# Stochastic simulation experiments on Model 5 of Bonn University

Calzolari, Giorgio

IBM Scientific Center, Pisa, Italy

August 1979

# Stochastic Simulation Experiments on Model 5 of Bonn University

by Giorgio Calzolari

Bonn, August 1979 N. 102

Address of the author:
Centro Scientifico IBM
Via S. Maria 67
56100 Pisa, Italia

## Contents


| | | |
|---|---|---|
| 1. | Introduction | p. 3 |
| 2. | Standard errors of the reduced form equations | p. 4 |
| 2.1. | Algorithms and computational note | p. 9 |
| 3. | Heteroschedasticity of the reduced form | p. 10 |
| 4. | Reduced form in dynamic models | p. 12 |
| 4.1. | Computational note | p. 15 |
| 5. | Deterministic simulation bias | p. 15 |
| 5.1. | Accurate measurement using antithetic variate sampling | p. 17 |
| 6. | A short note on the forecast errors | p. 20 |
| 7. | The program installed at Bonn University | p. 23 |
| | References | p. 26 |


## GLOSSARY

| | |
|---|---|
| YDP'NO | Gross Domestic Product MRDDM |
| YDP | Gross Domestic Product MRD62DM |
| Y'PP | Profit Income P MRDDM |
| Y'WG | Wage Income G MRDDM |
| C'PNO | Private Consumption MRDDM |
| I'PN | Gross Private Investment MRDDM |
| T | Total Tax Payments MRDDM |
| GS'GD | Gov. Demand Goods and Services MRDDM |
| P | Defl. Gross Dom. Product P B62 |
| HW | Hours of Work in the Economy MRDHR |
| FW | Foreign Workers |
| ETA | Productivity of Labour 62DM/HR |
| U | Unemployed Persons MIO |
| M'GSNO | Goods and Services Imports MRDDM |
| X'GSNO | Goods and Services Exports MRDDM |
| BP'GS | Bal. of Paym. Goods and Services MRDDM |

## 1. Introduction

When planing this work on the econometric model developed by the University of Bonn, it seemed that a suitable title could be "Stochastic Simulation and Dynamic Properties of Model 5". However, after most of the preparatory computational work had been done, it clearly appeared that the model was unable to undergo the long simulation runs which are required by dynamic analysis; in other words, even in its last version (Model 5.5 [20]) the model is probably instable, as already its previous version seemed to be [14].

It was therefore decided to concentrate the experiments on the stochastic properties of the one-period (static) and short-multi-period (dynamic) simulations and, in particular, on the problem of the distribution of the disturbances of the reduced form equation

It is purpose of this paper to present on one side, the numerical results obtained on Model 5 and, on the other side, to give some details (or at least some bibliographical references) on the algorithms which have been used.

The algorithms briefly discussed will be mainly presented "by analogy with the case of linear models. This, of course, leaves open any kind of doubts about the methodological consistency of the methods, but has, at the same time, the advantage of making the algorithms themselves immediately comprehensible to the"empirical econometrician"and of showing a simple computational method of solution.

It must be pointed out that the disturbance terms which are taken into account in these experiments are only the random disturbances of the structural form equations; no attempt is done to take into account the other possible sources of errors (briefly summarized in section 6) which involve problems considerably more complex.

The basic method to deal with the problem of the reduced form disturbances in nonlinear models is the stochastic simulation (Monte Carlo) In some sense, according to the standard hypotheses underlying an econometric model, it should reproduce the process which has generated the time series of the endogenous variables. It has, however, the disadvantage of involving a sampling variability due to the generation of pseudo-random numbers and, even more, of involving always a great deal of computational resources (time and storage).

Accurate substitutes of stochastic simulation can be, in some cases, obtained with a mixed use of numerical simulation and analytical methods; even if used for different purposes, it seems convenient to adopt, for these methods, the definition of "analytic simulation" proposed by Howrey and Klein [ 13 ] . These methods are based on linearisations of the model in the neighbourhood of the solution points (changing year by year); they involve, therefore, an approximation, but seem to be, in practice, more accurate than the stochastic simulation which only asymptotically would be able to lead to the correct results.

Most of the experiments here described have been performed either with stochastic or with analytic simulation.

## 2. Standard errors of the reduced form equations

Let

(2.1) $$Ay_t + Bz_t = u_t \qquad t=1,2,...,n$$

be a linear econometric model in its structural form, where $y_t$, $z_t$ and $u_t$ are, respectively, the vectors of the endogenous and predetermined variables and of the structural stochastic disturbances at time t, while A and B are matrices of structural coefficients (A is a nonsingular square matrix). Furthermore let $u_t$ be distributed as

(2.2) $$u_t \sim N\ (O,\Sigma) \qquad \text{cov}\ (u_t\ u_{t'}) = \delta_{tt'}\Sigma\ ;$$

in other words the vectors $u_t$ are supposed to be independent and identically distributed, with a multivariate normal distribution, zero means and covariance matrix constant over time.

The estimated structural model is

(2.3) $$\hat{A}y_t + \hat{B}z_t = \hat{u}_t$$

where $\hat{u}_t$ are the regression residuals and

$$(2.4) \quad \frac{1}{n} \sum_{t=1}^{n} \hat{u}_t \hat{u}_t' = \hat{\Sigma}$$

is an estimate of the covariance matrix of the structural form disturbances (or, which is the same, of the structural form equations).

The restricted reduced form (i.e. the reduced form derived from the structural form, thus taking into account all the restrictions on coefficients) is:

$$(2.5) \quad y_t = -A^{-1} B z_t + v_t$$

where

$$(2.6) \quad v_t = A^{-1} u_t$$

is the vector of the reduced form disturbances at time t.

It is clearly

$$(2.7) \quad v_t \sim N(O, A^{-1} \Sigma A^{-1'})$$

so that an estimate of the reduced form covariance matrix is immediately available as

$$(2.8) \quad \hat{A}^{-1} \hat{\Sigma} \hat{A}^{-1'}$$

provided the estimated $\hat{A}$ is nonsingular.

