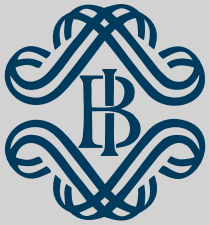
BANCA D'ITALIA

Temi di discussione

del Servizio Studi

Implementing Stochastic Optimal Control of Nonlinear Models: a Comparison with Alternative Solution Methods

by Andrea Cividini and Stefano Siviero



Numero 179 - Ottobre 1992

BANCA D'ITALIA

Temi di discussione

del Servizio Studi

Implementing Stochastic Optimal Control of Nonlinear Models: a Comparison with Alternative Solution Methods

by Andrea Cividini and Stefano Siviero

Numero 179 - Ottobre 1992

The purpose of the «Temi di discussione» series is to promote the circulation of working papers prepared within the Bank of Italy or presented in Bank seminars by outside economists with the aim of stimulating comments and suggestions.

The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank.

IMPLEMENTING STOCHASTIC OPTIMAL CONTROL OF NONLINEAR MODELS: A COMPARISON WITH ALTERNATIVE SOLUTION METHODS

by Andrea Cividini and Stefano Siviero (*)

Summary

Solving stochastic optimal control problems with a quadratic objective function for nonlinear models has attracted increasing attention in recent econometric literature. However, algorithms that implement the solution in its full form are very complex and require a large amount of computing resources. Consequently, optimal control problems are usually solved either by neglecting the stochastic nature of econometric models (i.e., by means of standard deterministic control), or by using algorithms that give approximate solutions.
This paper explores the full stochastic approach, presenting an operational iterative procedure that implements the full stochastic optimal control solution. This is done by using stochastic simulations to compute the multiplier matrix that maps changes of the control variables into changes of the objectives and of the variances of the objectives.

(*) Banca d'Italia, Research Department.

1. Introduction[1]

Optimal control techniques have received increasing attention in the economic literature, as economic policy design can be carried out by maximizing the policymaker's objective function subject to a set of constraints given by an estimated macroeconometric model.

Owing to the stochastic nature of models of this kind, optimal control techniques should be designed in such a way as to take this aspect into account. Only for linear models and quadratic cost functions does the certainty equivalence theorem apply, ensuring that the deterministic and stochastic solutions coincide.

Most econometric models are nonlinear, so that economic policy design should, in general, be carried out with stochastic optimal control techniques.

Algorithms of this kind usually work by iterating over linearizations of the entire model, using standard dynamic control theory to optimize the stochastic linearized model at each iteration: see for instance Chow (1976).

In view of the above-mentioned complexity of full stochastic solutions, optimal control problems are usually solved by neglecting the stochastic nature of the problem, i.e., by means of standard deterministic control.

Hall and Stephenson have recently (1990a, 1990b) proposed a simplified algorithm that transforms the original stochastic problem into a sequence of deterministic ones: an iterative procedure solves, at each iteration, a deterministic optimal control problem, where the objective function is modified to include the bias of the deterministic simulation, computed around the solution to the previous

---

1. Thanks for helpful suggestions go to Carlo Bianchi, participants in the Conference "Econometric Inference Using Simulation Techniques" (Rotterdam, June 1992), and an anonymous referee. The usual disclaimer applies.

deterministic problem. It is necessary to iterate because the bias is a function of the control variables.

In this paper we present an operational iterative procedure that implements the full stochastic optimal control solution. This is done by computing, by means of stochastic simulations, the multiplier matrices that locally map changes of the control variables into changes of the objectives and into changes of the variances of the objectives.

We then compare the properties of the three algorithms listed above (standard deterministic, Hall and Stephenson's, and our operational version of full stochastic).

We chose to carry out these comparisons with a simple two-equation dynamic nonlinear model having a closed-form solution.

This approach has two important consequences:

(i) We are able to find the exact solution of the full stochastic optimal control problem simply by minimizing the objective function subject to the closed-form of the nonlinear model rather than the nonlinear model itself. Analogously, we are able to find the exact solution of the Hall-Stephenson optimal control problem, which corresponds to the full stochastic algorithm being applied to an objective function where the variance of the objectives does not show up.

(ii) The operative versions of the two stochastic algorithms (i.e. those based on stochastic simulations) can then be compared with their respective exact solutions. This allows us to evaluate the consequences of increasing the number of replications.

The remainder of the paper is organized as follows: section 2 briefly describes the algorithms compared in this paper, and examines the approximation introduced by Hall and Stephenson. Section 3 discusses the exact solution to the problem, obtained by means of a blend of analytical and numerical techniques. Section 4 compares the performances of the operative versions of the stochastic algorithms.

Finally, three appendices discuss the following technical details: complete description of optimal control algorithms, closed-form solution of the model, empirical measurement of deterministic bias.

2. Alternative approaches to the solution of stochastic optimal control problems

In this paper we are concerned with the standard finite-horizon optimal control problem with a quadratic objective function and a set of constraints specified by a stochastic nonlinear model.

To simplify the exposition, this paragraph describes the simple one target - one instrument case; it is also assumed that the instrument can be moved at no cost. The general case (n targets, m instruments) is fully described in Appendix A.

The problem can thus be written as follows:

$$\min J^F = \min E \left\{ \sum_{t=1}^{T} \alpha_t \, (y_t - \bar{y}_t)^2 \right\} \qquad [1]$$

subject to the nonlinear model $\quad f_t(\mathbf{Y}_t, \mathbf{X}_t, x_t, \Theta, \varepsilon_t) = 0 \qquad [2]$

where y_t = objective variable, $y_t \in \mathbf{Y}_t$

$\mathbf{Y}_t$ = set of endogenous variables

$\bar{y}_t$ = target value for y_t

x_t = instrument variable

$\mathbf{X}_t$ = set of predetermined variables, other than x_t (including lagged endogenous variables)

ε_t = set of structural disturbances

Θ = set of parameters

α_t = non-negative weights, exogenously given

$[1, T]$ = time horizon of the optimization problem

Given that the set of constraints is specified as a nonlinear model, the Certainty Equivalence Theorem does not apply, and therefore deterministic optimal control is suboptimal with respect to stochastic techniques.

However, algorithms that implement the solution in its full form are very complex: the recent literature on this topic has therefore focused on finding approximated solution methods.

In this paper we give a detailed description of a procedure that implements the full stochastic solution and is still computationally tractable.

In this paragraph we highlight some basic relationships between our algorithm, the deterministic one, and the approximated algorithm recently proposed by Hall and Stephenson (1990a, 1990b). A detailed discussion of these three algorithms is given in Appendix A.

Making use of the well-known relationship

$$E(y_t^2) = (E(y_t))^2 + Var(y_t) \qquad [3]$$

the objective function can be written as:

$$J^F = \sum_{t=1}^{T} \alpha_t (E(y_t) - \bar{y}_t)^2 + \sum_{t=1}^{T} \alpha_t \, Var(y_t) \qquad [4]$$

The objective function can thus thought of as being made up of two parts:

- a quadratic component penalizing the deviations of the expected values of the objective variable from its target;
- a linear component penalizing the variability of the objective, i.e., the deviation of the variance from zero.

