Evaluating forecast uncertainty due to errors in estimated coefficients: empirical comparison of alternative methods

Bianchi, Carlo and Calzolari, Giorgio

Centro Scientifico IBM, Pisa

1982

Evaluating Forecast Uncertainty due to Errors in Estimated Coefficients: Empirical Comparison of Alternative Methods

C. BIANCHI[1]
Centro Scientifico IBM, Pisa
G. CALZOLARI
Centro Scientifico IBM, Pisa

13.1 INTRODUCTION

Forecasts produced by structural econometric models are subject to several sources of error; some of these are, for example, the random disturbance term of each stochastic equation, errors in estimated coefficients, errors in forecasts of exogenous variables, errors in preliminary data (before the final revisions become available), and possible misspecification of the model. Several analytical, empirical, numerical, and Monte Carlo methods have been proposed in the econometric literature for estimating the contribution to forecast errors of some or all of these error sources.[2]

This paper will be concerned with the contribution to forecast errors of errors in the estimated structural coefficients. Its main purpose is to perform, on several 'real world' models, an empirical comparison of alternative techniques available in the literature for this purpose.

For all models, the comparison will be performed in the case of one-period forecasts, outside the sample used for estimating the structural coefficients, under the assumptions of perfect knowledge of the predetermined variables and correct specification of the model structure; in other words, forecasts conditional on the predetermined variables and the model's specification. What remains is a forecast error that depends on the structural disturbances and on the coefficient estimates only. These two components, under general assumptions, are independent when

forecasting outside the sample period, so that their statistical properties can be studied separately.

The statistical techniques available for this purpose include explicit analytical formulae, but these are applicable to linear models only; furthermore, in the cases in which they are applicable, several computational difficulties are involved, so that, in empirical applications for real world nonlinear models, one must resort to simulation techniques.

Stochastic simulation is usually used for estimating the component due to the error terms, obtaining the so-called 'reduced form' variance.

Stochastic and analytic simulation techniques are used for the estimation of the component due to the coefficient estimates. In this paper, the empirical results obtained with three alternative methods will be discussed. These three methods, briefly described in Section 13.4, are discussed in detail in Schink (1971), Fair (1980), and Bianchi and Calzolari (1980).

The methods of Schink (1971) and of Fair (1980) perform a number of stochastic simulation runs, in each run starting from a set of different structural coefficients, and they differ in the way in which the 'new' coefficients are obtained. Hence, even if they could be considered as belonging to the same area (i.e. stochastic simulation with respect to the coefficient estimates), they are substantially different. In order to avoid misunderstanding in the present paper, they will be referred to in the following way:

— 'Stochastic simulation and re-estimation', for the method of Schink (1971), where the analysis of the forecast error is performed starting from the covariance matrix of the structural form disturbances; each run of stochastic simulation with respect to the error terms supplies a new pseudo-sample for the endogenous variables and the new set of structural coefficients is obtained by re-estimating the model.
— 'Monte Carlo on coefficients' for the method of Fair (1980), where the new set of structural coefficients is generated by random sampling from the covariance matrix of the structural coefficients.
— The last method, by Bianchi and Calzolari (1980), will be referred to as 'analytic simulation on coefficients'.

13.2 NOTATION AND ASSUMPTIONS

Even though the purpose of this paper is to compare algorithms applicable to nonlinear models, it is useful to give some consideration to the case of a linear simultaneous-equations system. For this reason, notation will be introduced in this section both for linear and nonlinear models.

A general structural econometric model, linear or nonlinear in the variables as well as in the coefficients (including coefficient restrictions), can be represented as

$$\mathbf{f}(\mathbf{y}_t, \mathbf{x}_t, \mathbf{a}) = \mathbf{u}_t \qquad t = 1, 2, \ldots, T \tag{2.1}$$

where:

$\mathbf{f}' = (f_1, f_2, \ldots, f_m)$ is a vector of functional operators, continuously differentiable with respect to the elements of $\mathbf{y}$, $\mathbf{x}$, and $\mathbf{a}$; $\mathbf{y}_t' = (y_{1t}, y_{2t}, \ldots, y_{mt})$ and $\mathbf{x}_t' = (x_{1t}, x_{2t}, \ldots, x_{nt})$ are the vectors of endogenous and predetermined variables at time $t = 1, 2, \ldots, T$; $\mathbf{a}' = (a_1, a_2, \ldots, a_s)$ 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); $\mathbf{u}_t' = (u_{1t}, u_{2t}, \ldots, u_{mt})$ is the vector of structural stochastic disturbances at time t, having zero mean and independently and identically distributed over time, with finite covariance matrix and independent of all the predetermined variables.

If the model is linear, its structural form can be represented as:

$$\boldsymbol{A}\mathbf{y}_t + \boldsymbol{B}\mathbf{x}_t = \mathbf{u}_t \qquad t = 1, 2, \ldots, T \tag{2.2}$$

In this case $\boldsymbol{A}$ and $\boldsymbol{B}$ are, respectively, $(m \times m)$ and $(m \times n)$ matrices of structural coefficients, including coefficients known *a priori*, such as zeros, ones, etc; the vector $\mathbf{a}$, therefore, will contain a subset of the elements of $\boldsymbol{A}$ and $\boldsymbol{B}$.

It is worth observing that the hypothesis on the structural disturbances, which are supposed to be independently distributed over time (serial independence), is indispensable for the derivation of the results in Section 13.3; it would be restrictive if the analysis were confined to linear models, but is less restrictive in the more general context of nonlinear models. In fact if, for example, the structural disturbances of the ith equation are assumed to be first-order serial correlated, i.e.

$$y_{it} + a_k x_{jt} = v_t \qquad v_t = \rho v_{t-1} + u_t \tag{2.3}$$

then the equation can be simply rewritten in the form

$$y_{it} - \rho y_{it-1} + a_k x_{jt} - \rho a_k x_{jt-1} = u_t \tag{2.4}$$

so that the equation becomes nonlinear in the coefficients, but no longer violates the above assumptions. In other words, serial correlation of any order is allowed, provided that the structural model is transformed leaving correlation-free disturbances.

The distribution function of the disturbances $\mathbf{u}_t$ must be specified in all the experiments that involve stochastic simulation. In all the experiments hereafter described, it will be assumed multivariate normal:

$$\mathbf{u}_t \sim N(0, \Sigma). \tag{2.5}$$

In the experiments which involve Monte Carlo or analytic simulation on coefficients, an estimate of the covariance matrix of the structural coefficients is also required. Such an estimate is a standard by-product of system estimation methods, like three-stage least squares (3SLS), full information maximum likelihood (FIML), or full information instrumental variables efficient (FIVE). If limited information methods are used, the derivation of an estimate of the full

covariance matrix of the structural coefficients generally requires some additional computations beyond those usually involved in the estimation of the structural coefficients. Since most of our experiments have been performed using limited information instrumental variables efficient estimates (LIVE) or two-stage least squares (with principal components), reference is made to Brundy and Jorgenson (1971, p. 215) and to Theil (1971, p. 500) for the construction of the coefficients' covariance matrix.

However, it is more appropriate to call such a matrix an estimated asymptotic covariance matrix. In fact, under general assumptions, the classical estimation methods for systems of simultaneous equations first of all provide an estimate of the structural coefficients, $\hat{\mathbf{a}}$, such that $\sqrt{T}(\hat{\mathbf{a}} - \mathbf{a})$ has asymptotically a multinormal distribution with zero means and finite covariance matrix. They then provide a consistent estimate of this covariance matrix divided by the actual sample period length. This point is well known and described in econometric textbooks (see, for example, Christ, 1966, p. 379; Schmidt, 1976, p. 254; or Theil, 1971, p. 378 for a detailed discussion). Nevertheless, it has been noted because the difference between two of the three methods considered later in this paper lies mainly in the different treatment given to this covariance matrix, or, more precisely, in a different choice of the stage where the asymptotic estimate is interpreted as an approximation of the small sample estimate. This point will be clarified below.

The covariance matrix of structural coefficients will be denoted by $\boldsymbol{\Psi}$; more exactly we suppose that, asymptotically,

$$\sqrt{T}(\hat{\mathbf{a}} - \mathbf{a}) \sim N(0, \boldsymbol{\Psi}) \tag{2.6}$$

and that, if T is the actual sample period length, the available estimated matrix is $\hat{\boldsymbol{\Psi}}/T$, where $\hat{\boldsymbol{\Psi}}$ is a consistent estimate of $\boldsymbol{\Psi}$.