If the model is nonlinear, obviously the above procedure cannot be applied. Let

$$(2.9) \quad f(y_t, z_t, a) = u_t$$

be the representation of a nonlinear structural model where, this time, a is a vector including all the structural coefficients (it is no more possible a clear distinction between coefficients of $y_t$ and coefficients of $z_t$). Even if no explicit reduced form can be derived, it is generally assumed that, for any set of values of $z_t$, a and $u_t$, a unique vector $y_t$ satisfies equation (2.9) [8, p.172]. This is equivalent to assume the existence of a reduced form, implicitly defined by the structural form. If we enter into equation (2.9) with the "true" values of the coefficients a, with the historical (or forecasted,but assumed exact) values of the predetermined variables $z_t$ and with a vector of random disturbances with the same distribution of $u_t$ and solve the model for $y_t$, the result is a random vector with the same statistical properties as the unknown reduced form of the model. If this process is repeated several times, the sample moments of the computed values of the endogenous variables should converge to the corresponding distribution moments. Nothing can be said, in general, about the existence of finite moments of the distribution of $y_t$. While in case of linear models no problem arises, in case of nonlinear models the transformations of the normally distributed $u_t$ can lead to distributions without finite moments.

From a purely empirical point of view, fortunately, this generally creates no trouble. The cause, in fact, of the possible non-existence of finite moments lays, generally, in one of the tails of the distribution of the structural disturbances, but so far from the central region that no generator of pseudo-random numbers will ever be able to generate numbers in that area. A simple example can better illustrate this problem.

A price variable (endogenous) appears on the left hand side of a linear equation, involving only exogenous regressors; for example:

$$p_t = a + bx_t + u_t$$

This price variable is used as deflator of another endogenous variable of the model $y_t/p_t$. If $p_t$ has values in the range 1÷2, and the standard deviation of $u_t$ is $\sigma=0.01$, $p_t$ will not assume values less than or equal to zero, unless $u_t$ assumes negative values besides $-100\,\sigma$ and no "reasonable" generator of normal deviates can generate values in such a tail of the distribution (as well as no model builder would hypotize for his price variable, negative values with nonzero probability! Quite similar would be the case of logarithms of possibly negative numbers).

The number of replications of stochastic simulations can be pre-fixed on the basis of the $\chi^2/df$ distribution[9] : for example 1000 replications assure that the computed sample standard deviations differ from the "true" reduced form standard errors no more than 5% with probability greater than 0.96

(provided a finite reduced form variance exists).

Alternative to stochastic simulation could be the use of analytic simulation procedures. These are based on a nonexplicit linearization of the model in the neighbourhood of the solution point corresponding to the year under examination. It is clear from equations (2.5) and (2.6) that the elements of the matrix $A^{-1}$ (such that $A^{-1}u_t=v_t$, reduced form disturbances) are the partial derivatives of the endogenous variables with respect to the structural disturbances at time t (elements of the vector $u_t$). These derivatives can be computed via numerical simulation, stored into a matrix $D_t$ (=$A^{-1}$ for linear models, but time-varying in case of nonlinearity) and the reduced form covariance matrix at time t can be computed as:

(2.10) $\hat{D}_t \hat{\Sigma} \hat{D}'_t$

Table 1 displays the results on Model 5 for the year 1978 (the first of the "example forecast" in [20]). For the main endogenous variables of the model, the computed values are displayed (they have been obtained with a one-period static simulation) together with the corresponding reduced form standard errors computed with 10, 50, 500 and 2000 replications of stochastic simulation and with analytic simulation. The last column displays the coefficients of variation, which are the ratios of each standard error with the computed value of the variable, in percentage form.

From Table 1 one could get the strong impression that the stochastic simulation results converge to those of analytic simulation as the number of replications goes to infinity. This is of course not exact due to the nonlinearity of the model, but clearly gives an idea of the great accuracy of the analytic simulation method.

Table 1

One-step simulation at 1978

| Variable | Computed Value | Reduced Form standard errors 10 Repl. | 50 Repl. | 500 Repl. | 2000 Repl. | Analytic simulation | Coefficient of variation |
|---|---|---|---|---|---|---|---|
| Y( 3)=YDP'NO | 1298. | 20.9 | 22.4 | 23.2 | 21.3 | 21.3 | 1.64 |
| Y( 4)=YDP | 652.9 | 8.68 | 11.7 | 11.4 | 10.9 | 10.9 | 1.67 |
| Y( 7)=Y'PP | 277.3 | 13.1 | 13.4 | 13.5 | 13.9 | 14.0 | 5.03 |
| Y( 9)=Y'WG | 149.0 | 2.71 | 2.50 | 2.57 | 2.68 | 2.72 | 1.83 |
| Y( 14)=C'PNO | 714.9 | 9.70 | 8.83 | 9.99 | 9.64 | 9.53 | 1.33 |
| Y( 18)=I'PN | 238.7 | 17.1 | 15.7 | 16.1 | 14.7 | 14.3 | 6.01 |
| Y( 27)=T | 508.7 | 15.8 | 15.2 | 14.5 | 14.1 | 14.2 | 2.79 |
| Y( 35)=GS'GD | 168.1 | 1.85 | 1.91 | 2.03 | 1.95 | 1.97 | 1.17 |
| Y( 36)=P | 189.0 | 4.22 | 3.54 | 3.38 | 3.18 | 3.16 | 1.67 |
| Y( 59)=HW | 45.03 | .807 | .706 | .706 | .667 | .673 | 1.49 |
| Y( 65)=FW | 2.013 | .203 | .210 | .218 | .211 | .213 | 10.6 |
| Y( 82)=ETA | 14.49 | .365 | .293 | .275 | .277 | .269 | 1.85 |
| Y( 83)=U | 1.100 | .151 | .169 | .173 | .174 | .174 | 15.8 |
| Y(107)=M'GSNO | 326.2 | 8.88 | 8.65 | 8.62 | 8.16 | 8.34 | 2.56 |
| Y(123)=X'GSNO | 344.5 | 6.05 | 5.39 | 5.09 | 5.27 | 5.30 | 1.54 |
| Y(126)=BP'GS | 18.25 | 12.2 | 11.1 | 10.3 | 10.0 | 10.2 | 56.0 |

## 2.1 Algorithms and computational note

The computation of the reduced form standard errors via stochastic simulation has been performed by means of a modified version of the program announced in [2], developed at the IBM Scientific Center of Pisa.

