXML	(-11% : +11%)	(-11% : +15%)	(-12% : +12%)
IFIT	(-18% : +18%)	(-19% : +17%)	(-20% : +21%)
LI	(-6% : +6%)	(-7% : +5%)	(-5% : +7%)
MT	(-21% : +21%)	(-24% : +28%)	(-19% : +29%)
PCL	(-16% : +16%)	(-17% : +20%)	(-17% : +17%)
VAP	(-6% : +6%)	(-7% : +7%)	(-7% : +8%)
XT	(-14% : +14%)	(-15% : +20%)	(-14% : +19%)

Programmazione Economica, Roma). The model, described in Sartori (1978), consists of 19 stochastic plus 15 definitional equations; there are 75 estimated coefficients. The initial estimates used for the experiments have been obtained by means of two stage least squares with principal components. For lack of data, the forecast period (1960-1977) is not external to the sample estimation period (1955-1976).

The results in the table are related to 8 of the main endogenous variables of the model. The variables are: CPNCF= private consumption; DXML= price deflator for exports; IFIT= private investment; LI= employees in industrial sector; MT= imports of goods and services; PCL= price deflator of private consumption; VAP= gross output of private sector; XT= exports of goods and services. The standard errors are displayed as percentages of the forecast value

of each endogenous variable, as well as the extreme points of the 90% forecast intervals. It is clear from Table 2 that, for this model, most methods remain approximately equivalent both for static simulation (1960) and after a sufficiently long dynamic simulation period; only Monte Carlo on coefficients converges, in dynamic simulation, to values of the second order moments slightly larger than those produced by the other methods, and the forecast intervals are slightly larger as well.

Table 3

IBM model of the United Kingdom. Dynamic simulation forecasts 1976/I-1978/IV. Initial estimates: iterative instrumental variables. Standard errors of forecasts

	at 1976/I (percentage);			at 1978/IV (percentage)		
	A.S.	I.V.	M.C.	A.S.	I.V.	M.C.
BI	3.2	3.2	3.2	8.3	19.	28.
CPI	1.4	1.4	1.5	9.6	11.	66.
EM	0.4	0.4	0.4	1.4	3.4	2.5
GNP	1.3	1.3	1.3	2.7	4.6	10.
GNPC	1.5	1.5	1.7	12.	15.	133.
IM	2.7	2.7	2.7	4.6	6.8	15.
XIP	1.9	2.2	1.9	4.6	18.	16.

Forecast intervals at 90%. Extreme points of the intervals are given as percentages of the corresponding forecasts

(1976/I)	A.S.	I.V.	M.C.
BI	(-5% : +5%)	(-6% : +5%)	(-6% : +5%)
CPI	(-2% : +2%)	(-2% : +2%)	(-2% : +3%)
EM	(-.6% : +.6%)	(-.9% : +.6%)	(-.6% : +.6%)
GNP	(-2% : +2%)	(-2% : +2%)	(-2% : +2%)
GNPC	(-2% : +2%)	(-3% : +3%)	(-3% : +3%)
IM	(-4% : +4%)	(-4% : +4%)	(-4% : +5%)
XIP	(-3% : +3%)	(-6% : +3%)	(-2% : +3%)

Forecast intervals at 75%

(1978/4)	A.S.	I.V.	M.C.
BI	(-9% : +9%)	(-19% : +14%)	(-11% : +15%)
CPI	(-11% : +11%)	(-7% : +20%)	(-13% : +17%)
EM	(-2% : +2%)	(-4% : +2%)	(-2% : +2%)
GNP	(-3% : +3%)	(-6% : +2%)	(-3% : +4%)
GNPC	(-14% : +14%)	(-12% : +21%)	(-16% : +19%)
IM	(-5% : +5%)	(-9% : +3%)	(-4% : +6%)
XIP	(-5% : +5%)	(-11% : +5%)	(-6% : +9%)

The results displayed in Table 3 are related to the macroeconomic model of the United Kingdom developed by the IBM Economics Department. It is a quarterly model, with 120 equations, 21 of which are stochastic, and with 32 exogenous variables; there are 68 structural estimated coefficients. The start of the sample period varies from 1956/II to 1969/I but always ends at 1975/IV. The

initial estimates used for the experiments have been obtained by means of iterative instrumental variables (obtained by iterating I.V. method till convergence is reached).

The results in the table are related to 7 of the main endogenous variables of the model; standard errors are displayed as percentages of the value forecast for each variable. The forecast period is external to the sample estimation period: from 1976/I to 1978/IV. The variables are: BI= private fixed investment; CPI= consumer price index; EM= private sector employees; GNP= gross national product; GNPC= GNP in current prices; IM= imports; XIP= index of industrial production.