There are difficulties in proving that equation (2.6) holds when the nonlinear model includes lagged endogenous variables among the predetermined variables. In some sense, however, this work starts after structural estimation problems have been solved, for example by assuming as reasonable the considerations in Gallant (1977, pp. 73–74) and Hatanaka (1978, fn. 8).

13.3 DECOMPOSITION OF THE FORECAST ERRORS

It is usually assumed that a simultaneous-equation system like (2.1) uniquely defines the values of the elements of $\mathbf{y}_t$ once values for the coefficients, the predetermined variables, and the disturbance terms are given (at least in some range); in the case of the linear model (2.2) this is equivalent to assuming nonsingularity of the matrix $\boldsymbol{A}$. This means that the structural form equations (2.1) implicitly define a system of reduced form equations

$$\mathbf{y}_t = \mathbf{y}(\mathbf{x}_t, \mathbf{a}, \mathbf{u}_t) \tag{3.1}$$

where the vector of functional operators $\mathbf{y}$ is generally unknown in the case of nonlinear models.

If the model is linear, equation (3.1) simply becomes

$$\mathbf{y}_t = \boldsymbol{\Pi}\mathbf{x}_t + \mathbf{v}_t \qquad \boldsymbol{\Pi} = -\boldsymbol{A}^{-1}\boldsymbol{B} \qquad \mathbf{v}_t = \boldsymbol{A}^{-1}\mathbf{u}_t. \tag{3.2}$$

Let $\hat{\mathbf{a}}$ be an estimate of the vector $\mathbf{a}$, as described in Section 13.2, obtained by using the data for $\mathbf{y}_t$ and $\mathbf{x}_t$ in the sample period $t = 1, 2, \ldots, T$, and let h be the forecast period, not belonging to $1, 2, \ldots, T$. Under the assumption of serial independence, the disturbances at time h, $\mathbf{u}_h$, are independent of the disturbance terms in the sample period and, therefore, $\mathbf{u}_h$ and $\hat{\mathbf{a}}$ are two independent random variables.

The usual forecast supplied by the model is obtained by inserting, in the structural form equations (2.1), the estimated vector $\hat{\mathbf{a}}$ and the values of the predetermined variables $\mathbf{x}_h$ (supposed exact, for the purposes of this paper, as already observed), dropping the disturbance term, and solving the resulting system

$$\mathbf{f}(\hat{\mathbf{y}}_h, \mathbf{x}_h, \hat{\mathbf{a}}) = \mathbf{0} \tag{3.3}$$

by means of some numerical method. In terms of the (unknown) reduced form, this means that the vector of forecasts at time h can be represented as

$$\hat{\mathbf{y}}_h = \mathbf{y}(\mathbf{x}_h, \hat{\mathbf{a}}, \mathbf{0}). \tag{3.4}$$

The vector of the 'true' values of endogenous variables in the forecast period can be represented, using the reduced form notation, as

$$\mathbf{y}_h = \mathbf{y}(\mathbf{x}_h, \mathbf{a}, \mathbf{u}_h). \tag{3.5}$$

The vector of forecast errors is the difference between $\hat{\mathbf{y}}_h$ and $\mathbf{y}_h$. It is now convenient to introduce an auxiliary vector, $\bar{\mathbf{y}}_h$, defined as the vector of forecasts that would be produced by the model if the structural coefficients were known with certainty; in other words $\bar{\mathbf{y}}_h$ is the solution of the model free of errors at time h:

$$\mathbf{f}(\bar{\mathbf{y}}_h, \mathbf{x}_h, \mathbf{a}) = \mathbf{0} \tag{3.6}$$

that is, with reduced form notation,

$$\bar{\mathbf{y}}_h = \mathbf{y}(\mathbf{x}_h, \mathbf{a}, \mathbf{0}). \tag{3.7}$$

Returning to the vector of forecast errors, we now have

$$\begin{aligned} \hat{\mathbf{y}}_h - \mathbf{y}_h &= [\hat{\mathbf{y}}_h - \bar{\mathbf{y}}_h] + [\bar{\mathbf{y}}_h - \mathbf{y}_h] \\ &= [\mathbf{y}(\mathbf{x}_h, \hat{\mathbf{a}}, \mathbf{0}) - \mathbf{y}(\mathbf{x}_h, \mathbf{a}, \mathbf{0})] + [\mathbf{y}(\mathbf{x}_h, \mathbf{a}, \mathbf{0}) - \mathbf{y}(\mathbf{x}_h, \mathbf{a}, \mathbf{u}_h)]. \end{aligned} \tag{3.8}$$

In the case of a linear model, equation (3.8) assumes the well-known form

$$\begin{aligned} \hat{\mathbf{y}}_h - \mathbf{y}_h &= [\hat{\boldsymbol{\Pi}}\mathbf{x}_h - \boldsymbol{\Pi}\mathbf{x}_h] + [\boldsymbol{\Pi}\mathbf{x}_h - (\boldsymbol{\Pi}\mathbf{x}_h + \mathbf{v}_h)] \\ &= [\hat{\boldsymbol{\Pi}} - \boldsymbol{\Pi}]\mathbf{x}_h - \boldsymbol{A}^{-1}\mathbf{u}_h. \end{aligned} \tag{3.9}$$

In both cases, the vector of forecast errors is the sum of two random vectors: the first is a function of several variables, among which only the vector of estimated coefficients, $\hat{\mathbf{a}}$, is random; the second is also a function of several variables, among which only the vector of structural disturbances, $\mathbf{u}_h$, is random. Since, by assumption, $\hat{\mathbf{a}}$ and $\mathbf{u}_h$ are independent, so also are the two components of the vector of forecast errors.

Therefore the two components can be separately analysed and, in particular, an estimate of the variances of the forecast errors can be obtained by summing the estimated variances of the two components.

What we have stated above is not exactly true if lagged endogenous variables are present among the predetermined variables; in this case, in fact, the two terms of the sum are both functions of the (random) lagged endogenous variables. The above considerations, however, still hold 'conditional' on a given value of the lagged endogenous variables (for example, it could be the historical value, in the case of one-period forecasts).

There is substantial agreement about the methods of analysing the component $(\bar{\mathbf{y}}_h - \mathbf{y}_h)$, which is a function of the random structural disturbances; in most papers that deal with this problem, stochastic simulation is proposed as the basic computational method. By means of replicated solutions of the model, each time introducing a vector of pseudo-random disturbances in place of $\mathbf{u}_h$, it is possible to compute approximate values of the conditional means and variances of the elements of $(\bar{\mathbf{y}}_h - \mathbf{y}_h)$. The approximation improves, usually, as the number of replications increases; if finite moments of the first two orders exist, a very high number of replications would lead, in practice, to the exact values of means and variances, if the parameters of the model (the vector $\mathbf{a}$ and the covariance matrix of the structural disturbances) were known with certainty.

As, however, we are assuming only estimates of these parameters, stochastic simulation will lead to an estimate of the means and variances of the elements of $(\bar{\mathbf{y}}_h - \mathbf{y}_h)$ (a consistent estimate, in particular, if the available estimate of the structural form parameters is consistent).[3]

Of course, if the model is linear, the mean of this component is zero and the covariance matrix of its elements is

$$A^{-1}\Sigma(A^{-1})' \tag{3.10}$$

if Σ is the covariance matrix of $\mathbf{u}_h$; the estimated covariance matrix is

$$\hat{A}^{-1}\hat{\Sigma}(\hat{A}^{-1})'. \tag{3.11}$$

Since, as already mentioned, this method is widely accepted, it is unnecessary to go into further detail.

The rest of the paper will deal only with the other component, for which the literature presents methods that differ from one another both computationally and conceptually. To be more precise, in the works by Schink (1971) and Fair (1980), numerical results are not presented for this component by itself (as it is, for example, in Bianchi and Calzolari, 1980), but directly for the complete forecast error due to the two error sources together. However, since our assumptions allow us to separate the two components, the analysis of the first component by itself should allow us to compare in detail the empirical results obtained with the different methods.