The generation of the pseudo-random disturbances to be inserted, during the solution phase, into the structural equations, has been performed in three main steps (see [7] for details):

1) Generation of independent pseudo-random variables with uniform distribution, using the power residue method [15] and a final shuffling.
2) Transformation of the uniform numbers into independent standard normal deviates, using the algorithm by Box and Muller [6].
3) Transformation of the standard normal deviates into pseudo-random disturbances with zero means and covariance matrix equal to the sample covariance matrix of the regression residuals (covariances are between equations at the same period of time, while zero covariances are assumed between disturbances in different time periods, as in assumption (2.2)). The algorithm by McCarthy [16] has been used.

The computation of the $\hat{D}_t$ matrix (partial derivatives of the endogenous variables, with respect to the structural disturbances, in the solution point at time t) required by the analytic simulation method, has been performed using finite increments on the structural disturbances. More exactly, first a deterministic control solution has been computed, at time t, with all the $u_t$ set to zero. Then a value $\varepsilon$ is assigned to the disturbance of the first equation, all the other having still zero, and the model is solved again. The same is then repeated for all the structural stochastic equations and the differences between the disturbed solutions and the control solution, divided by the values adopted for $\varepsilon$, supply the numerical values of the partial derivatives. Variations of $\varepsilon$ from $0.001\sigma_n$ to $0.000001\sigma_n$, for each stochastic equation, did not modify the results in the first 3 decimal significant digits, thus showing a good stability of the algorithms in the application to Model 5. Of course, to appreciate the small differences between the disturbed solutions and the control solution, a quite small tolerance had to be fixed for the convergence of the Gauss-Seidel solution algorithm; a relative tolerance $10^{-13}$ has been chosen in these experiments.

## 3. Heteroschedasticity of the reduced form

When a model is linear, an alternative method can be used instead of equation (2.8) (or (2.10), or stochastic simulation) to compute the standard errors of the reduced form equations. It follows, in fact, from equations (2.4), (2.6) and (2.8) that

$$(3.1) \quad \hat{A}^{-1}\hat{\Sigma}\hat{A}^{-1'} = \frac{1}{n}\sum_{t=1}^{n} (\hat{A}^{-1}\hat{u}_t)(\hat{A}^{-1}\hat{u}_t)' = \frac{1}{n}\sum_{t=1}^{n} \hat{v}_t\hat{v}_t'$$

where the values of $\hat{v}_t$, as it is clear from equation (2.5), are simply the differences between the historical values of $y_t$ and the values of the endogenous variables computed as $-\hat{A}^{-1}\hat{B}z_t$ all over the sample period. Therefore the diagonal elements of the matrix in (3.1) are simply the mean squared errors of the one-step (static) simulation over all the sample period. And this is, probably, the simplest computational approach to the problem of the variances of the reduced form equations.

Unfortunately this method cannot be applied to nonlinear models. A simple example will make this clear.

If $y_t$ is an endogeneous variable which appears on the left hand side of a structural equation in the form $y_t/z_t$, where $z_t$ is a time varying predetermined variable, and if the structual form equation is supposed to be homoschedastic, clearly $y_t$ will be heteroschedastic in the reduced form, with a variance changing over time with the square of $z_t$. Therefore the mean squared error of $y_t$ over the sample period will be simply an average estimate of goodness of fit, but will not be equal to the reduced form variance, as this changes with time. Moreover, since the economic time series generally increase with time, for some variables it can happen that the reduced form variance in the last years of the sample(or in the forecast periods) is greater than the mean squared error computed over the sample period. This means that the use of the mean squared errors as approximate measurements of the variances of forecasts will probably lead to an underestimate of the magnitude of the forecast errors of several variables.

Table 2 exemplifies this problem. For some of the main variables of Model 5 the root mean squared errors over the period 1960-1975 and the reduced form standard errors in some years are displayed.

Table 2

| Variable | Root Mean Squared Error 1960-1975 | Reduced form standard errors 1960 | 1970 | 1975 | 1978 |
|---|---|---|---|---|---|
| Y( 3)=YDP'NO | 13.1148 | 14.5 | 16.8 | 17.7 | 21.3 |
| Y( 4)=YDP | 14.5708 | 11.2 | 11.9 | 9.78 | 10.9 |
| Y( 7)=Y'PP | 10.0924 | 7.24 | 9.86 | 10.7 | 14.0 |
| Y( 9)=Y'WG | 2.01722 | .821 | 1.38 | 2.20 | 2.72 |
| Y( 14)=C'PNO | 7.55442 | 7.74 | 7.88 | 7.98 | 9.53 |
| Y( 18)=I'PN | 9.76139 | 8.33 | 12.1 | 12.3 | 14.3 |
| Y( 27)=T | 6.63770 | 4.47 | 7.66 | 11.4 | 14.2 |
| Y( 35)=GS'GD | 1.96941 | 1.97 | 1.97 | 1.97 | 1.97 |
| Y( 36)=P | 2.95902 | 3.15 | 3.05 | 2.99 | 3.16 |
| Y( 59)=HW | 1.27758 | 1.69 | .914 | .684 | .673 |
| Y( 65)=FW | .347212 | .424 | .306 | .229 | .213 |
| Y( 82)=ETA | .135602 | .570D-01 | .143 | .199 | .269 |
| Y( 83)=U | .202692 | .210 | .145 | .165 | .174 |
| Y(107)=M'GSNO | 4.28544 | 3.05 | 4.61 | 6.52 | 8.34 |
| Y(123)=X'GSNO | 2.69351 | 1.85 | 2.80 | 4.46 | 5.30 |
| Y(126)=BP'GS | 5.64082 | 3.06 | 5.72 | 8.17 | 10.2 |