The full stochastic algorithm thus minimizes an objective function where the number of targets exceeds that of the instruments: hence, none of the targets will be exactly reached.

Without going into the details, it is worth noting that our full stochastic algorithm works by linearizing the relationships between the instruments and the expected value and the variance of the objectives.[2]

The algorithm proposed by Hall and Stephenson can be proved to give the solution of the optimal control problem where the objective function (4) is replaced with (5) below:

$$J^{HS} = \sum_{t=1}^{T} \alpha_t (E(y_t) - \bar{y}_t)^2 \qquad [5]$$

In order to prove this claim, we shall briefly describe Hall and Stephenson's algorithm.

They define d_t as the deterministic simulation bias, i.e.:

$$d_t = E(y_t) - \hat{y}_t \qquad [6]$$

where $\hat{y}_t$ is the deterministic model solution implicitly defined by $f_t(\hat{Y}_t, \hat{X}_t, x_t, \Theta, 0) = 0$.

Substituting eq. [6] into the original objective function (rewritten as in eq. [4]), they obtain:

$$J^F = \sum_{t=1}^{T} \alpha_t \left\{ \hat{y}_t^2 + d_t^2 + 2\, \hat{y}_t\, d_t + Var(y_t) + \bar{y}_t^2 + \right.$$

$$\left. - 2\, \bar{y}_t\, \hat{y}_t - 2\, \bar{y}_t\, d_t \right\} \qquad [7]$$

---

2. We do this by computing numerically (by means of stochastic simulations) the multiplier matrices mapping changes of x_t into changes of $E(y_t)$ and changes of $Var(y_t)$.
In the general case, with n instruments and T periods, a full stochastic approach requires nT+1 stochastic simulations at each iteration. This is so because one needs to compute the partial derivatives of all the objective variables and their variances with respect to all the instruments in each time period (see Appendix A for more details).

They then (1990b, p. 250) suggest "a solution algorithm which involves iterating between conventional [i.e., deterministic] optimal control exercises and stochastic simulation exercises"; their algorithm involves three steps:

(1) At each iteration i, apply conventional optimal control techniques to:

$$J = \sum_{t=1}^{T} \alpha_t \left\{ \hat{y}_t^2 + 2\, \hat{y}_t\, \hat{d}_t^{(i-1)} - 2\, \bar{y}_t\, \hat{y}_t \right\} + k \qquad [8]$$

where the deterministic simulation bias from previous iteration ($\hat{d}_t^{(i-1)}$, $d_t^0 = 0$) and the variance are treated as constants. The term k includes all the additive constants that obviously do not affect the result:

$$k = (\hat{d}^{(i-1)})^2 + \mathrm{Var}(y_t) + \bar{y}_t^2 - 2\, \bar{y}_t\, \hat{d}_t^{(i-1)}$$

(2) Use stochastic simulation techniques in order to re-estimate d_t around the solution obtained in step (1), that is, $x_t^{*(i)}$.

(3) The iterative procedure comes to an end when a fixed point is found between x_t^* and $\hat{d}_t$.

Clearly, the variance plays no role in such an iterative procedure: it only appears in the constant k, and therefore does not affect the result of the deterministic optimization of eq. [8].

Neglecting the variance is equivalent to switching from the original objective function to the following one:

$$J^{HS} = \sum_{t=1}^{T} \alpha_t \left[E(y_t) - \bar{y}_t \right]^2 \qquad [9]$$

To see this, substitute $d_t = E(y_t) - \hat{y}_t$ into eq. [9], to get an objective function that looks like eq. [8], except that the constant term is now different.

This approximation has the attractive feature that, unlike full stochastic algorithms, it requires only one stochastic simulation at each iteration: this is needed to compute $\hat{d}_t$.[3]

As the number of instruments and the number of time periods rise, full stochastic algorithms therefore become extremely expensive in terms of computing resources. By contrast, the Hall-Stephenson procedure is not affected and still requires only one stochastic simulation at each iteration.

There is clearly a tradeoff between saving computer resources and keeping the problem in its full form.

Hall and Stephenson are aware of the implicit modification they introduce in the original objective function, but feel that this is a reasonable approximation to make, at least when nonlinearity is mild.

When it is not, they suggest to evaluate a linear approximation of the variance of y_t, as a function of the instruments.

Given the nonlinearity of the model, this relationship cannot, in general, be assumed to be stable during the iterative process.

Therefore, this "extended" algorithm can be guaranteed to improve upon the basis one only if the relationship between the variance of y_t' and the instruments is linearized at each iteration: this can be done by performing step (2) above twice, with two different sets of stochastic simulations; the first one around $x^{*(i)}$, the second around

---

3. On the other hand, the number of stochastic simulations needed at each iteration of the full stochastic algorithm is a function of the number n of instruments and periods within the time horizon (nT+1).

$x^{*(i)} + \Delta x^{*}$.

However, in the general case of n instruments and T periods, while the first set consists of just one stochastic simulation, the second one involves nT simulations.

Consequently, nothing is saved in comparison with full stochastic algorithms, as there would not be any reduction in the number of stochastic simulations.[4]

3. An example based on a model having a closed-form solution

In order to compare the performances of the three solution methods briefly described above, we chose a simple two-equation dynamic nonlinear model that has been used extensively in the literature on nonlinear deterministic bias (see, e.g., Mariano and Brown, 1983 and 1988).

The structural form of the model is given by:

$$\begin{cases} \log y_t = \alpha_0 \log x_t + \alpha_1 \log y_{t-1} + u_t & [10] \\ z_t = \beta_0 x_t + \beta_1 y_t + v_t & [11] \end{cases}$$

$$(u_t, v_t)' \sim \text{IIN}(\underline{0}, \Sigma)$$

where we assume $\sigma_v^2 = \sigma_{uv} = 0$.

The closed-form solution and the analytical properties of the model (including the deterministic bias) are fully described in Appendix B.

The fact that the model has a closed-form solution allows us to compute, with a single deterministic optimal

---

4. It is in fact quite likely that a full stochastic algorithm would be more computationally efficient: at each step of the full stochastic algorithm, the derivatives of both the objective variables and their variances would suggest the direction towards the solution. The "extended" Hall-Stephenson algorithm would probably require a larger number of iterations as the information content of the bias is probably smaller than that of the derivatives of the objectives.

control exercise, the values of the instrument (x_t) which solve the stochastic optimal control problem. As we show below, this is done by modifying the model structure so that the expected value and the variance of the objective (z_t) appear explicitly as identities in the reduced form of the model.[5]

In the discussion below we present the detailed results relative to the model generated with the following set of parameter values:

$$\alpha_0 = .8,\ \alpha_1 = .2,\ \beta_0 = 1,\ \beta_1 = .1,\ \sigma_u^2 = .01,\ y_0 = 1000,$$

$x_1 = 1000$, x is growing at .5 per cent per period.