It is clear from Table 3 that, while for static simulation (1976/I) all methods are approximately equivalent, after some periods of dynamic simulation Monte Carlo on coefficients method and stochastic simulation and re-estimation with instrumental variables converge to values of the second order moments which are rather different from each other and from those produced by the other method.

Unfortunately (and this is the only case we encountered in our experiments) the number of outliers, in the Monte Carlo on coefficients method at the end of the dynamic simulation, is so large that the forecast interval at 90% is not enough to leave them out. Comparable results were obtained only with narrower forecast intervals (for example, at 75%; in such a case, the extreme points of the analytic simulation forecast intervals have been obtained as plus and minus 1.15 times the standard errors).

The results displayed in Table 4 are related to the real sector sub-model of the Forecasting System No.10 of the German economy, developed by the University of Bonn. For a description of the model, reference should be made to Krelle (1976). The sub-model used for these experiments consists of 136 equations, 59 of which are stochastic; it includes 39 exogenous variables and 163 estimated coefficients (data are annual). For most of the equations, the estimation period is 1960-1977.

The initial estimates used for the experiments have been obtained by instrumental variables (two iterations after an initial ordinary least squares estimate). Since for this model the number of structural equations is considerably larger than the length of the sample period, the estimated covariance matrix of structural coefficients is not positive definite, so that the triangular decomposition of the matrix, required by the Monte Carlo on coefficients method, cannot be performed. Therefore, experiments with Monte Carlo on coefficients have been performed assuming the matrix to be block diagonal, as in Cooper and Fischer (1974), Fair (1980); of course, this introduces an additional approximation.

Analytic simulation on coefficients has been performed both with the complete covariance matrix and with the block diagonal matrix; as the results did not change significantly, at least for most variables, only results obtained with the complete matrix are displayed in Table 4.

The results are displayed for 7 of the main variables: P'C= price index of consumption; WR'P= wage rate private; C'PR= private consumption; YDP'P= gross domestic product private; M'GSNO= imports of goods and services; T= total tax payment; FW= foreign workers.

In the first forecast period (static simulation), the results produced by all methods are quite similar to each others. On the contrary, in the last forecast period (dynamic simulation from 1970 to 1977; also in this case, the forecast period is not external to

Table 4

Bonn Forecasting System No.10 (real sector).
Dynamic simulation forecasts 1970-1977.
Initial estimates: I.V. Standard errors of forecasts

	at 1970 (percentage);			at 1977 (percentage)		
	A.S.	I.V.	M.C.	A.S.	I.V.	M.C.
P'C	1.1	1.1	1.1	3.5	3.4	3.4
WR'P	2.6	2.6	2.6	8.7	9.1	9.3
C'PR	1.5	1.5	1.5	3.1	5.8	3.4
YDP'P	1.4	1.5	1.5	2.2	8.3	2.4
M'GSNO	3.2	3.2	3.2	5.3	19.	5.6
T	3.1	3.1	3.1	6.8	9.4	6.9
FW	11.	11.	11.	25.	24.	28.

Forecast intervals at 90%. Extreme points of the intervals are given as percentages of the corresponding forecasts

(1970)	A.S.	I.V.	M.C.
P'C	(-2% : +2%)	(-2% : +2%)	(-2% : +2%)
WR'P	(-4% : +4%)	(-5% : +4%)	(-4% : +5%)
C'PR	(-2% : +2%)	(-2% : +2%)	(-2% : +3%)
YDP'P	(-2% : +2%)	(-2% : +2%)	(-3% : +2%)
M'GSNO	(-5% : +5%)	(-4% : +4%)	(-4% : +5%)
T	(-5% : +5%)	(-4% : +5%)	(-5% : +5%)
FW	(-18% : +18%)	(-16% : +14%)	(-19% : +19%)

(1977)	A.S.	I.V.	M.C.
P'C	(-6% : +6%)	(-9% : +3%)	(-6% : +5%)
WR'P	(-14% : +14%)	(-19% : +12%)	(-15% : +14%)
C'PR	(-5% : +5%)	(-6% : +4%)	(-5% : +5%)
YDP'P	(-4% : +4%)	(-2% : +7%)	(-4% : +4%)
M'GSNO	(-9% : +9%)	(-9% : +10%)	(-7% : +9%)
T	(-11% : +11%)	(-13% : +8%)	(-11% : +10%)
FW	(-41% : +41%)	(-35% : +45%)	(-42% : +45%)

the sample estimation period) two methods produce quite similar values of the second order moments, while moments produced by stochastic simulation and instrumental variables re-estimation diverge substantially from the others. The 90% forecast intervals, however, are quite similar also for this case.