## 4. Reduced form in dynamic models

A model is dynamic if, among the predetermined variables $z_t$ of equation (2.1), some are lagged endogenous variables. In these cases a convenient representation of the structural form of the model can be:

$$(4.1) \quad A\, y_t + B\, x_t + C\, y_{t-1} = u_t$$

where no loss of generality is caused by the explicit presence of only lag-one endogenous variables as also higher order lags can be reconducted to lag-one by the proper insertion of additional definitional (nonstochastic) equations [25]. If the model is dynamic, the simulation at time t can be performed in several different ways. The one-step (or one-period, or static) simulation considers $y_{t-1}$ as predetermined variables, and assigns to them the historical values; this is the case already discussed in section 2. The dynamic (or multiperiod dynamic) simulation starting from the period t-1 uses for $y_{t-1}$ the values computed in a one-step solution (therefore considering $y_{t-2}$ as given); in formula:

$$(4.2.) \quad y_t = -A^{-1}Bx_t - A^{-1}Cy_{t-1} + A^{-1}u_t =$$

$$= -A^{-1}Bx_t + A^{-1}CA^{-1}Bx_{t-1} + A^{-1}CA^{-1}Cy_{t-2} - A^{-1}CA^{-1}u_{t-1} + A^{-1}u_t$$

Equation (4.2) can be clearly extended to the case of simulations starting from the period t-2, t-3,...,t-s. It is clear from equation (4.2) that the reduced form error depends either on $u_t$ or on $u_{t-1}$ (in general $u_{t-k}$, k=0,1,2,...,s if the dynamic simulation starts from the period t-s). Under the assumptions (2.2) the covariance matrix of $y_t$ in equation (4.2), given $y_{t-2}$, is the sum of two components, respectively due to $u_{t-1}$ and $u_t$:

$$(4.3) \quad (A^{-1}CA^{-1})\Sigma(A^{-1}CA^{-1})' + A^{-1}\Sigma\, A^{-1'}$$

If the model is nonlinear, resort can be done to stochastic simulation or to an analytic simulation procedure. The latter computes something similar to equation (4.3), considering that

$-A^{-1}CA^{-1}$ is nothing but the matrix of the partial derivatives of $y_t$ with respect to $u_{t-1}$, as it is clear from (4.2); in case of dynamic simulation from t-s to t, the reduced form covariance matrix at time t will be:

$$(4.4) \quad \sum_{k=0}^{s} (D_{t,t-k} \Sigma D'_{t,t-k})$$

where the elements of the matrix $D_{t,t-k}$ are the partial derivatives of the endogenous variables at time t with respect to the structural disturbances at time t-k.

Table 3 displays, for some of the main variables of Model 5, the standard errors of the reduced form in the dynamic simulations from 1978 to 1980. As already in section 2, the results have been obtained either by means of stochastic simulation, with various numbers of replications, or by means of analytic simulation; only those obtained by means of analytic simulation are displayed, as they seemed to be more accurate.

Table 3

DYNAMIC SIMULATION 1978-1980
REDUCED FORM STANDARD ERRORS

| VARIABLE | VALUE | STD.ERR. | CFF.VAR. | |
|---|---|---|---|---|
| Y( 3)=YDP'NO | 1298.70 | 21.3 | 1.64 | 1978 |
| Y( 4)=YDP | 652.938 | 10.9 | 1.67 | |
| Y( 7)=Y'PP | 277.397 | 14.0 | 5.03 | |
| Y( 9)=Y'WG | 149.009 | 2.72 | 1.83 | |
| Y( 14)=C'PNO | 714.977 | 9.53 | 1.33 | |
| Y( 18)=I'PN | 238.704 | 14.3 | 6.01 | |
| Y( 27)=T | 508.744 | 14.2 | 2.79 | |
| Y( 35)=GS'GD | 168.145 | 1.97 | 1.17 | |
| Y( 36)=P | 189.051 | 3.16 | 1.67 | |
| Y( 59)=HW | 45.0360 | .673 | 1.49 | |
| Y( 65)=FW | 2.01355 | .213 | 10.6 | |
| Y( 82)=ETA | 14.4981 | .269 | 1.85 | |
| Y( 83)=U | 1.10054 | .174 | 15.8 | |
| Y(107)=M'GSNO | 326.261 | 8.34 | 2.56 | |
| Y(123)=X'GSNO | 344.521 | 5.30 | 1.54 | |
| Y(126)=BP'GS | 18.2598 | 10.2 | 56.0 | |
| Y( 3)=YDP'NO | 1401.88 | 28.1 | 2.01 | 1979 |
| Y( 4)=YDP | 678.633 | 14.6 | 2.16 | |
| Y( 7)=Y'PP | 291.738 | 18.8 | 6.44 | |
| Y( 9)=Y'WG | 165.588 | 4.11 | 2.48 | |
| Y( 14)=C'PNO | 761.494 | 12.7 | 1.66 | |
| Y( 18)=I'PN | 251.036 | 19.2 | 7.63 | |
| Y( 27)=T | 554.741 | 16.7 | 3.02 | |
| Y( 35)=GS'GD | 185.191 | 4.63 | 2.50 | |
| Y( 36)=P | 195.412 | 3.88 | 1.98 | |
| Y( 59)=HW | 44.6246 | .805 | 1.80 | |
| Y( 65)=FW | 1.59543 | .241 | 15.1 | |
| Y( 82)=ETA | 15.2076 | .304 | 2.00 | |
| Y( 83)=U | .898125 | .204 | 22.7 | |
| Y(107)=M'GSNO | 359.547 | 9.82 | 2.73 | |
| Y(123)=X'GSNO | 387.488 | 6.07 | 1.57 | |
| Y(126)=BP'GS | 27.9416 | 11.7 | 41.8 | |
| Y( 3)=YDP'NO | 1534.55 | 36.2 | 2.36 | 1980 |
| Y( 4)=YDP | 709.058 | 18.5 | 2.61 | |
| Y( 7)=Y'PP | 304.982 | 23.6 | 7.75 | |
| Y( 9)=Y'WG | 186.728 | 5.67 | 3.04 | |
| Y( 14)=C'PNO | 814.497 | 16.0 | 1.97 | |
| Y( 18)=I'PN | 280.242 | 26.2 | 9.34 | |
| Y( 27)=T | 622.350 | 20.5 | 3.29 | |
| Y( 35)=GS'GD | 202.480 | 5.64 | 2.78 | |
| Y( 36)=P | 203.647 | 4.36 | 2.14 | |
| Y( 59)=HW | 44.5801 | .961 | 2.16 | |
| Y( 65)=FW | 1.42006 | .366 | 25.8 | |
| Y( 82)=ETA | 15.9053 | .339 | 2.13 | |
| Y( 83)=U | .590272 | .235 | 39.8 | |
| Y(107)=M'GSNO | 394.443 | 11.4 | 2.89 | |
| Y(123)=X'GSNO | 433.353 | 6.78 | 1.56 | |
| Y(126)=BP'GS | 38.9100 | 13.4 | 34.5 | |