The one step ahead weighted bias is then:

$$\Psi^d_{z_t} = .0025$$

This bias (around .2 per cent of the total variance of z_t) is very small, but nevertheless sufficient to produce significant differences between Hall and Stephenson's solution and the full stochastic solution.

The optimal control problem is defined by:

$$\min_{x_t} J^F = \min_{x_t} \sum_{t=81}^{100} E(\bar{z}_t - z_t)^2 \qquad [12]$$

subject to the model above, where $\bar{z}_t$ is growing at 1 per cent per period (against .5 per cent in the baseline).[6]

---

5. In order to evaluate the sensitivity of the results with respect to the parameter values, we chose to generate 27 different data sets, each consisting of 100 observations; see Appendix B for details. In order to reduce the influence of the initial conditions, the optimization experiments were performed over the last 20 periods.

6. It is worth remarking that the model is homogeneous and that in the steady state all the variables grow at the same rate as x.

The exact solution of the stochastic optimal control problem defined by [12]-[10]-[11] is equivalent to the solution of the deterministic optimal control problem defined by [12] and [13]-[14]-[15]-[16] below:

$$\begin{cases} \xi_t = \alpha_0 \log x_t + \alpha_1 \xi_{t-1} & [13] \\ Ey_t = \exp\left(\xi_t + \alpha_1^t \log y_0 + \frac{\sigma_u^2}{2} \frac{1-\alpha^{2t}}{1-\alpha^2}\right) & [14] \\ Ez_t = \beta_0 x_t + \beta_1 Ey_t & [15] \\ Vz_t = \beta_1^2 \exp\left(2\xi_t + 2\alpha_1^t \log y_0\right) \cdot \\ \quad \cdot \exp\left(\sigma_u^2 \frac{1-\alpha^{2t}}{1-\alpha^2}\right) \left[\exp\left(\sigma_u^2 \frac{1-\alpha^{2t}}{1-\alpha^2}\right) - 1\right] & [16] \end{cases}$$

where $\xi_0 = 0$.

Ey_t, Ez_t and Vz_t, which are treated here as ordinary endogenous variables, are in fact the closed-form expressions of $E(y_t)$, $E(z_t)$, $E(z_t - E(z_t))^2$, given the relevant information available at $t = 0$, i.e. y_0.

In order to solve this problem, it is necessary to compute the numerical derivatives of Ez_t and Vz_t with respect to the instrument x_t, as should be clear after rewriting J^F as follows:

$$J^F = \sum_{t=81}^{100} (E(z_t) - \bar{z}_t)^2 + \sum_{t=81}^{100} E(z_t - E(z_t))^2$$

$$= \sum_{t=81}^{100} (Ez_t - \bar{z}_t)^2 + \sum_{t=81}^{100} (Vz_t - 0) \qquad [17]$$

Thus, while the Hall-Stephenson algorithm solves a one target-one instrument problem (and hence reaches exactly the target), the full stochastic algorithm minimizes an objective

function that depends on two targets (a target for z_t and one for its variance): hence, neither will be reached.

We compare the solutions of the following three problems:

(1) Minimize

$$J^D = \sum_{t=81}^{100} (z_t - \bar{z}_t)^2$$

subject to the nonlinear model defined by the deterministic part of [10] and [11].

The solution of this **deterministic optimal control** exercise is $\{x_t^D\}_{t=81}^{100}$.

(2) Minimize

$$J^{HS} = \sum_{t=81}^{100} (Ez_t - \bar{z}_t)^2$$

subject to [13]-[14]-[15].

This gives **Hall and Stephenson's solution** of the stochastic optimal control exercise, $\{x_t^{HS}\}_{t=81}^{100}$.

(3) Minimize

$$J^F = \sum_{t=81}^{100} (Ez_t - \bar{z}_t)^2 + \sum_{t=81}^{100} Vz_t$$

subject to [13]-[14]-[15]-[16].

This gives the **full stochastic solution** of the optimal control problem, $\{x_t^F\}_{t=81}^{100}$.

We then evaluate our results by computing the values of the three objective functions J^D, J^{HS} and J^F evaluated around each of the three optimal paths.

Obviously, given that both problem (1) and problem (2) are of the one target-one instrument type, the objective functions J^D and J^{HS}, evaluated at the optimum, are zero. This is not the case for J^F.

The most relevant comparison is, of course, the one obtained by feeding the values of the optimal paths $\{x_t^D\}$,

$\{x_t^{HS}\}$ and $\{x_t^F\}$ into J^F (first column of Table 1a).

The improvement with respect to the deterministic solution produced by the approximation suggested by Hall and Stephenson is the difference between J^F evaluated at $\{x_t^D\}$ and $\{x_t^{HS}\}$: this difference is slightly larger than 4,000, that is, around .8 per cent of the value of the objective function. This improvement can be attributed in part to the fact that z_t is closer to $\bar{z}_t$ (which induces a reduction of about 1,500 in the value of the objective function) and in part to the reduction in the variance of z_t (about 2,500).

Moving from Hall and Stephenson's solution to the full stochastic solution, there is a further reduction of more than 1,000 in the value of the objective function, i.e., about .2 per cent.

Summing up, using a solution method that takes account of the stochastic nature of the problem (either Hall-Stephenson or full stochastic), a significant reduction of the objective function is obtained in comparison with the deterministic solution, even when the degree of nonlinearity is, by usual standards, quite low (see above).

Moreover, the marginal improvement produced by moving from Hall-Stephenson to full stochastic is far from negligible, as it amounts to about 24 per cent of the total improvement (row 3 - row 1 in first column of Table 1a).

Figures 1a, 1b and 1c show the paths of the instrument x_t, of the objective variable Ez_t and of the variance of the objective Vz_t for the deterministic and the exact Hall-Stephenson procedures, as deviations with respect to the exact full stochastic solutions.

It can be seen that the deterministic solution is suboptimal because the instrument is raised excessively: this is so because the deterministic solution does not take account of the fact that $z_t < E(z_t)$.

The difference between the deterministic and the full stochastic solutions amounts to about 0.5 per cent of the

latter. Using the approximated Hall-Stephenson algorithm, the difference is reduced by about half, as can be seen in Figure 1a.

When evaluated on the objectives and the variances of the objectives, the relative improvement produced by using the Hall-Stephenson algorithm is still half the total improvement produced by moving from the deterministic to the full stochastic algorithm; see Figures 1b and 1c.

Finally, it is worth noting that the finite-horizon characteristic of these experiments implies a mild instrument instability in the last 2 or 3 periods: this is obviously so only in the full stochastic case, where the number of targets exceeds that of the instruments. This effect tends to reduce the difference between the deterministic and Hall-Stephenson solutions and the full stochastic one towards the end of the time horizon. This is again clear from Figures 1a, 1b, 1c.