Finally it must be pointed out that the decomposition of the forecast errors outlined in this section is not confined to the case of forecasts obtained by means of one-step (or static) simulation, but could be extended to the case of conditional forecasts obtained by dynamic simulation (always outside the sample estimation period). Since, however, we are not going to present, in this paper, results related to the forecast errors in dynamic simulation, this problem will not be discussed further. For a detailed analytical treatment of this problem in the case of linear models, reference can be made to Schmidt (1974).

13.4 THREE METHODS FOR ANALYSING THE COMPONENT OF FORECAST ERRORS DUE TO ERRORS IN THE ESTIMATED COEFFICIENTS

Three different methods for analysing the component $(\hat{\mathbf{y}}_h - \bar{\mathbf{y}}_h)$ are briefly described in this section. They will be referred to as:

— Stochastic simulation and re-estimation.

— Monte Carlo on coefficients.

— Analytic simulation on coefficients.

It will be clear that not only are there technical differences in the computational algorithms, but there are some basic conceptual differences among the methods.

Stochastic simulation and re-estimation tries to deal with the 'small sample' distribution of $(\hat{\mathbf{y}}_h - \bar{\mathbf{y}}_h)$ directly.

Monte Carlo on coefficients starts from the estimate of the asymptotic covariance matrix of the structural coefficients, treats this matrix as an approximation of the small-sample covariance matrix of the coefficients, and derives the consequences of this assumption on $(\hat{\mathbf{y}}_h - \bar{\mathbf{y}}_h)$.

Analytic simulation on coefficients also starts from the estimated asymptotic covariance matrix of the structural coefficients and derives the asymptotic covariance matrix of $(\hat{\mathbf{y}}_h - \bar{\mathbf{y}}_h)$; only after this computation is performed is the resulting matrix interpreted as an approximation of the small-sample covariance matrix of $(\hat{\mathbf{y}}_h - \bar{\mathbf{y}}_h)$. In some sense, with respect to Monte Carlo on coefficients, the approximation is performed at a later stage.

From a purely empirical point of view, however, all methods lead to the same information, i.e. an estimated covariance matrix of the given component of forecast errors.

13.4.1 Stochastic simulation and re-estimation

This method can be summarized as follows (see Schink, 1971, for more details). Let $\hat{\Sigma}$ be the available estimate of the covariance matrix of the structural disturbances.

1. T vectors of pseudo-random numbers, $\tilde{\mathbf{u}}_t$, $t = 1, 2, \ldots, T$ (each of which having multinormal distribution, zero means, and covariance matrix equal to the available $\hat{\Sigma}$), are generated. The method of Nagar (1969) can be applied if $\hat{\Sigma}$ is positive definite; otherwise, the method by McCarthy (1972a) can be used.
2. The vectors $\tilde{\mathbf{u}}_t$ are inserted into the model, where the structural coefficients are maintained fixed at their originally estimated values, and the model is solved over all the sample period, obtaining for the endogenous variables the vectors $\tilde{\mathbf{y}}_t$, $t = 1, 2, \ldots, T$.
3. The vectors $\tilde{\mathbf{y}}_t$ are treated as a new set of observations of the endogenous variables and are used to re-estimate the model, thus obtaining a new vector, $\tilde{\mathbf{a}}$, of pseudo-estimated coefficients (or new matrices $\tilde{A}$ and $\tilde{B}$ if the model is linear).
4. The coefficients $\tilde{\mathbf{a}}$ are inserted into the model to produce, via deterministic solution, a vector of pseudo-forecasts at time h, $\tilde{\mathbf{y}}_h$.

The process is repeated from step 1 to 4 and the desired results follow from the computation of the sample variances of the elements of all the $\tilde{\mathbf{y}}_h$ computed in the various replications.

Some complications arise from the treatment of lagged endogenous variables in the simulation phase (in other words simulation can be static or dynamic) and in the re-estimation phase (they can be maintained 'static', i.e. fixed at some given (historical) value, or their simulation value can be chosen). This problem is discussed in Schink (1971, pp. 101–108); in all the experiments here performed the static–static combination has been adopted.

This method is frequently used in the literature to derive small-sample distributions of estimators for simultaneous-equation systems when analytical methods are not available The main theoretical limitation is in the possible nonexistence of finite moments in the small-sample distribution of the structural form or reduced form coefficients (these last directly related to forecasts); this topic is discussed, for example, in Dhrymes (1970, p. 182), McCarthy (1972b), Sargan (1976), and Mariano (1980).

As pointed out in McCarthy (1972b, p. 761), '... it should be noted that the non-existence of moments has some implications for those engaged Monte Carlo studies. Outliers can be expected. Computation of mean squared forecast errors and the mean squared errors of the restricted reduced form coefficient estimates will not converge as the number Monte Carlo runs increases. These computations really will not yield meaningful information. Throwing out the outliers in making these calculations is also of questionable value. What is accomplished by throwing them out?...'.

Whether the nonconvergence of this method is a purely theoretical problem or, on the contrary, is an actual problem for real world models will be clear from the examples in Sections 13.5 and 13.6.

13.4.2 Monte Carlo on coefficients

This method can be summarized as follows. Let $\hat{\Psi}/T$ be the available estimate of the covariance matrix of the structural coefficients $\hat{\mathbf{a}}$.

1. A vector $\tilde{\mathbf{a}}$ of pseudo-random numbers, with mean $\hat{\mathbf{a}}$ and covariance matrix $\hat{\Psi}/T$, is generated.
2. The pseudo-random coefficients vector $\tilde{\mathbf{a}}$ replaces the original estimates $\hat{\mathbf{a}}$ and the model is solved deterministically in the forecast period h, obtaining the vector of pseudo-forecasts $\tilde{\mathbf{y}}_h$.

The process is repeated from step 1 to 2 and the desired results follow from the computation of the sample variances of the elements of all the $\tilde{\mathbf{y}}_h$ computed in the various replications.

A difficulty may arise in the generation of the pseudo-random vectors $\tilde{\mathbf{a}}$. The usual generation methods are, in fact, based on Choleski triangularization of the matrix $\hat{\Psi}/T$ (see Cooper and Fischer, 1974, or Nagar, 1969, for example) and this is possible only if such a matrix is positive definite. Unfortunately, this is not always the case. For example when, in a large-scale model, the length of the time series does not allow the application of system estimation methods, the matrix $\hat{\Psi}/T$ must be built block by block (see, for example, Brundy and Jorgenson, 1971, p. 215, for LIVE estimates) and it is not necessarily of full rank. In this case the generation of the pseudo-random coefficients vectors $\tilde{\mathbf{a}}$ should pass through the generation of shorter vectors with full rank covariance matrix (see, for example, Rao, 1965, pp. 498–501) with some additional computational difficulties.

This problem clearly does not arise if only the diagonal blocks of the $\hat{\Psi}/T$ matrix are taken into account, as in the work of Cooper and Fischer (1974), Haitovsky and Wallace (1972), and Fair (1980). In the experiments that will be described in the following sections, the complete matrix $\hat{\Psi}/T$ will be taken into account whenever possible, otherwise only its diagonal blocks will be used. It must, however, be pointed out that all the experiments performed with the complete $\hat{\Psi}/T$ matrix have also been repeated with the diagonal blocks only, usually obtaining quite similar results (a similar conclusion is in Bianchi *et al.*, 1981).

With respect to stochastic simulation and re-estimation this method seems to be more sensitive to outliers, so that a kind of 'instability' in the convergence of the Monte Carlo process has been encountered more often; examples are given in Section 13.6 and more details on the problem of outliers are given in Section 13.6.5.

13.4.3 Analytic simulation on coefficients

This method, described in detail in Bianchi and Calzolari (1980), is an extension to nonlinear models of the fully analytical method developed for linear models by Goldberger *et al.* (1961).

The method relies on the property, well known in large-sample theory (see, for example, Rao, 1965, p. 322), that asymptotic normality of sample statistics can be maintained through transformations, even nonlinear, provided they are continuous and differentiable.