### 4.1 Computational note

No special problem arises if stochastic simulation is used to compute the reduced form variances in dynamic simulation. The simulations to be performed must be, at the same time, stochastic and dynamic. Of course, if the model is not stable, convergence problems will arise as soon as the simulation period becomes sufficiently long. For this reason the experiments on Model 5 were confined to simulation runs of 3 years: from 1978 to 1980.

The use of analytic simulation involves some problems which were not present in the case of one-period simulation and are also not present in the dynamic simulation of linear models.

Since simulation is always performed foreward, it would be much simpler to compute the elements of $D_{t',t-k}$ of equation (4.4) as partial derivatives, with respect to a disturbance at the origin of simulation (time t-s), of the endogenous variables "after" k simulation periods. This would lead to correct results on linear models, but, unfortunately, not on nonlinear models, whose behaviour is not symmetric over time and, moreover, heteroschedastic. Therefore, to compute numerically the elements of $D_{t,t-k}$, something like a set of "backwards" simulations must be performed; all of them must finish at time t, one must start at time t, one must start at time t-1 and so on,untill the last one, that must start at time t-s.

The procedure is rather laborious and time consuming, so that the advantages with respect to the stochastic simulation method are no more so clearly evident as in the case of one-period simulation.

## 5. Deterministic Simulation Bias

It is a well known statement that, in nonlinear models, the deterministic simulation values "can be expected to diverge systematically from the corresponding" historical values [12,p.309]. In fact the nonlinear transformations of the random disturbances, when passing from the structural form to the reduced form, do not maintain zero means, so that the conditional expectation (given coefficients and values of the predetermined variables) of the solution error (computed minus observed value of each endogenous variable) will be generally non-zero.

As already observed in section 2 , the existence of a finite value for this conditional expectation must be assumed. More exactly, it must be assumed that the distribution of the structural disturbances is truncated in such a way as to guarantee the existence of a finite conditional expectation of the endogenous variables (and, therefore , of the solution error); a very strong assumption, in theory, but not particularly troublesome in practice, as already pointed out.

To check the existence of a bias for some variables, the number of replications of stochastic simulation must be generally very high. A convenient stopping rule could be the following: increase the number of replications until the difference between the deterministic and the mean stochastic solution for the examined variable is greater than the standard deviation, computed across the replications, of the mean stochastic solution (which decreases and goes to zero as the number of trials increases) or, even better, greater than double the standard deviation.

As far as the practical experience of the IBM Scientific Center of Pisa is concerned, no experiment of this kind (on models of practical interest)has been found successful with less than several thousands replications.
As also on Model 5 an experiment with 2000 replications was not enough to get, for most of the variables, an "estimated bias" sufficiently larger (in absolute value) than its estimated standard deviation, an alternative method has been experimented.

## 5.1. Accurate measurement using antithetic variate sampling

The stochastic simulation experiments previously performed are based on repeated solutions of the model, each time inserting a new vector of disturbances, independently generated at each new solution; they could be called Monte Carlo experiments with simple (or independent) random sampling. In the case of experiments to measure the deterministic simulation bias, the variance which is associated to the estimate of such a bias is too high, when compared with the value of the bias itself. It would be,therefore, convenient to use a method which strongly reduces the variance associated with a single simulation run.
The method which has been experimented is based on antithetic variate sampling [18]. In case of one-step simulation at time t, it can be applied in the following way.

1) The model is solved deterministically at time t.
2) A vector of additive pseudo-random structural disturbances $\tilde{u}_t$ is inserted and the model solved.
3) The same vector, with the opposite sign $(-\tilde{u}_t)$, is inserted, the model again solved and the computed values of the endogenous variables are averaged with those obtained at step 2.
4) The so obtained means of the endogenous variables, subtracted from the deterministic simulation values, supply an estimate of the bias with a variance which is, in this case, much smaller than in the case of a couple of replications with simple random sampling (for example, the variance is exactly zero for all the nonsimultaneous endogenous variables of any model with additive disturbances and it would be zero for all the variables of a linear model).

The process from step 2 to 4 can be replicated several times, thus further reducing the variance and allowing, at the same time, to compute the sample standard deviation of the mean.