REFERENCES

Amemiya,T. (1977), "The Maximum Likelihood and the Nonlinear Three-Stage Least Squares in the General Nonlinear Simultaneous Equation Model", Econometrica 45, 955-968.

Bianchi,C. and G.Calzolari (1980), "The One-Period Forecast Errors in Nonlinear Econometric Models", International Economic Review 21, 201-208.

Bianchi,C. and G.Calzolari (1982), "Evaluating Forecast Uncertainty Due to Errors in Estimated Coefficients: Empirical Comparison of Alternative Methods", in Evaluating the Reliability of Macro-Economic Models, ed. by G.C.Chow and P.Corsi. New York: John Wiley, 251-277.

Bianchi,C. and G.Calzolari (1983), "Standard Errors of Forecasts in Dynamic Simulation of Nonlinear Econometric Models: Some Empirical Results", in Time Series Analysis - Theory and Practice 3, ed. by O.D.Anderson. Amsterdam: North Holland (forthcoming).

Brundy,J.M. and D.W.Jorgenson (1971), "Efficient Estimation of Simultaneous Equations by Instrumental Variables", The Review of Economics and Statistics 53, 207-224.

Calzolari,G. (1979), "Antithetic Variates to Estimate the Simulation Bias in Non-Linear Models", Economics Letters 4, 323-328.

Calzolari,G. (1981), "A Note on the Variance of Ex-Post Forecasts in Econometric Models", Econometrica 49, 1593-1595.

Cooper,J.P. and S.Fischer (1974), "Monetary and Fiscal Policy in the Fully Stochastic St.Louis Econometric Model", Journal of Money, Credit, and Banking 6, 1-22.

Fair,R.C. (1980), "Estimating the Expected Predictive Accuracy of Econometric Models", International Economic Review 21, 355-378.

Gallant,A.R. (1977), "Three-Stage Least-Squares Estimation for a System of Simultaneous, Nonlinear, Implicit Equations", Journal of Econometrics 5, 71-88.

Goldberger,A.S., A.L.Nagar and H.S.Odeh (1961), "The Covariance Matrices of Reduced-Form Coefficients and of Forecasts for a Structural Econometric Model", Econometrica 29, 556-573.

Intriligator,M.D. (1978), Econometric Models, Techniques, and Applications. Amsterdam: North Holland.

Klein,L.R. (1950), Economic Fluctuations in the United States, 1921-1941. New York: John wiley, Cowles Commission Monograph 11.

Klein,L.R. and R.M.Young (1980), An Introduction to Econometric Forecasting and Forecasting Models. Lexington: D.C.Heath and Company.

Krelle,W. (1976), "Das Bonner Prognosemodell 10", Teil I, "Reales Modell". Universitaet Bonn, Institut fuer Gesellschafts- und Wirtschaftswissenschaften, Wirtschaftstheoretische Abteilung, discussion paper.

Mariano,R.S. (1982), "Analytical Small-Sample Distribution Theory in Econometrics: The Simultaneous-Equations Case", International Economic Review 23, 503-533.

Mariano,R.S. and B.W.Brown (1980), "Asymptotic Behavior of Predictors in a Nonlinear Simultaneous System". University of Pennsylvania, The Wharton School, Department of Economics, discussion paper No.410.

McCarthy,M.D. (1972), "Some Notes on the Generation of Pseudo-Structural Errors for Use in Stochastic Simulation Studies", in Econometric Models of Cyclical Behavior, ed. by B.G.Hickman. New York: NBER, Studies in Income and Wealth No.36, 185-191.

McCarthy,M.D. (1981), "Notes on the Existence of Moments of Restricted Reduced Form Estimates". University of Pennsylvania, discussion paper.

Rao,C.R. (1973), Linear Statistical Inference and its Applications, New York: John Wiley.

Sartori,F. (1978), "Caratteristiche e Struttura del Modello", in Un Modello Econometrico dell'Economia Italiana; Caratteristiche e Impiego. Roma: Ispequaderni 1, 9-36.

Schink,G.R. (1971), "Small Sample Estimates of the Variance Covariance Matrix of Forecast Error for Large Econometric Models: The Stochastic Simulation Technique". University of Pennsylvania, Ph.D. dissertation.

Schmidt,P. (1974), "The Asymptotic Distribution of Forecasts in the Dynamic Simulation of an Econometric Model", Econometrica 42, 303-309.

Schmidt,P. (1976), Econometrics. New York: Marcel Dekker.

Theil,H. (1971), Principles of Econometrics. New York: John Wiley.