If we assume that, as T increases, asymptotically

$$\sqrt{T}(\hat{\mathbf{a}} - \mathbf{a}) \sim N(0, \boldsymbol{\Psi}) \tag{4.3.1}$$

(and $\hat{\boldsymbol{\Psi}}$ is a consistent estimate of $\boldsymbol{\Psi}$) then, asymptotically,

$$\sqrt{T}(\hat{\mathbf{y}}_h - \bar{\mathbf{y}}_h) = \sqrt{T}\,[\mathbf{y}(\mathbf{x}_h, \hat{\mathbf{a}}, \mathbf{0}) - \mathbf{y}(\mathbf{x}_h, \mathbf{a}, \mathbf{0})] \sim N(0, \boldsymbol{G}_h \boldsymbol{\Psi} \boldsymbol{G}_h') \tag{4.3.2}$$

where $\boldsymbol{G}_h$ is the $(m \times s)$ matrix of first-order partial derivatives of the elements of $\mathbf{y}$ with respect to the elements of $\mathbf{a}$, computed at the point $(\mathbf{x}_h, \mathbf{a}, \mathbf{0})$.

If the computation is performed at the point $(\mathbf{x}_h, \hat{\mathbf{a}}, \mathbf{0})$ and $\hat{\boldsymbol{\Psi}}$ is used in equation (4.3.2), then $\hat{\boldsymbol{G}}_h \hat{\boldsymbol{\Psi}} \hat{\boldsymbol{G}}_h'$ is a consistent estimate of $\boldsymbol{G}_h \boldsymbol{\Psi} \boldsymbol{G}_h'$; the division by the sample period length, T, leads to the result we are looking for, the estimate of the covariance matrix of a multinormal distribution which approximates the small sample distribution of the random vector $(\hat{\mathbf{y}}_h - \bar{\mathbf{y}}_h)$.

Continuity and differentiability of the elements of the (unknown) vector of reduced form functional operators $\mathbf{y}$ is ensured by the implicit function theorem, which also provides the way of computing the partial derivatives

$$\frac{\partial \mathbf{y}}{\partial \mathbf{a}'} = -\left(\frac{\partial \mathbf{f}}{\partial \mathbf{y}'}\right)^{-1} \left(\frac{\partial \mathbf{f}}{\partial \mathbf{a}'}\right) \tag{4.3.3}$$

where the derivatives of the structural form operators vector $\mathbf{f}$ (known) can be also analytically computed, once the deterministic solution of the model at time h is known.

For medium- or large-scale models it can be more convenient to perform the computation of the above derivatives with numerical methods (finite differences), rather than analytically; this criterion has been followed for all the models used here, with the exception of the Klein-I model, where both analytical and numerical derivation has been performed (of course, with coincident results).

If the model is linear, recalling equation (3.9) and making use of the formula proposed in Nissen (1968), the above method can be made more explicit as follows:

$$\begin{aligned}\sqrt{T}(\hat{\mathbf{y}}_h - \bar{\mathbf{y}}_h) &= \sqrt{T}(\hat{\boldsymbol{\Pi}} - \boldsymbol{\Pi})\mathbf{x}_h = \sqrt{T}\,\mathrm{vec}\,[\boldsymbol{I}(\hat{\boldsymbol{\Pi}} - \boldsymbol{\Pi})\mathbf{x}_h] \\ &= \sqrt{T}\,(\mathbf{x}_h' \otimes \boldsymbol{I})\,\mathrm{vec}(\hat{\boldsymbol{\Pi}} - \boldsymbol{\Pi})\end{aligned} \tag{4.3.4}$$

where $\boldsymbol{I}$ is the $(m \times m)$ unit matrix.

Equation (4.3.4) represents a linear combination of the elements of $(\hat{\boldsymbol{\Pi}} - \boldsymbol{\Pi})$ with fixed coefficients, so that the asymptotic covariance matrix of $\sqrt{T}(\hat{\mathbf{y}}_h - \bar{\mathbf{y}}_h)$ can be computed with no difficulty as soon as the asymptotic covariance matrix of $\sqrt{T}\,\mathrm{vec}(\hat{\boldsymbol{\Pi}} - \boldsymbol{\Pi})$ has been computed, and this can be done with the methods proposed by Goldberger *et al.* (1961) or Dhrymes (1973).

13.5 FIRST SET OF EXAMPLES

This section is concerned with the numerical results obtained from experiments performed on three nonlinear models, plus the linear Klein-I model. Some general considerations hold for all the experiments and must be taken into account for a clear understanding of the tables of results.

1. For all the models, forecasts are related to the first period (year or quarter) outside the sample estimation period.
2. In all the tables, the first numerical column displays the forecast values of some of the main endogenous variables of the model.
3. Analytic simulation on coefficients, not being a Monte Carlo method, leads to results that are not affected by simulation sampling errors. Only minor changes can be expected from the third decimal digit, due to approximations in the numerical computation (for example due to the choice of the finite increments to compute the derivatives). The estimated standard errors of the component of the forecast errors under consideration, computed with this method, are displayed in the second column of each table.
4. Results based on stochastic simulation and re-estimation have been obtained with a number of replications varying from 100 to 1000 and are, therefore, affected by the sampling variability characteristic of Monte Carlo methods. In no case, however, have significant changes in the results been observed after the first 40 or 50 replications.
5. What has just been observed for stochastic simulation and re-estimation means also that in no case has the generation of outliers produced any kind of 'instability' in the computed sample moments (at least up to the second order). The standard errors of the component of the forecast error under consideration, computed with this method, are displayed in column 3 of each table.
6. Also for the Monte Carlo on coefficients method, no problem of 'stability' in the computation of sample variances has been encountered for any of the models used in this section; in other words, no trouble due to outliers has been encountered. The results (standard errors of the component of the forecast error under consideration), all obtained with 1000 replications, are displayed in the fourth column of each table.
7. The last column (fifth) on the right-hand side of each table displays the reduced form standard errors, i.e. the standard errors of that component of the

forecast errors which is not treated in detail in this paper. They are displayed only for the sake of completeness. In Section 13.3, we observe that the standard error of the (complete) forecast error can be obtained, for each variable, by taking the square root of the sum of the square of the values in column 5 and columns 2, 3, or 4 (according to the method adopted).

8. For all the models in this section, the results obtained with the three different methods (columns 2, 3, and 4) are quite similar, but this statement is not of general validity, as will be clear from the examples of Section 13.6.

13.5.1 Klein-I model: LIVE estimates 1921–1941

The model proposed in Klein (1950) consists of three stochastic plus three definitional equations; there are 12 estimated coefficients, 4 for each equation. Estimation has been performed with the limited information instrumental variables efficient method proposed in Brundy and Jorgenson (1971), on the sample period 1921–1941. Forecasts are related to year 1948, using, for the predetermined variables, the values in Goldberger *et al.* (1961) (Table 13.1).

Table 13.1 Klein-I model: LIVE estimates 1921–1941. One-period forecasts at 1948

	(1)	(2)	(3)	(4)	(5)
C	78.1	1.5	1.7	1.7	2.0
I	9.22	0.91	0.91	1.2	1.4
W_1	59.8	1.2	1.4	1.4	1.7
Y	95.5	2.3	2.4	2.8	3.3
P	27.0	1.4	1.4	1.7	1.9
K	207.0	0.91	0.91	1.2	1.4

Glossary

C = Consumption
I = Net investment
W_1 = Private wage bill
Y = National income
P = Profits
K = End-of-year capital stock

13.5.2 ISPE model of Italian economy

The nonlinear model analysed in this section is an annual model of the Italian economy developed by a team led by ISPE (Istituto Studi Programmazione Economica).

The model, originally described in Sartori (1978), consists of 19 stochastic plus 15 definitional equations; there are 75 estimated coefficients. It has been re-estimated for the period 1955–1976 using limited information instrumental variables efficient method (LIVE) (Table 13.2).

The name LIVE has been maintained since the estimation method is exactly the method of Brundy and Jorgenson (1971), but it must be recalled that instrumental variables estimates are generally not efficient when applied to nonlinear models; the problem is discussed in Amemiya (1977) and Hatanaka (1978).

The results in column 2 (analytic simulation) and 4 (Monte Carlo on coefficients) have been obtained with the complete covariance matrix of coefficients.