If the variance plays a role in the loss function, then it is possible to solve the optimal control problem under different assumptions regarding the degree of risk aversion of the policy maker, as suggested by Mitchell (1979).

Mitchell suggests to give different weights to the first and second component of the loss function in eq. [17], i.e., to solve the following problem:

$$\min J^M = \lambda \sum_{t=81}^{100} (E(z_t)-\bar{z}_t)^2 + \sum_{t=81}^{100} E(z_t-E(z_t))^2 \qquad [18]$$

subject to [13]-[14]-[15]-[16]

where λ is the risk aversion parameter;[7] "in general, the higher is λ the less the policy-marker is concerned with the predictability (variance) of the target" (Mitchell, 1979, p. 914).

---

7. Mitchell gives examples, taken from the economic literature, for both of the two extreme cases $\lambda \to \infty$ and $\lambda=0$. Notice that, with $\lambda \to \infty$, we are back to Hall and Stephenson's problem.

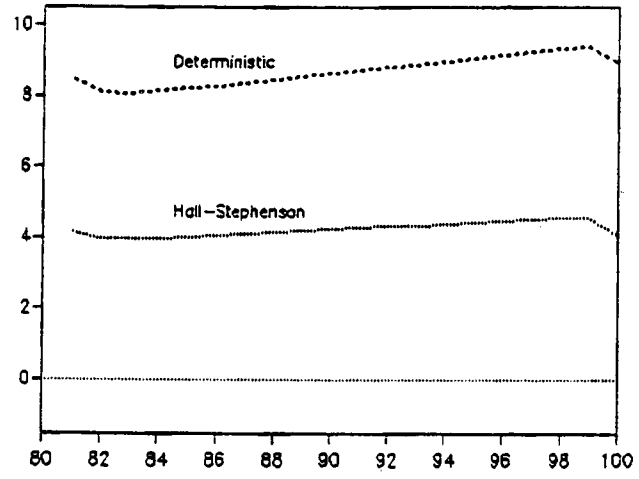
Table 1a: Comparison of the performance of three different solution methods: deterministic, Hall-Stephenson, and full stochastic optimal control

Objective function / Solution algorithm	$J^F = \sum_{t=81}^{100} E(z_t-\bar{z}_t)^2 = \sum_{t=81}^{100} [(E(z_t)-\bar{z}_t]^2+Var(z_t)$	$J^{HS} = \sum_{t=81}^{100} [E(z_t)-\bar{z}_t]^2$	$J^D = \sum_{t=81}^{100} (z_t-\bar{z}_t)^2$
Full stochastic (x^F)	551,376	1,283	5,392
Hall-Stephenson (x^{HS})	552,662	0	1,422
Deterministic (x^D)	556,807	1,429	0

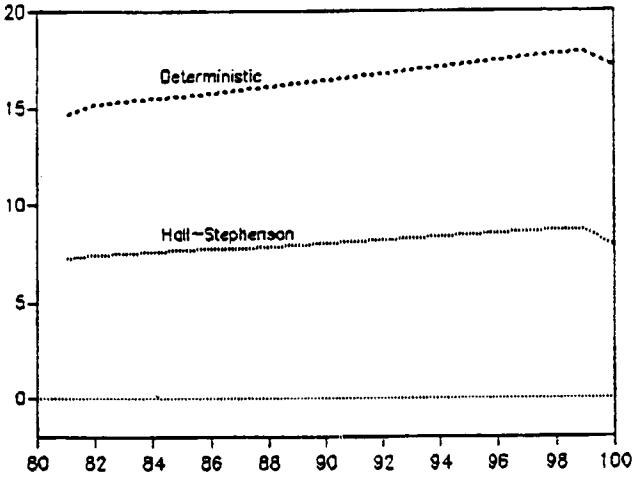
Table 1b: Comparison of the performance of Hall-Stephenson and full stochastic optimal control: consequences of increasing the number of replications (all performances evaluated with J^F).

Number of antithetic replications / Solution algorithm	100	1,000	10,000
Full stochastic (x^F)	551,469	551,385	551,378
Hall-Stephenson (x^{HS})	552,818	552,680	552,668

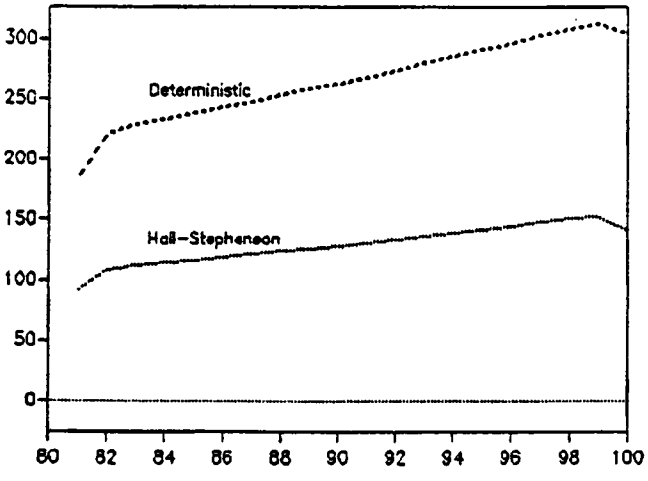
Figure 1: Instrument, objective and variance for the deterministic and exact Hall-Stephenson procedures (Deviations with respect to the exact full stochastic solution)



(a)



(b)



(c)

Clearly one can allow for a varying degree of risk aversion only if the variance term is not dropped from the loss function.

Since the economic theory is very often concerned with the effects of risk aversion in the policy-marker's objective function, it seems interesting to keep the possibility of solving optimal control problems with different values of λ.

We solved the problem above under a set of different assumptions concerning the parameter λ (λ = .01, .05, .1, .5, 1, 5, 10). In table 2a we compare the results obtained applying full stochastic optimal control to these problems, with the value of the objective function J^M evaluated along the Hall-Stephenson and deterministic solutions.

Quite obviously, the differences between the full stochastic and the Hall-Stephenson methods decrease as λ increases. On the other hand, for λ smaller than unity, the Hall-Stephenson solution can be considerably outperformed by the full stochastic one. For instance, with λ=.1 the value of the loss function evaluated around the Hall-Stephenson solution is more than 2 per cent higher than that obtained with the full stochastic solution (as a term of comparison, it is slightly more than .2 per cent with λ=1).

It can be seen that the advantage of using Hall and Stephenson's algorithm rather than standard deterministic optimal control vanishes as λ approaches zero: the loss of performance with either of these two methods is of about the same order of magnitude.

Table 2b compares the detailed results for the case λ=.1: while, with λ=1, the relative improvement in performance produced by using the Hall-Stephenson algorithm was about half the total improvement produced by moving from the deterministic to the full stochastic algorithm, it is now much less than that (it is in fact about 1/10).