Table 4 first displays the results related to the simulation of Model 5 at 1978. The estimated bias is the difference between the deterministic simulation value and the mean of the stochastic solutions; therefore a positive value (if significantly positive, i.e. with a small standard deviation and a large T value) means that deterministic simulation systematically will overestimate the conditional expectation of the endogenous variables.

This in principle; in practice the estimated bias of all the variables is so small, when compared with the values of the variables (in table 1) to be probably of no usefulness to the model's user (the largest bias is for the variable BP'GS and does not exceed 1% of the value of the variable). As an analogous conclusion has been drawn in all the previous experiences of the author, it could be possible to conclude that the problem of the deterministic simulation bias in nonlinear econometric models, though of extremely high theoretical interest, is of quite poor practical usefulness. A quite similar conclusion can be derived from the results of the computation of the bias in case of dynamic simulation, which are displayed always in table 4.

Table 4

DYNAMIC SIMULATION 1978-1980. DETERMINISTIC SIMULATION
BIAS COMPUTED WITH 500 COUPLES OF ANTITHETIC SAMPLES

| VARIABLE | BIAS | STD.DEVIAT. | T | |
|---|---|---|---|---|
| Y( 3)=YDP'NO | .114 | .669D-01 | 1.71 | 1978 |
| Y( 4)=YDP | .154 | .242D-01 | 6.37 | |
| Y( 7)=Y'PP | -.879D-01 | .342D-01 | -2.56 | |
| Y( 9)=Y'WG | .623D-02 | .270D-02 | 2.30 | |
| Y( 14)=C'PNO | .103 | .274D-01 | 3.76 | |
| Y( 18)=I'PN | -.158 | .392D-01 | -4.02 | |
| Y( 27)=T | -.467D-01 | .207D-01 | -2.25 | |
| Y( 35)=GS'GD | .355D-14 | .0 | .0 | |
| Y( 36)=P | -.596D-01 | .510D-02 | -11.6 | |
| Y( 59)=HW | .187D-01 | .147D-02 | 12.7 | |
| Y( 65)=FW | .422D-02 | .338D-03 | 12.4 | |
| Y( 82)=ETA | -.460D-02 | .371D-03 | -12.4 | |
| Y( 83)=U | -.301D-02 | .241D-03 | -12.4 | |
| Y(107)=M'GSNO | -.320 | .138D-01 | -23.1 | |
| Y(123)=X'GSNO | -.157 | .629D-02 | -25.0 | |
| Y(126)=BP'GS | .163 | .135D-01 | 12.0 | |
| Y( 3)=YDP'NO | -.147 | .804D-01 | -1.83 | 1979 |
| Y( 4)=YDP | .133 | .335D-01 | 3.99 | |
| Y( 7)=Y'PP | -.170 | .407D-01 | -4.17 | |
| Y( 9)=Y'WG | -.398D-01 | .492D-02 | -8.09 | |
| Y( 14)=C'PNO | .754D-01 | .330D-01 | 2.28 | |
| Y( 18)=I'PN | -.351 | .546D-01 | -6.42 | |
| Y( 27)=T | -.171 | .298D-01 | -5.73 | |
| Y( 35)=GS'GD | -.133D-01 | .593D-02 | -2.25 | |
| Y( 36)=P | -.999D-01 | .674D-02 | -14.8 | |
| Y( 59)=HW | .214D-01 | .176D-02 | 12.1 | |
| Y( 65)=FW | .590D-02 | .436D-03 | 13.5 | |
| Y( 82)=ETA | -.619D-02 | .450D-03 | -13.7 | |
| Y( 83)=U | -.506D-02 | .339D-03 | -14.9 | |
| Y(107)=M'GSNO | -.399 | .227D-01 | -17.6 | |
| Y(123)=X'GSNO | -.217 | .875D-02 | -24.7 | |
| Y(126)=BP'GS | .182 | .219D-01 | 8.30 | |
| Y( 3)=YDP'NO | -.328 | .113 | -2.90 | 1980 |
| Y( 4)=YDP | .283 | .564D-01 | 5.02 | |
| Y( 7)=Y'PP | -.125 | .610D-01 | -2.05 | |
| Y( 9)=Y'WG | -.127 | .968D-02 | -13.1 | |
| Y( 14)=C'PNO | .114D-01 | .475D-01 | .240 | |
| Y( 18)=I'PN | -.375 | .807D-01 | -4.64 | |
| Y( 27)=T | -.298 | .422D-01 | -7.06 | |
| Y( 35)=GS'GD | -.522D-01 | .919D-02 | -5.68 | |
| Y( 36)=P | -.175 | .105D-01 | -16.5 | |
| Y( 59)=HW | .331D-01 | .291D-02 | 11.3 | |
| Y( 65)=FW | .143D-01 | .955D-03 | 14.9 | |
| Y( 82)=ETA | -.606D-02 | .511D-03 | -11.8 | |
| Y( 83)=U | -.603D-02 | .576D-03 | -10.4 | |
| Y(107)=M'GSNO | -.483 | .312D-01 | -15.5 | |
| Y(123)=X'GSNO | -.265 | .106D-01 | -24.8 | |
| Y(126)=BP'GS | .218 | .297D-01 | 7.34 | |

## 6. A Short Note on the Forecast Errors

Unadjusted forecasts of an econometric model are affected by errors whose sources can be summarized as follows:

### 6.1 Models's structure

Some or all the equations may be not correctly specified. Generally one deals with the problem of the forecast errors "conditional on the model's structure". Nevertheless some attempts have been done to get some empirical evidence of the errors due to incorrect specification of a model [10] .

### 6.2 Exogenous variables

Specially in ex-ante forecasts, the values of some exogenous variables may be wrong. This problem can be empirically solved by adding to the model some stochastic equations, where the "suspicious" exogenous variables are endogenized and explained in terms of a "sure" exogenous component plus a random term with pre-assigned distribution. This method is rather simple, but, of course, involves a good degree of arbitrariness. Some theoretical works on this topic are available [24].