Table 13.2 ISPE model of Italian economy: LIVE estimates 1955–1976. One-period forecasts at 1977

	(1)	(2)	(3)	(4)	(5)
CPNCF	36735	277	278	281	470
DXML	2.92	0.07	0.07	0.07	0.07
IFIT	6949	206	209	209	268
LI	7735	67	70	66	148
MT	13806	293	288	312	391
PCL	3.02	0.03	0.03	0.03	0.04
VAP	55818	460	449	487	599
XT	16971	519	522	541	603

Glossary

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

13.5.3 IBM model of United Kingdom

The model analysed in this section is the 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 includes 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. Estimation has been performed by means of the limited information instrumental variables efficient method (LIVE) (Table 13.3).

For the experiments with analytic simulation and Monte Carlo on coefficients, the complete covariance matrix of the structural coefficients has been used.

Table 13.3 IBM model of United Kingdom: LIVE estimates. One-period forecasts at 1976 I

	(1)	(2)	(3)	(4)	(5)
BI	2.37	0.05	0.05	0.05	0.07
CPI	108	0.57	0.65	0.59	1.40
EM	22.7	0.03	0.04	0.03	0.09
GNP	26.7	0.14	0.15	0.14	0.33
GNPC	28.9	0.21	0.23	0.22	0.40
IM	7.33	0.05	0.05	0.05	0.20
XIP	102.0	0.87	1.00	0.89	1.80

Glossary

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

13.5.4 Bonn model 10 of Germany (real sector)

The submodel, for the real economy, of the Bonn Forecasting System n. 10, 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. For a detailed description of the model, reference should be made to Krelle (1976) and to Conrad and Kohnert (1979).

The model has been re-estimated by means of limited information instrumental variables efficient method (LIVE) (Table 13.4).

As the estimated covariance matrix of structural coefficients could not be triangularized, experiments with Monte Carlo on coefficients have been performed assuming the matrix to be block diagonal.

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 column 2.

13.5.5 Klein–Goldberger model: 2SLS–8PC estimates

The model on which the experiments have been performed is the revised Klein–Goldberger model described in Klein (1969). It is nonlinear and consists of 16 stochastic and 4 definitional equations, with 54 estimated coefficients; data are

Table 13.4 Bonn forecasting system n. 10 for Germany (real sector): LIVE estimates. One-period forecasts at 1978

	(1)	(2)	(3)	(4)	(5)
BP'BPGS	29.1	3.4	3.4	3.5	6.6
C'PR	463.0	2.3	2.4	2.8	4.3
M'GSNO	347.0	3.6	3.5	3.8	6.8
P'C	1.54	0.004	0.004	0.004	0.01
WR'P	18.6	0.14	0.14	0.13	0.22
U'DIR	1.33	0.24	0.25	0.23	0.39
T	510.0	5.7	5.7	5.0	11
FW	1.70	0.10	0.10	0.10	0.17
YDP'P	734.0	4.3	4.9	5.5	6.1

Glossary

BP'BPGS = Bal. of paym. goods and services
C'PR = Private consumption
M'GSNO = Imports of goods and services
P'C = Price index of consumption
WR'P = Wage rate private
U'DIR = Unemployed persons
T = Total tax payment
FW = Foreign workers
YDP'P = Gross domestic product private

annual and the sample period is divided into two parts: 1929–1941 and 1947–1964 (Table 13.5).

Estimation has been performed by means of 2SLS with 8 principal components, as in Klein (1969); forecasts are related to 1965.

13.6 COUNTEREXAMPLES

This section is concerned with the numerical results obtained on three models, one of which is again Klein-I estimated in two different ways.

The results obtained in Section 13.5 still hold for the analytic simulation method; they hold for stochastic simulation and re-estimation in all cases except one: they do not hold at all for the Monte Carlo on coefficients method. With Monte Carlo on coefficients the sample variances did not converge as the number of replications increased; outliers caused, from time to time, abnormal fluctuations in the experimental results. This effect is demonstrated by inserting, in the tables, more columns of results, indicated by 4/1, 4/2, etc., with the indication of the numbers of replications to which they refer.

The same problem has occurred in one case with stochastic simulation and re-estimation and also in this case the effect of outliers has been demonstrated by inserting columns 3/1, 3/2, etc., with the indication of the number of replications to which they refer.

Table 13.5 Klein–Goldberger revised model 2SLS–8PC estimates. One-period forecasts at 1965

	(1)	(2)	(3)	(4)	(5)
C_d	56.5	1.1	1.2	1.1	2.4
C_n	304	1.5	1.6	1.5	3.7
R	22.9	0.58	0.56	0.57	1.2
H	7.60	0.10	0.10	0.10	0.25
I_m	29.9	0.56	0.55	0.54	1.1
X	534	3.6	3.7	3.6	8.4
$h.100$	102	0.78	0.88	0.81	2.6
W	313	1.9	2.0	1.9	4.6
w	5.51	0.02	0.02	0.02	0.07
r	4.86	0.13	0.12	0.13	0.36
I	46.7	1.1	1.0	1.1	2.4
D	59.1	0.70	0.71	0.70	1.3
r_s	4.63	0.15	0.14	0.16	0.35
P_c	43.7	3.0	3.6	3.1	6.9
N_w	68.0	0.48	0.61	0.50	1.1
Y	375	2.6	2.7	2.6	5.9
p	1.22	0.014	0.018	0.015	0.04
S_c	3.61	2.8	3.4	2.9	6.6
Π	88.2	2.9	3.3	2.9	6.5
Π_r	33.9	0.35	0.45	0.38	0.86

Glossary

When not otherwise specified the variable is expressed in billions of 1954 dollars

C_d = Consumption of durables
C_n = Consumption of nondurables and services
R = Residential construction
H = Stock of inventories
I_m = Imports
X = Gross national product
h = Index of hours worked per week, 1954 = 1.
W = Wages and salaries
w = Annual earnings, thousands of dollars
r = Average yield on corporate bonds, percent
I = Investments in plant and equipment
D = Capital consumption allowances
r_s = Yield on prime commercial paper, percent
P_c = Corporate profits
N_w = Wage and salary workers, millions
Y = Personal disposable income
p = Implicit GNP deflator, 1954 = 1
S_c = Corporate savings
$\Pi - P_c$ = Proprietors' income
Π_r = Rental income and net interest

One of the experiments is based on ordinary least squares estimates. In this case not all the assumptions concerning consistency of the estimates still hold. From a purely technical point of view, however, the three methods can be applied without any change. The statistical properties of the results will be completely unpredictable, but this does not mean that they cannot be of practical use.

13.6.1 Bonn model 5 of Germany

The model used in these experiments is the Bonn University Model 5, described in Krelle (1974) and Quinke (1978). It consists of 126 equations, 50 of which are stochastic, with 20 exogenous variables; the estimation period is 1960–1975 for most equations, but some equations have been estimated on 1960–1977; the first year outside the sample period is, therefore, 1978; data are annual (Table 13.6).

The re-estimation of the model has created some difficulties, giving rise to some convergence problems in the solution of the model inside the sample estimation period. LIVE estimates have been performed, but they led to some unacceptable values of a few coefficients and to unreasonable forecasts at 1978. It was, therefore, decided to maintain the model in its original form, with ordinary least squares estimates.

Monte Carlo on coefficients has been performed using a block-diagonal covariance matrix for the structural coefficients.

Stochastic simulation and re-estimation has been performed replicating OLS estimates.

Analytic simulation on coefficients has been first performed using the block-diagonal covariance matrix for the coefficients, then has been repeated with a complete matrix obtained with the same formula used for instrumental variables, but with the historical values of endogenous variables instead of the instruments. The results slightly changed; those obtained with the full matrix were closer to those obtained with stochastic simulation and re-estimation, so those displayed in column 2 are referred to the complete matrix, whatever the statistical properties of such a matrix may be.

13.6.2 Klein-I model: FIML estimates

The Klein-I model has been estimated by means of full information maximum likelihood (Table 13.7). The estimated coefficients and standard errors can be found, for example, in Hausman (1974).

13.6.3 Klein-I model: LIVE estimates 1921–1939

For this experiment the Klein-I model has been re-estimated by means of instrumental variables on a sample period shortened two years (Table 13.8).