Table 2a: Evaluation of the objective function J^M with different values for the risk aversion parameter λ (percentage deviation from the full stochastic result in parentheses)

Solution method \ λ	.01	.05	.1	.5	1	5	10
Full stochastic (x^F)	448,094	528,035	540,069	550,096	551,376	552,404	552,533
Hall-Stephenson (x^{HS})	552,662 (23.3)	552,662 (4.7)	552,662 (2.3)	552,662 (0.5)	552,662 (0.2)	552,662 (0.05)	552,662 (0.02)
Deterministic (x^D)	555,392 (23.9)	555,449 (5.2)	555,520 (2.9)	556,092 (1.1)	556,807 (1.0)	562,522 (1.8)	569,667 (3.1)

Table 2b: Comparison of deterministic, Hall-Stephenson and full-stochastic solutions with λ=.1
(Hall and Stephenson and full-stochastic solutions are computed as percentage deviations from the deterministic solution).

Period	Optimal path for x_t			Optimal path for Ez_t			Optimal path for Vz_t		
	Determ.	Hall and Steph.	Full-stoch. with λ=.1	Determ.	Hall and Steph.	Full-stoch. with λ=.1	Determ.	Hall and Steph.	Full-stoch. with λ=.1
81	1621	-0.26	-2.78	3114	-0.24	-2.52	22390	-0.42	-4.41
82	1652	-0.25	-2.59	3145	-0.25	-2.55	23331	-0.48	-4.97
83	1670	-0.25	-2.55	3177	-0.25	-2.55	23786	-0.49	-5.03
84	1687	-0.25	-2.55	3209	-0.25	-2.55	24259	-0.49	-5.03
85	1703	-0.25	-2.55	3241	-0.25	-2.55	24746	-0.49	-5.03
86	1721	-0.25	-2.55	3273	-0.25	-2.55	25244	-0.49	-5.03
87	1738	-0.25	-2.55	3306	-0.25	-2.55	25751	-0.49	-5.03
88	1755	-0.25	-2.55	3339	-0.25	-2.55	26269	-0.49	-5.03
89	1773	-0.25	-2.55	3372	-0.25	-2.55	26797	-0.49	-5.03
90	1790	-0.25	-2.55	3406	-0.25	-2.55	27335	-0.49	-5.03
91	1808	-0.25	-2.55	3440	-0.25	-2.55	27885	-0.49	-5.03
92	1826	-0.25	-2.55	3475	-0.25	-2.55	28445	-0.49	-5.03
93	1845	-0.25	-2.55	3509	-0.25	-2.55	29017	-0.49	-5.03
94	1863	-0.25	-2.55	3544	-0.25	-2.55	29600	-0.49	-5.03
95	1882	-0.25	-2.55	3580	-0.25	-2.55	30195	-0.49	-5.03
96	1901	-0.25	-2.55	3616	-0.25	-2.55	30802	-0.49	-5.03
97	1920	-0.25	-2.55	3652	-0.25	-2.55	31421	-0.49	-5.03
98	1939	-0.25	-2.54	3688	-0.25	-2.55	32053	-0.49	-5.03
99	1958	-0.25	-2.52	3725	-0.25	-2.52	32697	-0.49	-4.98
100	1978	-0.25	-2.26	3762	-0.25	-2.28	33354	-0.49	-4.56

4. Operative solutions based on stochastic simulation techniques

While the comparisons presented in the previous paragraph are interesting per se, their relevance is limited by the fact that they are only possible when a model has a closed-form solution. Since this is not generally the case for nonlinear econometric models, we now compare the performances of an operative version of the procedures discussed above.

By operative versions we mean algorithms that are based on stochastic simulation techniques and can be used for any kind of econometric model.

Appendix C describes the details of the experiments of stochastic simulation that we need in this paragraph, including the particular Monte Carlo technique which we use.

Appendix C also stresses the need to take proper care in computing the deterministic simulation bias: the fact that this is a key factor in determining the properties of the results is clearly shown by the functioning of the Hall-Stephenson procedure.

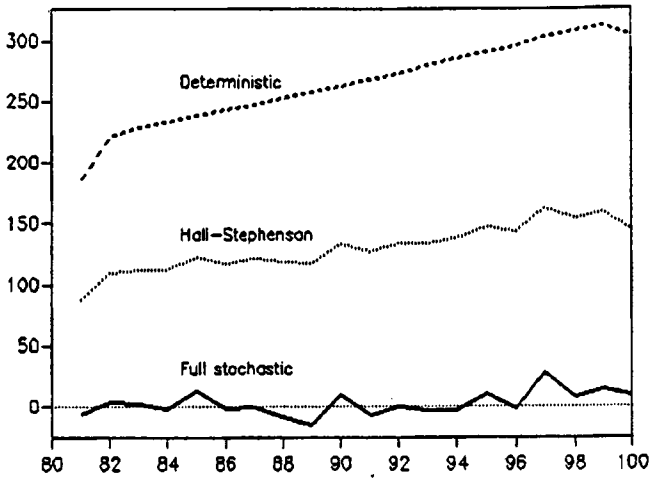
We have computed the deterministic bias with different number of replications (100, 1,000 and 10,000). We also explored the consequences of introducing methods of variance-reduction: we found that they can greatly improve the precision of the estimates of the bias, as we discuss in Appendix C.

All in all, in our experiments the choice of 1,000 antithetic replications appears to be a reasonable compromise between precision and computing costs.

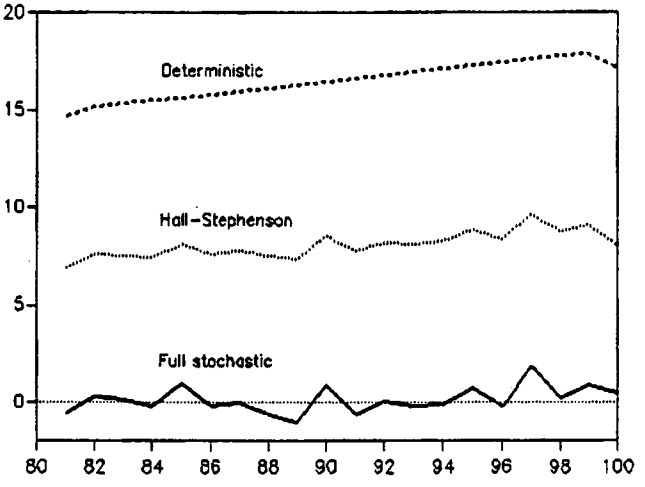
The results discussed below are based on the optimization exercises of Section 3: we are thus able to make a comparison of the operative versions of the algorithms and to compare them with the exact ones.