### 6.3 Random structural disturbances

It is exactly the problem of passing from the structural form; disturbances to the reduced form disturbances, discussed in the previous sections.

### 6.4 Errors in the coefficient estimates

Without any difference for linear and nonlinear models, this problem can be faced at least in 4 different ways, 3 of which are based on Monte Carlo methods, while one could be called "analytic simulation" (and can be even completely analytical if the model is linear, but the completely analytical method is not recommandable in practice, due to computational complexity also for small models). These four methods are briefly discussed below.

### 6.4.1. Monte Carlo on coefficients [10] , [23].

A vector of pseudo-random numbers with multivariate normal distribution can be added to the structural coefficients several times, and each time a new solution is computed. This method has some strong theoretical drawbacks also for linear models, as it should be clear from the following example.

$$C = a + bY + u$$
$$Y = C + I$$

Let the estimate of the coefficient b and of its standard error be, for example:

$\hat{b} = 0.8;\quad \hat{\sigma}_b = 0.2$ (so to have a "good" T = 4)

If we solve the model several times, after adding to 0.8 a random number with zero mean and standard deviation 0.2, we get meaningless results, since in the reduced form,

$$C = \frac{a}{1-b} + \frac{b}{1-b} A + \frac{u}{1-b}$$

the denominator 1-b can quite easily assume values close to zero (the truncation, which should be done to avoid it,should be too narrow, when compared with the nonlinear example of section 2). More exactly, if $\hat{b}$ has a normal distribution

$\frac{\hat{b}}{1-b}$ has a distribution with no finite moments, so that it is completely meaningless to compute sample means and sample variances of the results of simulation (in any case $\hat{b}$ has not a normal distribution, and the same holds, in general, for the coefficients of a structural model properly estimated by means of some method for simultaneous equations systems [8, p.182]).

### 6.4.2. Stochastic simulation and re-estimate [22]

This method is of quite difficult practical implementation and also has some theoretical drawbacks, but it gives interesting indications on the "small sample" properties of the estimation methods (and, therefore, of the forecasts produced by the model).

Vectors of pseudo structural disturbances are added to the model's equations and the model is solved over all the sample period. The computed values of the endogenous variables are treated "as if they were a new set of historical values" and a re-estimation of the structural coefficients is performed. The "new" set of estimated coefficients leads to a new forecast, and so on.

This method allows a good insight into the "small sample" properties of estimation and forecast errors. It has a strong practical drawback, as it is of difficult implementation even for small models. It also involves several theoretical problems (even if not so big as in case 6.4.1) since, under assumption of normality of the structural disturbances, the reduced form coefficients and the forecasts may have no finite moments [17],[21].

### 6.4.3. Monte Carlo to derive asymptotic variances

This method starts from the consideration that the estimation methods for simultaneous equations systems (2SLS, 3SLS, FIML, etc.) produce neither unbiased estimates of the coefficients, nor estimates of the variances of the coefficients, but just produce consistent and asymptotically normally distributed estimates of the coefficients. More exactly, with reference to the small two equations model, they produce $\hat{a}$, $\hat{b}$ and $\hat{\Sigma}/n$, such that [8,p. 199].

$$\operatorname*{p\,lim}_{n \to \infty} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \qquad \sqrt{n} \begin{bmatrix} \hat{a} - a \\ \hat{b} - b \end{bmatrix} \xrightarrow[n \to \infty]{\text{distrib.}} N\,(0,\Sigma)$$

A Monte Carlo method, which takes into account this starting assumption, is described in [5]. It seems (so I hope!) theoretically correct, but is not really more than an exercise.

### 6.4.4. Analytic simulation

This method takes properly into account the asymptotic properties of the estimation methods, as the algorithm of section 6.4.3., but has the advantage of avoiding the Monte Carlo sampling errors.

In case of linear models this method can be completely analytical [11]; however the completely analytical computation, even for a small model, is so complicated that computational errors have been quite frequent in the literature [4]. The mixed analytical and numerical method (therefore called "analytic simulation") is considerably simpler and faster and has the advantage of being applicable also to nonlinear models. It is described in details in [1].

It must be pointed out that the methods 6.4.1., 6.4.3. and 6.4.4. require, among the input data, an estimate of the complete asymptotic covariance matrix of all the structural estimated coefficients of the model. This is a standard outcome, for small models, of system estimation methods, such as 3SLS or FIML, but it is not produced by single equation methods, such as 2SLS or LIML; in the last cases, therefore, some additional preliminary computations must be performed.

## 7. The program installed at Bonn University

The program for stochastic and analytic simulation of econometric models, installed by the author of the Computer Center of Bonn University, is a modified version of the program announced in [2], whose installation and user's procedures are described in [3].

Data, programs and procedures are contained in 5 partitioned data sets on direct access device, under the code 'UJW411'.

UJW411.UJW.PISA.FORTRAN; this data set contains the source modules of all the programs and models (Model 5, Klein-I and Klein-Goldberger). All the programs are written in FORTRAN-G, excepted one ASSEMBLER subroutime (UNIFOR).

UJW411.UJW.PISA.LIBRARY; this data set contains the load modules of all the "standard" subroutimes of the program (which must be linked when no special option is specified); each member name is equal to the subroutine's entry name.

UJW411.UJW.PISA.DATA; this data set contains, in three members, the data for the three available models; the format of the data is described in [3].

UJW.411.UJW.PISA.TEXT; this data set contains the object modules of the main program, of the three models and of all the subroutines whose entry name is different from the member name, but is equal to a member name of the LIBRARY data set. The member name is the same of the corresponding source module in the FORTRAN data set.

UJW411.UJW.PISA.EXEC; this data set contains only one member, which is the execution procedure.

TSO execution command (with 512 K):
EX 'UJW411.UJW.PISA.EXEC(STOCSIM)'

4 parameters are required; the last 3 can be chosen among the following:

DUMMY; the standard stochastic simulation is performed, as described in [3];

INVERSE; the generation of the univariate standard normal deviates is performed by the inverse algorithm [7] rather than by the algorithm of Box and Muller.