Table 13.6 Model 5 of Bonn University for Germany: OLS estimates. One-period forecasts at 1978

	(1)	(2)	(3)	370 repl. (4/1)	380 repl. (4/2)	1000 repl. (4/3)	(5)
YDP	653.0	16.0	13.0	28.0	162.0	102.0	11.0
Y'PP	277.0	21.0	15.0	22.0	626.0	386.0	14.0
C'PNO	715.0	17.0	11.0	21.0	369.0	228.0	9.5
I'PN	239.0	19.0	15.0	33.0	901.0	556.0	14.0
T	508.0	12.0	12.0	19.0	413.0	255.0	14.0
GS'GD	168.0	1.9	1.7	1.8	1.80	1.80	1.9
HW	45.0	1.0	0.70	1.5	8.60	5.40	0.67
M'GSNO	326.0	7.8	6.4	10.0	171.0	106.0	8.3
X'GSNO	345.0	8.0	8.9	8.0	8.10	8.40	5.3
BP'GS	18.3	8.2	7.2	10.0	170.0	105.0	10.0

Glossary

YDP = Gross domestic product
Y'PP = Profit income private
C'PNO = Private consumption
I'PN = Gross private investment
T = Total tax payments
GS'GD = Gov. demand goods and services
HW = Hours of work
M'GSNO = Imports of goods and services
X'GSNO = Exports of goods and services
BP'GS = Bal. of paym. goods and services

Table 13.7 Klein-I model: FIML estimates 1921–1941. One-period forecasts at 1948

	(1)	(2)	(3)	380 repl. (4/1)	390 repl. (4/2)	500 repl. (4/3)	1000 repl. (4/4)	(5)
C	76.5	1.4	1.3	1.7	9.2	8.2	6.1	2.3
I	6.27	1.0	1.0	1.2	6.8	6.1	4.5	1.9
W_1	58.2	1.4	1.3	1.5	5.8	5.2	3.9	2.2
Y	91.0	2.3	2.2	2.7	16	14	11	4.2
P	24.1	1.2	1.2	1.6	10	9.2	6.8	2.2
K	204	1.0	1.0	1.2	6.8	6.1	4.5	1.9

Table 13.8 Klein-I model: LIVE estimates 1921–1939. One-period forecasts at 1940

	(1)	(2)	(3)	310 repl. (4/1)	320 repl. (4/2)	440 repl. (4/3)	1000 repl. (4/4)	(5)
C	65.6	1.0	1.1	1.3	2.0	2.7	2.3	2.0
I	2.97	0.67	0.64	0.98	1.4	2.6	2.4	1.5
W_1	45.9	0.80	0.84	0.95	1.3	1.9	1.7	1.8
Y	74.3	1.6	1.6	2.2	3.3	5.2	4.6	3.4
P	20.4	1.1	1.0	1.5	2.2	3.5	3.0	1.9
K	204	0.67	0.64	0.98	1.4	2.6	2.4	1.5

13.6.4 Klein–Goldberger model: 2SLS–4PC estimates

For this experiment the Klein–Goldberger model has been estimated by means of 2SLS with 4 principal components, as in Klein (1969) (Table 13.9).

Table 13.9 Klein-Goldberger revised: model 2SLS–4PC estimates. One-period forecasts at 1965

	(1)	(2)	10 repl. (3/1)	50 repl. (3/2)	100 repl. (3/3)	40 repl. (4/1)	200 repl. (4/2)	1000 repl. (4/3)	(5)
C_d	55.3	1.1	0.72	6.4	4.6	1.1	1.4	16.0	2.4
C_n	303.0	1.5	1.5	6.1	4.5	1.4	2.2	15.0	3.6
R	22.4	0.55	0.54	1.5	1.2	0.48	0.58	2.9	1.2
H	7.55	0.10	0.06	0.59	0.42	0.08	0.15	1.4	0.25
I_m	30.4	0.70	0.71	35.0	25.0	0.77	6.5	87.0	1.1
X	530.0	3.5	3.4	54.0	39.0	3.6	11.0	136.0	8.5
$h.100$	102.0	0.83	1.4	29.0	21.0	1.1	4.3	50.0	2.3
W	311.0	2.0	1.9	25.0	18.0	1.9	4.5	67.0	4.8
w	5.52	0.02	0.02	0.82	0.59	0.03	0.21	1.1	0.08
r	4.89	0.14	0.17	0.14	0.13	0.13	0.15	0.14	0.36
I	46.7	1.1	1.2	1.0	1.0	1.1	1.1	1.1	2.4
D	59.1	0.70	0.85	0.84	0.72	0.69	0.72	0.70	1.3
r_s	4.63	0.15	0.19	0.14	0.14	0.18	0.15	0.15	0.35
P_c	42.0	3.5	5.3	14.0	11.0	5.3	6.1	76.0	6.1
N_w	67.9	0.71	1.0	36.0	26.0	1.3	6.5	51.0	1.4
Y	369	2.6	1.9	35.0	25.0	2.6	7.1	68.0	6.1
p	1.22	0.022	0.035	6.2	4.4	0.043	0.52	5.0	0.036
S_c	1.90	3.6	5.3	18.0	14.0	5.7	7.5	75.0	6.0
Π	86.2	3.2	4.8	9.7	7.6	4.4	4.0	75.0	5.8
Π_r	33.8	0.57	0.83	6.4	4.7	0.95	1.9	1.8	1.1

13.6.5 Comments

The Klein-I model, being linear and sufficiently manageable, can be used to try to explain the abnormal fluctuations in the results obtained, in Section 13.6, with Monte Carlo on coefficients. It seems quite reasonable to believe that a similar explanation can be extended to the other models.

The structural form of the model is:

$$
\begin{aligned}
C &= a_1 + a_2 P + a_3 P_{-1} + a_4 (W_1 + W_2) + u_1 \\
I &= a_5 + a_6 P + a_7 P_{-1} + a_8 K_{-1} + u_2 \\
W_1 &= a_9 + a_{10}(Y + T - W_2) + a_{11}(P + W_1 + T)_{-1} + a_{12} t + u_3 \\
Y &= C + I + G - T \\
P &= Y - W_1 - W_2 \\
K &= K_{-1} + I.
\end{aligned}
$$

By substitution into the three stochastic equations, the three identities can be eliminated, obtaining:

$$
\begin{aligned}
&C - a_1 - a_2(C + I - W_1 - W_2 - T + G) - a_3 P_{-1} - a_4 (W_1 + W_2) = u_1 \\
&I - a_5 - a_6(C + I - W_1 - W_2 - T + G) - a_7 P_{-1} - a_8 K_{-1} = u_2 \\
&W_1 - a_9 - a_{10}(C + I - W_2 + G) - a_{11}(P + W_1 + T)_{-1} - a_{12} t = u_3.
\end{aligned}
$$

In the notation of equation (2.2), the matrix A is

$$
A = \begin{bmatrix} 1 - a_2 & -a_2 & a_2 - a_4 \\ -a_6 & 1 - a_6 & a_6 \\ -a_{10} & -a_{10} & 1 \end{bmatrix}.
$$

The inverse of the matrix A, involved in the reduced form equations (3.2) (and therefore, in forecasts), has the denominator of each element equal to the determinant of A, which is

$$
D = 1 - a_2 - a_6 + a_2 . a_{10} - a_4 . a_{10} + a_6 . a_{10}.
$$

Performing Monte Carlo on coefficients, assuming for the coefficients a multinormal distribution with zero mean and covariance matrix equal to that of FIML estimates (as well as FIML estimates of the coefficients), produces first of all a control value of the determinant equal to 1.6, then a distribution of the determinant with mean, approximately, 1.6 (practically equal to the control value) and standard error, approximately, 0.62; values of this determinant between plus and minus 1/5 of its control value occur approximately in 1.7% of cases. If the numerators of the elements of the inverse of A and of B were all constants (and this is clearly not the case, but can be supposed for a while), there would be 1.7% replications that would produce forecasts larger, in absolute value, than 5 times the control forecasts. A smaller percentage would produce forecasts at least 10 times larger, 20 times larger, and so on. It is clear that these outliers produce a kind of instability in the Monte Carlo process, so that there is no warranty about convergence by increasing the number of replications. The standard error and the value 1.7% have been computed using the asymptotic normal approximation to the distribution of the determinant, then they have been confirmed by means of Monte Carlo with 10^5 replications.

The same can be said for the model estimated by LIVE on the period 1921–1939. In this case the control value of the determinant is 0.67 and its approximate standard error is 0.22; approximately 0.8% of the values fall between plus and minus 1/5 of the control value. As above, the standard error and the value 0.8% have been computed by means of a normal approximation to the distribution of the determinant and then have been confirmed by means of Monte Carlo with 10^5 replications.