Figures 2a, 2b and 2c show respectively the behaviour of the instrument, the objective and the variance of the

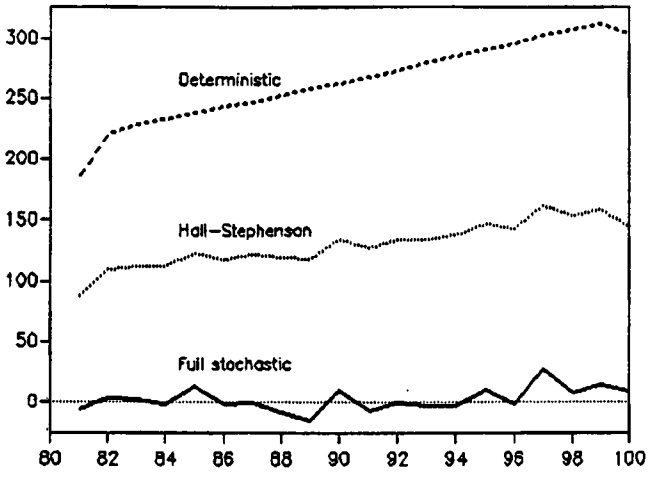
Figure 2: Instrument, objective and variance for the three solution procedures (1,000 replications, deviations with respect to the exact full stochastic solution)



(a)

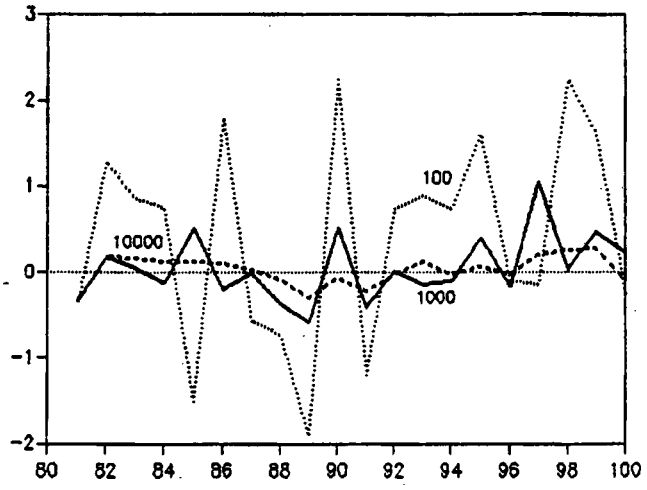


(b)

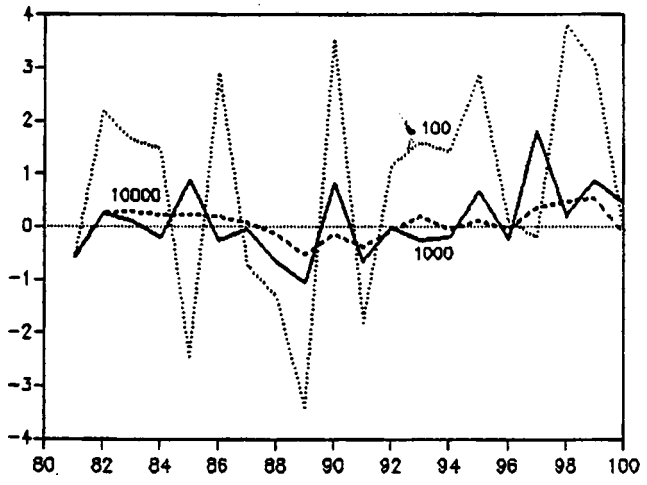


(c)

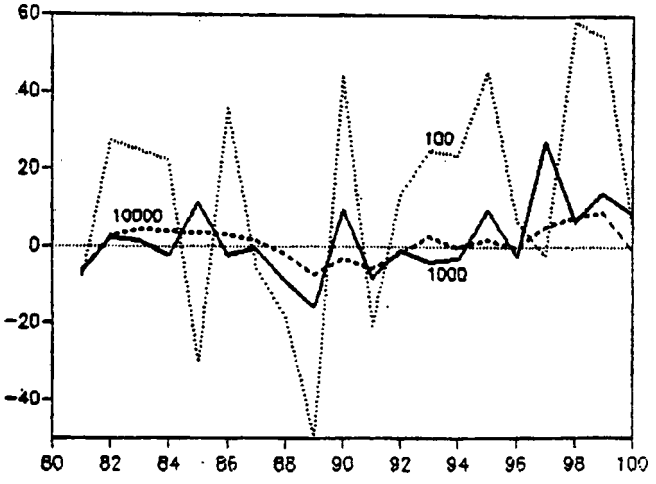
Figure 3: Instrument, objective and variance for the full stochastic procedure: effects of increasing the number of replications (Deviations with respect to the exact full stochastic solution)



(a)



(b)



(c)

objectives for the deterministic, Hall-Stephenson and full stochastic procedures (the last two computed with 1,000 antithetic replications), as deviations with respect to the true full stochastic values.

Comparing these figures with figures 1a, 1b and 1c, we see that the operative versions give results that are reasonably close to those of the exact algorithms. The largest deviations are those of the variance of the objectives (Figure 1c).

The relative performances of the three algorithms are basically the same as before, with the Hall-Stephenson procedure catching up roughly half of the total difference between deterministic and full stochastic methods.

The central column of Table 1b shows the value of the objective function (J^F) evaluated around the full stochastic and Hall-Stephenson solutions shown in the figures: it basically confirms the findings of the previous chapter.

The table also shows the effect of changing the number of replications: one can see that not much is gained by moving from 1,000 to 10,000 replications (of course, much is lost in terms of computer time). This confirms that the choice of 1,000 antithetic replications, justified in Appendix C, is quite reasonable in this case.

Figures 3a, 3b and 3c show, for the full stochastic case only, how the objectives and the variances are affected by the increase in the number of replications. It is clear that the main effect is to reduce their variability.

Summing up, the findings of the previous paragraph are still valid when referred to the operative algorithms.

5. Concluding remarks

Given our choice of a simple two-equation nonlinear form having an analytical solution, we have been able to compare three algorithms for the solution of optimal control problems

(deterministic, Hall-Stephenson, full stochastic) in both their exact and their operative versions.

This has allowed us both to verify the properties of these algorithms per se and to make a direct evaluation of the effects of using stochastic simulation techniques, which are the only viable ones for the most general nonlinear case.

Our comparisons highlight the following considerations:

- as one would expect, the loss due to the use of deterministic optimal control techniques is not relevant when the degree of nonlinearity (as measured by the bias) is very low;
- before choosing the most appropriate algorithm, a preliminary analysis of the bias is advisable;
- in order to obtain a good estimate of the bias, a large number of replications is necessary. We show that using variance-reduction techniques can considerably improve the estimate of the bias (it is worth recalling again that we can compare estimates of the bias with its true value, computed with the closed-form solution of the model);
- even with a model where nonlinearity is mild, the exact Hall-Stephenson procedure produces a significant improvement over the deterministic solution. As far as the operative version is concerned, the estimate of the bias becomes a crucial issue;
- however, the basic version of this algorithm neglects the role of the variance, and therefore implicitly modifies the objective function. If adjusted in order to overcome this limit (along the lines proposed in Hall and Stephenson, 1990a and 1990b), it becomes at least as costly (in terms of computing resources) as a full stochastic algorithm;
- our version of the full stochastic algorithm involves a further reduction in the cost function, which is of the same order of magnitude as the one obtained when moving from the deterministic to the Hall-Stephenson algorithm. That is, neglecting the variance has a significant cost even when the degree of nonlinearity is mild;

- finally, neglecting the role of the variance puts heavy constraints on the kind of loss function that one can postulate. In particular, if the variance term is not explicitly included in the loss function, one cannot allow for any of a range of degrees of risk aversion, as suggested by Mitchell (1979). We show that the introduction of an arbitrary amount of aversion to uncertainty in the loss function can make non trivial differences in the optimal solution.