NAGAR; the generation of the multivariate normal deviates is performed by the algorithm proposed by Nagar [19],[7], rather than by the algorithm of Mc Carthy [16]; Nagar's algorithm cannot be used for Model 5, since the number of stochastic equations is greater than the sample period length.

MCSERIAL; the generation of the multivariate normal deviates is performed by the algorithm of Mc Carthy [16] which takes into account serial correlation.

ANTIBIAS; the deterministic simulation bias is computed by means of the antithetic variate sampling technique.

REDVAR; the reduced form variances are computed by means of the analytic simulation procedure.

BONNGS; a special version of Gauss-Seidel is loaded; its execution speed is generally smaller than the standard, but the probabilities to get convergence are higher.

VERIFY; equations check with residuals is performed.

The first of the 4 parameters must be the name of the model:

MODEL5

KLEIN2 (is the Klein-I model estimated by 2SLS, as in [11]).

KLEINFML (is the Klein-I model estimated by FIML).

KLEINGOL (is the revised Klein-Goldberger model, estimated by 2SLS with 4 principal components, as in [1]).

When the execution begins, 5 integer numbers are printed at the terminal. With reference to [3] they are:

NREP : number of replications of stochastic simulation to be performed.

NPRINT : a submultiple of NREP for intermediate outputs corresponding to a number of replications less than or equal to NREP.

IFROM : initial year of simulation.

ITO : final year of simulation

IDYNAM: flag 0 if one-step simulation must be performed, 1 if the simulation must be dynamic.

These 5 numbers are read by the program at the beginning of the model's data set (a member of DATA) and can be either left unchanged by entering at the terminal a /*, or can be overwritten (with the same format).

The standard output is on the printer, but some results are, in some cases, also displayed at the terminal.

# R E F E R E N C E S


[ 1] Bianchi,C. and G. Calzolari, "The One-Period Forecast Errors in Nonlinear Econometric Models", International Economic Review, 20 (1979 or 1980, forthcoming).

[ 2] Bianchi, C., G. Calzolari and P. Corsi, "A Program for Stochastic Simulation of Econometric Models", Econometrica, 46 (1978), 235-236.

[ 3] Bianchi,C., G. Calzolari and P. Corsi, "Stochastic Simulation of Econometric Models:Installation's Procedures and User's Instructions", IBM Technical Report G513-3568, Pisa, (1978).

[ 4] Bianchi,C.,G. Calzolari and P.Corsi, "A Note on the Numerical Results by Goldberger, Nagar and Odeh", Econometrica 47 (1979), 505-506.

[ 5] Bianchi,C.,G. Calzolari and P. Corsi, "A Monte Carlo Approach to Compute the Asymptotic Standard Errors of Dynamic Multipliers", Economics Letters, (1979, forthcoming).

[ 5] Box, G.E.P., and M.E. Muller, "A Note on the Generation of Random Normal Deviates", Ann. Math.Stat., 29 (1958), 610-611.

[ 7] Calzolari,G., T.A. Ciriani and P. Corsi, "Generation and Testing of Pseudo-Random Numbers with Assigned Statistical Properties to be used in the Stochastic Simulation of Econometric Models", IBM Technical Report, CSP034/513-3544, Pisa, (1976).

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

[ 9] Dixon, W.T., Introduction to Statistical Analysis, New York: Mc. Graw-Hill, (1969).

[10] Fair, R.C., "Estimating the Expected Predictive Accuracy of Econometric Models", Cowles Foundation Discussion Paper n.480, (1978).

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

[12] Howrey,E.P. and H.H. Kelejian, "Simulation Versus Analytical Solutions: the Case of Econometric Models", in Computer Simulation Experiments with Models of Economic Systems, ed by T.H. Naylor, New York: John Wiley, (1971), 299-319.

[13] Howrey,E.P. and L.R. Klein,"Dynamic Properties of Nonlinear Econometric Models",International Economic Review, 13 (1972),599-618.

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

[15] Lewis,P.A.W., A.S. Goodman and J.M. Miller:"A Pseudo-Random Number Generator for the System/360", IBM Systems Journal,8 (1969),136-146

[16] McCarthy,M.D.: "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, Studies in Income and Wealth No.36, New York: National Bureau of Economic Research (1972), 185-191.

[17] McCarthy,M., "A Note on the Forecasting Properties of Two-Stage Least Squares Restricted Reduced Forms: The Finite Sample Case", International Economic Review, 13 (1972), 757-761.

[18] Moy,W.A., "Variance Reduction", in , Computer Simulation Experiments with Models of Economic Systems,ed by T.H. Naylor, New York: John Wiley, (1971), 269-289.

[19] Nagar,A.L., "Stochastic Simulation of the Brookings Econometric Model", in The Brookings Model: Some Further Results, ed. by T.S. Duesenberry, et al., Amsterdam: North-Holland,(1969).

[20] Quinke,H.,"Bonner Modell 5.5, Variablen,Gleichungen,Daten 1965-1977, Prognosen 1978-1985, Zusammenstellung" Institut für Gesellschafts-u. Wirtschaftswissenschaften der Universität Bonn, (1978).

[21] Sargan,J.D.,"The Existence of the Moments of Estimated Reduced Form Coefficients ",London School of Economics & Political Science, Discussion Paper, 46 (1976).

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

[23] Schmidt,P., "Some Small Sample Evidence on the Distribution of Dynamic Simulation Forecasts", Econometrica, 45 (1977), 997-1005.

[24] Schmidt,P., "A Note on Dynamic Simulation Forecasts and Stochastic Forecast-Period Exogenous Variables", Econometrica, 46 (1978), 1227-1230.

[25] Theil,H. and J.C.G. Boot, "The Final Form of Econometric Equation Systems", Review of the International Statistical Institute, 30 (1962), 136-152.