If, as in Section 13.5.1, LIVE estimates are performed on the period 1921–1941, the control value of the determinant is 0.66; using a normal approximation, the standard error of the determinant would be 0.16 and only 0.04% of the values would be expected to fall between plus and minus 1/5 of the control value (again, Monte Carlo with 10^5 replications confirms the value 0.04%); the probability of encountering outliers is, therefore, so small (or at least so much smaller than in the previous two cases) as to exclude, in practice, problems of instability in the convergence of Monte Carlo.

This exemplifies and, in some sense, explains with numerical evidence the presence or absence of dangerous outliers in the experiments with Monte Carlo on coefficients. The same examples can be used to show why, in the same cases, no outlier has ever produced trouble in the experiments with stochastic simulation and re-estimation. For the Klein-I model, estimated by LIVE on the sample period 1921–1939, 10^4 replications of stochastic simulation and re-estimation have been performed, each time computing the determinant of the matrix A. The results, as far as moments of the first two orders are concerned, are almost the same as those obtained with Monte Carlo on coefficients or with the asymptotic normal approximation; however, the distribution of the determinant was strongly skewed (Figure 13.1), with the left tail much shorter than the right tail, so that no value of the determinant was ever less than 1/5 of the control value.

We can reasonably presume that something similar must happen for the other estimates of the Klein-I model and for Bonn's model 5, even if, since the latter is nonlinear, the transformations from the structural form to the reduced form cannot be followed explicitly; but it clearly does not happen for the Klein–Goldberger model estimated by means of 2SLS with 4 principal components. Among all the experiments described in this paper, only this case demonstrates evidence that the nonconvergence risk involved in stochastic simulation and re-estimation method is actual and not purely theoretical; however, the entire set of experiments suggests that such a risk seems to be smaller than in the case of Monte Carlo on coefficients.

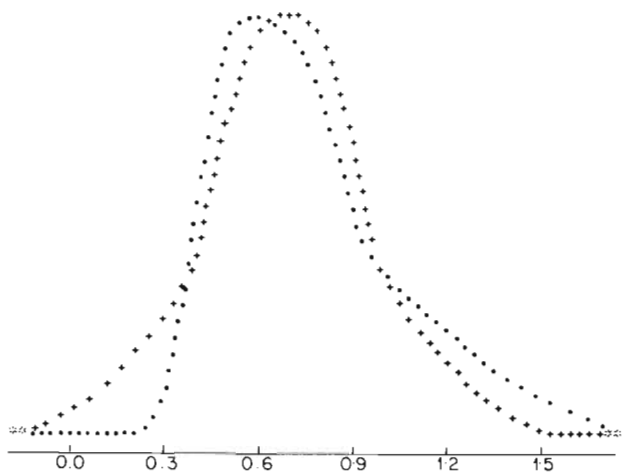

Figure 13.1 Klein-I model: LIVE estimates 1921–1939. Probability density of the determinant of matrix A. $^+$ Monte Carlo on coefficients; • stochastic simulation and re-estimation

13.7 ANALYTIC SIMULATION VERSUS MONTE CARLO ON COEFFICIENTS

It has been pointed out in Section 13.4 that the basic conceptual difference between analytic simulation and Monte Carlo on coefficients lies in the choice of the stage of the computation where asymptotic estimates are considered as approximating the corresponding small sample estimates; the effects of this different choice can be dramatic, as in the examples of Section 13.6.

The two methods are also computationally quite different. However, it is possible to introduce a slight modification into the Monte Carlo method which makes it exactly equivalent to the analytic simulation method (and, in fact, leads to identical results, apart from the sampling variability). This modified Monte Carlo is derived from the method described and applied, for different purposes (standard errors of dynamic multipliers), in Bianchi *et al.* (1979). It is proposed again simply because it shows a kind of logical connection between the two methods.

This Monte Carlo method is carefully designed to take into account explicitly that the normality of the estimated coefficients is asymptotic; the procedure is the following (recall that $\hat{\mathbf{a}}$ and $\hat{\mathbf{y}}_h$ are kept fixed over all the experiment).

1. A tentative value n (greater than 1) is fixed: for example $n = 10$.
2. A vector $\tilde{\mathbf{a}}$ of pseudo-random numbers with multivariate normal distribution, mean equal to $\hat{\mathbf{a}}$ and covariance matrix equal to $\hat{\boldsymbol{\Psi}}/nT$, is generated.
3. The elements of $\tilde{\mathbf{a}}$ are inserted into the model and the model is solved in the forecast period h, producing $\tilde{\mathbf{y}}_h$.
4. The vector of differences $(\tilde{\mathbf{y}}_h - \hat{\mathbf{y}}_h)$ is computed.
5. Remultiplying by the scaling factor $\sqrt{n}$, we compute $\sqrt{n}(\tilde{\mathbf{y}}_h - \hat{\mathbf{y}}_h)$.
6. The steps from 2 to 5 are repeated a certain number of times (for example 1000); each time a new vector $\sqrt{n}(\tilde{\mathbf{y}}_h - \hat{\mathbf{y}}_h)$ is computed.
7. The sample variances (or the covariance matrix) can now be computed for the elements of $\sqrt{n}(\tilde{\mathbf{y}}_h - \hat{\mathbf{y}}_h)$; they are a first approximation to the estimated asymptotic variances of the component of the forecast errors under consideration.
8. After the process has been performed for the first time, n must be increased (for example, by a factor of 10) and a new iteration must be started from step 2; in the other cases, the sample variances obtained in the last iteration must be compared with those of the previous one; if significant differences are found, a new iteration must be started from step 2, after increasing n (for example, again by a factor of 10).

The process is stopped when no differences are found in the first three or four significant digits of the sample variances in two consecutive iterations. To avoid the problem of differences caused by the sampling variability, it is warmly recommended that the same set of pseudo-random numbers are used for each iteration.

The role of the scaling, and final re-scaling, by a value $\sqrt{n}$ should be clear at this point. The above experiment with $n = 1$ is nothing but the Monte Carlo on coefficients method described in Section 13.4.2. A large (and increasing, in theory, to infinite) value of n simulates the behaviour of an increasing sample size (T); it generates a narrow and collapsing distribution of the generated pseudo-random numbers $\tilde{\mathbf{a}}$ around $\hat{\mathbf{a}}$ (as asymptotically the distribution of $\hat{\mathbf{a}}$ collapses around $\mathbf{a}$); for $\tilde{\mathbf{y}}_h$, this generates a distribution collapsing around $\hat{\mathbf{y}}_h$ (as asymptotically the distribution of $\hat{\mathbf{y}}_h$ collapses around $\bar{\mathbf{y}}_h$); finally the collapsed differences $(\tilde{\mathbf{y}}_h - \hat{\mathbf{y}}_h)$ are re-scaled by $\sqrt{n}$ (as the covariance matrix to be estimated is that of the asymptotic distribution of $\sqrt{T}(\hat{\mathbf{y}}_h - \bar{\mathbf{y}}_h)$, divided by the actual value of the sample period length T). Moreover, a large value of n generates pseudo-random numbers with a very narrow distribution, thus avoiding in practice the instability effects due to outliers.

Numerical results obtained with this method do not need to be displayed; they are, in fact, equal to those obtained by means of analytic simulation on coefficients, apart from small differences due to the sampling variability of the Monte Carlo method.

13.8 CONCLUSIONS

This paper has been mainly concerned with empirical comparisons of three methods of analysing the effects of coefficient estimation errors, on forecasts produced by structural econometric models.

The experiments have been designed to include a sufficiently wide and representative set of real-world models. However, it has been pointed out that the methods present not only technical, but also conceptual, differences, so that we cannot conclude with precise criteria for ranking among the three methods. We shall, therefore, confine ourselves to these concluding remarks.