Appendix A: Description of the optimization algorithms

Let us consider a dynamic nonlinear econometric model in structural form:

$$f_t(Y_t, X_t, x_t, \Theta, \varepsilon_t)=0 \qquad t = 1,2,\ldots,T \qquad (A.1)$$

where the symbols are defined as in eq. [2].

Given a set of m control variables (or instruments) with values x_{jt} $j=1,\ldots,m$, $t=1,\ldots,T$ and a set of n objectives with values y_{it} $i=1,\ldots,n$, $t=1,\ldots,T$ the problem can be stated as:

$$\min_{x_{jt}} \left\{ E\left[\sum_{i=1}^{n} \sum_{t=1}^{T} h_{it}(y_{it}-\bar{y}_{it})^2 \right] + \sum_{j=1}^{m} \sum_{t=1}^{T} k_{jt}(x_{jt}-\bar{x}_{jt})^2 \right\} \qquad (A.2)$$

subject to

- the structural nonlinear model (A.1);
- linear intertemporal equality and inequality constraints on control variables and/or objectives;

where

$\bar{y}_{it}$ are the desired values (targets) for the objectives

$\bar{x}_{it}$ are the desired values (targets) for the control variables

h_{it}, k_{jt} are non-negative weights, exogenously given.

The optimization algorithms presented here treat the problem as static by stacking all the periods of the time horizon. In this way we obtain the following vectors of control variables, objectives and targets over the entire time horizon:

$$\begin{aligned}
y &= [y_{11}\ y_{21}\ \cdots\ y_{n1}\ y_{12}\ y_{22}\ \cdots\ y_{n2}\ \cdots\ y_{1T}\ y_{2T}\ \cdots\ y_{nT}]' \\
x &= [x_{11}\ x_{21}\ \cdots\ x_{m1}\ x_{12}\ x_{22}\ \cdots\ x_{m2}\ \cdots\ x_{1T}\ x_{2T}\ \cdots\ x_{mT}]' \\
\bar{y} &= [\bar{y}_{11}\ \bar{y}_{21}\ \cdots\ \bar{y}_{n1}\ \bar{y}_{12}\ \bar{y}_{22}\ \cdots\ \bar{y}_{n2}\ \cdots\ \bar{y}_{1T}\ \bar{y}_{2T}\ \cdots\ \bar{y}_{nT}]' \\
\bar{x} &= [\bar{x}_{11}\ \bar{x}_{21}\ \cdots\ \bar{x}_{m1}\ \bar{x}_{12}\ \bar{x}_{22}\ \cdots\ \bar{x}_{m2}\ \cdots\ \bar{x}_{1T}\ \bar{x}_{2T}\ \cdots\ \bar{x}_{mT}]'
\end{aligned}$$

and the problem can be written as:

$$\min_{x} E\left\{ \left[y'-\bar{y}', x'-\bar{x}' \right] \begin{bmatrix} H_{11} & 0 \\ 0 & H_{22} \end{bmatrix} \begin{bmatrix} y-\bar{y} \\ x-\bar{x} \end{bmatrix} \right\} \qquad (A.3)$$

subject to

- the structural nonlinear model (A.1)
- the linear constraints specified explicitly in the form

$$B^l \leq C_1 y + C_2 x \leq B^u \qquad (A.4)$$

where

H_{11} is a (nT,nT) diagonal matrix with elements h_{it} $i=1,\ldots,n$, $t=1,\ldots,T$;

H_{22} is a (mT,mT) diagonal matrix with elements k_{jt} $j=1,\ldots,m$, $t=1,\ldots,T$;

B^l and B^u are p vectors of lower and upper bounds

C_1 is a (p,nT) matrix of coefficients for the constraints on objectives;

C_2 is a (p,mT) matrix of coefficients for the constraints on control variables.

The basic idea (see Holbrook, 1974, and Brandsma, Hughes Hallet, Van der Windt, 1983) is to use an iterative procedure approximating the model by the (nT,mT) multiplier matrix that maps changes in the control variables into changes of the objectives (see Cividini, 1990).

Our algorithm for full stochastic optimal control extends this idea to the stochastic context: at each iteration the multiplier matrix evaluation is based on the numerical derivatives of the expected values of the objectives with respect to the control variables computed via stochastic simulation. As already discussed in the text, full stochastic optimal control implies also computing the multiplier matrix based on the numerical derivatives of the variances of the objectives with respect to the control variables.

The $[(t_1-1)n+i,(t_2-1)m+j]$ element of the multiplier matrix for the expected values of the control variables is approximated by:

$$\frac{\partial E(y_{it_1})}{\partial x_{jt_2}} \simeq \frac{\frac{1}{N}\sum_{r=1}^{N} y^{d}_{it_1,r} - \frac{1}{N}\sum_{r=1}^{N} y^{c}_{it_1,r}}{x^{d}_{jt_2} - x^{c}_{jt_2}} \qquad \begin{array}{l} t_1,t_2=1,\ldots,T \\ t_1 \geq t_2 \\ i=1,\ldots,n \\ j=1,\ldots,m \end{array} \qquad (A.5)$$

where

N is the number of replications produced via stochastic simulation;

$y^c_{it_1}$, . are the control solutions for y_i at time t_1 obtained via stochastic simulation using the exogenous variables as they are;

$y^d_{it_1}$, . are the shocked solutions for y_i at time t_1 obtained via stochastic simulation increasing the exogenous variable x_j at time t_2 by a finite increment $\Delta x^c_{jt_2}$:

$$x^d_{jt_2} = x^c_{jt_2} + \Delta x^c_{jt_2}, \quad \Delta x^c_{jt_2} = \begin{cases} 10^{-5}\, x^c_{jt_2} & \text{if } |x^c_{jt_2}| > 10^{-5} \\ 10^{-5} & \text{if } |x^c_{jt_2}| \leq 10^{-5} \end{cases}$$

The multiplier matrix for the variances of the control variables is computed in an analogous fashion, using the same set of replications.

To improve the accuracy of the evaluation of the expected values of the objectives in (A.5), the antithetic variates method proposed by Calzolari (1979) has been used. This solves the model with pseudostructural disturbances of opposite sign and therefore requires two simulations for each replication of the stochastic simulation (for more details see Appendix C).