1. From a purely technical point of view, all methods can be applied to macroeconomic models of practical interest.
2. The applicability of analytic simulation strictly depends on the assumption of asymptotic normality of estimated coefficients; in practice, however, the method can be applied also when this assumption is violated and the results seem to be 'numerically' reasonable, even if with unpredictable statistical properties.
3. Monte Carlo on coefficients is not confined to the assumption of coefficient normality. Its correct application, however, would require the knowledge of the distribution of estimated coefficients and the technical availability of a pseudo-random numbers generator from such a distribution. If, in the absence of this information, which usually involves overwhelming difficulties, a normal distribution is adopted, problems like those discussed in Section 13.6 can be encountered.
4. Analytic simulation requires the model to be solved as many times as the number of structural estimated coefficients; furthermore, to compute derivatives with sufficient precision, it needs a computational accuracy much greater than in the usual simulation experiments. Therefore, it could be more efficient to use Monte Carlo for models with more than 40 or 50 coefficients; if outliers do not affect the stability of the Monte Carlo process, in fact, 40 or 50 replications (with fewer problems of accuracy) are usually enough to provide a sufficient approximation to the desired standard errors.
5. Stochastic simulation and re-estimation is computationally the most inefficient method. However, it does not need, among input data, the covariance matrix of all the structural coefficients, thus avoiding the cumbersome problems of its estimation and storing and the problem of its triangularization, which is required by Monte Carlo on coefficients; furthermore, it avoids the problem of storing the matrix of partial derivatives (usually large), which is required by analytic simulation.
6. The method discussed in Section 13.7 can be of some interest for the experimenter; it is, in fact, technically similar to Monte Carlo on coefficients, but conceptually equivalent to analytic simulation.
7. The main problem in applying stochastic simulation and re-estimation or Monte Carlo on coefficients is, in any case, the possible instability in the convergence process due to outliers; in practice stochastic simulation and re-estimation seems to be less sensitive than the other to the presence of outliers. In any case the experimenter cannot simply prefix a number of replications to be performed and then take uncritically the final results; the experimenter, on the contrary, should carefully take into account intermediate results as well.
8. An alternative empirical way of detecting the presence of outliers could be the comparison of the standard errors due to the structural coefficients with the standard errors due to the structural disturbances (the so called 'reduced form' standard errors); it seems reasonable, when forecasting in the first period outside the sample estimation period, to expect two standard errors with the same size, or the error due to structural disturbances slightly larger than the other; when, on the contrary, the error due to coefficients is larger than the other (2, 3, 4 times or more, as in Tables 13.6–13.9), problems of outliers have probably occurred.
9. When problems of instability in the convergence process have not been encountered, the three methods always produced similar results, thus suggesting a kind of empirical equivalence of the methods themselves.

NOTES

[1] The authors gratefully acknowledge helpful comments and criticisms from John Bradley, Ray C. Fair, Lawrence R. Klein, Jacques Mairesse, Joel L. Prakken, and Peter K. Schoenfeld, but retain sole responsibility for any errors.

[2] Reference can be made, for example, to Goldberger *et al.* (1961), Denton and Kuiper (1965), Feldstein (1971), Schink (1971), Haitovsky and Wallace (1972), Cooper and Fischer (1974), Schmidt (1974), Oliver (1979), Bianchi and Calzolari (1980), and Fair (1980).

[3] It is perhaps worth mentioning that these means and variances are sometimes referred to as deterministic simulation bias and reduced form variances, respectively.

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 Calzolari, G. (1980). 'The one-period forecast errors in nonlinear econometric models.' *International Economic Review*, **21**, 201–208.

Bianchi, C., Calzolari, G., and Corsi, P. (1979). 'A Monte Carlo approach to compute the asymptotic standard errors of dynamic multipliers.' *Economics Letters*, **2**, 161–164.

Bianchi, C., Calzolari, G., and Corsi, P. (1981). 'Standard errors of multipliers and forecasts from structural coefficients with block-diagonal covariance matrix.' In J. M. L. Janssen, L. F. Pau, and A. J. Straszak, (Eds), *Dynamic Modelling and Control of National Economies* (IFAC), Pergamon, Oxford, pp. 311–316.

Brundy, J. M., and Jorgenson, D. W. (1971). 'Efficient estimation of simultaneous equations by instrumental variables.' *Review of Economics and Statistics*, **53**, 207–224.

Christ, C. F. (1966). *Econometric Models and Methods*, Wiley, New York.

Conrad, K., and Kohnert, P. (1979). 'Economic activity, interest rates and the exchange rate in the Bonn forecasting system no. 10.' Universität Bonn, Institut für Gesellschafts- und Wirtschaftswissenschaften, Wirtschaftstheoretische Abteilung, discussion paper.

Cooper, J. P., and Fischer, S. (1974). 'Monetary and fiscal policy in the fully stochastic St. Louis econometric model.' *Journal of Money, Credit, and Banking*, **6**, 1–22.

Denton, F. T., and Kuiper, J. (1965). 'The effects of measurement errors on parameter estimates and forecasts.' *Review of Economics and Statistics*, **47**, 198–206.

Dhrymes, P. J. (1970). *Econometrics: Statistical Foundations and Applications*, Harper & Row, New York.

Dhrymes, P. J. (1973). 'Restricted and unrestricted reduced forms: Asymptotic distribution and relative efficiency.' *Econometrica*, **41**, 119–134.

Fair, R. C. (1980). 'Estimating the expected predictive accuracy of econometric models.' *International Economic Review*, **21**, 355–378.

Feldstein, M. S. (1971). 'The error of forecast in econometric models when the forecast-period exogenous variables are stochastic.' *Econometrica*, **39**, 55–60.

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., Nagar, A. L., and Odeh, H. S. (1961). 'The covariance matrices of reduced-form coefficients and of forecasts for a structural econometric model.' *Econometrica*, **29**, 556–573.

Haitovsky, Y., and Wallace, N. (1972). '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*, National Bureau of Economic Research, New York, pp. 261–309.

Hatanaka, M. (1978). 'On the efficient estimation methods for the macro-economic models nonlinear in variables.' *Journal of Econometrics*, **8**, 323–356.

Hausman, J. A. (1974). 'Full information instrumental variables estimation of simultaneous equations systems.' *Annals of Economic and Social Measurement*, **3**, 641–652.

Klein, L. R. (1950). *Economic Fluctuations in the United States, 1921–1941*, Cowles Commission Monograph 11, Wiley, New York.

Klein, L. R. (1969). 'Estimation of interdependent systems in macroeconometrics.' *Econometrica*, **37**, 171–192.

Krelle, W. (1974). *Erfahrungen mit einem Ökonometrischen Prognosemodell für die Bundesrepublik Deutschland*, Verlag Anton Hain, Meisenheim/Glan.

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

Mariano, R. S. (1980). 'Analytical small sample distribution theory in econometrics: The simultaneous equations case.' Université Catholique de Louvain, Center for Operations Research & Econometrics, discussion paper.

McCarthy, M. D. (1972a). 'Some notes on the generation of pseudo-structural errors for use in stochastic simulation studies.' In B. G. Hickman (Ed.), *Econometric Models of Cyclical Behavior*, National Bureau of Economic Research, Studies in Income and Wealth no. 36, New York, pp. 185–191.

McCarthy, M. D. (1972b). 'A note on the forecasting properties of two-stage least squares restricted reduced forms: The finite sample case.' *International Economic Review*, **13**, 757–761.

Nagar, A. L. (1969). 'Stoichastic 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*, North-Holland, Amsterdam, pp. 425–456.

Nissen, D. H. (1968). 'A note on the variance of a matrix.' *Econometrica*, **36**, 603–604.

Oliver, F. R. (1979). 'Rebased and revised data: Their effects on econometric estimates.' *Economics Letters*, **4**, 355–359.

Quinke, H. (1978). 'Bonner Modell 5.5, Variablen, Gleichungen, Daten 1965–1977, Prognosen 1978–1985, Zusammenstellung.' Universität Bonn, Institut für Gesellschafts- und Wirtschaftswissenschaften, Wirtschaftstheoretische Abteilung, discussion paper.

Rao, C. R. (1965). *Linear Statistical Inference and its Applications*, Wiley, New York.

Sargan, J. D. (1976). 'The existence of the moments of estimated reduced form coefficients.' London School of Economics and Political Science, discussion paper A6.

Sartori, F. (1978). 'Caratteristiche e struttura del modello.' In *Un Modello Econometrico dell'Economia Italiana: Caratteristiche e Impiego, 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. (1976). 'The asymptotic distribution of forecasts in the dynamic simulation of an econometric model.' *Econometrica*, **42**, 303–309.

Schmidt, P. (1976). *Econometrics*, Dekker, New York.

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