The number of dynamic simulations required to compute the elements of each column of the multiplier matrix is thus 4N (2N control solutions and 2N shocked solutions); the total number of simulations is therefore 2N (mT+1).

Replacing E(y) with Ax (where A is the (nT,mT) multiplier matrix computed as in (A.5)), and Var(y) with Vx (where V is the (nT,mT) multiplier matrix) in the objective function, we obtain after some algebraic manipulations the following expression:

$$\min_x \left(\frac{1}{2} x'Qx + Lx \right) \tag{A.6}$$

$$\text{s.t. } B^l \leq Dx \leq B^u$$

where

$$Q = 2\,(A'H_{11}\,A + H_{22}) \tag{A.7}$$

$$L = -2\bar{y}'H_{11}\,A - 2\bar{x}'H_{22} + H_{11}\text{vec}V \tag{A.8}$$

$$D = C_1\,A + C_2 \tag{A.9}$$

The matrices Q, L, D, B^l, B^u are directly supplied to a quadratic programming algorithm.[8]

Because of the nonlinearities of the model the solution requires an iterative procedure that consists of the following steps:

a) shock the structural model for each control variable and each time period, solve the model for each shock via stochastic simulation and compute the multiplier matrices $A^{(k)}$ and $V^{(k)}$;
b) compute the matrices $Q^{(k)}$, $L^{(k)}$, $D^{(k)}$ in equations (A.7), (A.8), (A.9);
c) solve the Quadratic Programming problem (A.6) to find the optimal values of the control variables;
d) evaluate the results; if the difference of the control variables from their values in the previous iteration does not satisfy a given convergence criterion go back to a) with the new values of the control variables, otherwise the solution procedure is over.

It is worth pointing out that the proposed algorithm does not introduce approximations in the full stochastic optimal control problem and is still computationally tractable even for large-size econometric models using the simulation package produced at the Bank of Italy.[9]

Even if tractable the problem is still computationally formidable. A simplification of the algorithm, proposed by Hall and Stephenson (1990), can be used to reduce the number of stochastic simulations and therefore the CPU time.

The objective function of problem (A.2) can be stated as:

$$E(J) = \sum_{i=1}^{n} \sum_{t=1}^{T} h_{it} \left\{ \left[E(y_{it}) \right]^2 + \mathrm{Var}\left[y_{it} \right] + \bar{y}_{it}^2 - 2\bar{y}_{it} E\left[y_{it} \right] \right\}$$

---

8. The routine used in this work is the E04NAF routine of the NAG Fortran library designed to solve the Quadratic Programming problem, which is assumed to be as stated in (A.6).

9. This package is a revised version of Modeleasy (Federal Reserve System, 1986) implemented as VS Fortran functions, designed as part of the Speakeasy software (Speakeasy Corp., Chicago, Illinois). Each equation is interpreted as a vector relation whose elements correspond to different datasets that are processed concurrently on a vector machine, thereby reducing the computational costs. A stochastic simulation can be viewed as a single vector simulation. For more details see Cividini and Petersen (1987) and (1989).

$$+ \sum_{j=1}^{m} \sum_{t=1}^{T} k_{jt} \left(x_{jt} - \bar{x}_{jt}\right) \qquad \text{(A.10)}$$

Let us now define:

$$E(y_{it}) = \hat{y}_{it} + d_{it}$$

where

$\hat{y}_{it}$ is the deterministic model solution (i.e., the solution obtained setting the error terms to zero);
d_{it} is the bias of the deterministic solution from the mean value.

The objective function can thus be written as:

$$E(J) = \sum_{i=1}^{n} \sum_{t=1}^{T} h_{it} \left\{ \hat{y}_{it}^2 + 2\left(d_{it} - \bar{y}_{it}\right)\hat{y}_{it} + Var\left(y_{it}\right) + d_{it}^2 + \right.$$

$$\left. \bar{y}_{it}^2 - 2\bar{y}_{it}\, d_{it} \right\} + \sum_{j=1}^{m} \sum_{t=1}^{T} k_{jt} \left(x_{jt} - \bar{x}_{jt}\right)^2 \qquad \text{(A.11)}$$

The idea of Hall and Stephenson is to consider an iterative solution procedure where the terms d_{it} and $Var(y_{it})$ are constant within each iteration even if they are functions of the control variables. However, the terms d_{it} are recomputed at each iteration, and modify the linear term of the objective function as written below (A.14); on the other hand, the variance is also implicitly re-computed, but enters only in the constant term (not shown below) and thus does not affect the results.

Using this approximation and treating the problem as before, the optimal control problem can be formulated as follows:

$$\min_{x} \ \frac{1}{2} x'Qx + Lx \qquad \text{(A.12)}$$

$$\text{s.t. } B^l \leq Dx \leq B^u$$

where

$$Q = 2(A'H_{11}\, A + H_{22}) \qquad \text{(A.13)}$$

$$L = 2(d - \bar{y})'H_{11}\, A - 2\bar{x}'H_{22} \qquad \text{(A.14)}$$

$$D = C_1 A + C_2 \tag{A.15}$$

$$d = [d_{11}\ d_{21}\cdots d_{n1}\ d_{12}\ d_{22}\cdots d_{n2}\cdots d_{1T}\ d_{2T}\cdots d_{nT}] \tag{A.16}$$

and A is the (nT,mT) multiplier matrix based on the numerical derivatives of the objectives with respect to the control variables, computed via deterministic (rather than stochastic) simulation.

It is worth pointing out that, in this formulation, only one stochastic simulation is required to compute the vector of biases d at each iteration.[10]

The iterative procedure consists of the following steps:

a) shock the structural model for each control variable and each time period, solve the model for each shock via deterministic simulation and compute the multiplier matrix $A^{(k)}$;
b) compute the vector of biases $d^{(k)}$ as the difference between the results of the deterministic simulation and the mean of the stochastic simulation; in the first iteration set d to zero;
c) compute the matrices $Q^{(k)}, L^{(k)}, D^{(k)}$ as in equations (A.13), (A.14), (A.15);
d) solve the Quadratic Programming problem (A.12) to find the optimal values of the control variables;
e) in the first iteration go back to a) with the new values of the control variables, in the following iterations evaluate the results: if the difference of the control variables from their values in the previous iteration does not satisfy a given convergence criterion go back to a), otherwise the solution procedure is over. This amounts to finding a fixed point between the instruments and the deterministic bias.

---

10. As in the full stochastic algorithm the antithetic variates method has been used to improve the accuracy of the evaluation of the bias.

Appendix B: Analytical properties of the model

The model used to compare the performance of alternative optimization algorithms is a slightly modified version of a simple two-equation nonlinear model that has been used extensively in the literature investigating the properties of alternative predictors for nonlinear models: see for instance Mariano and Brown (1983) and (1988).

The structural form is given by:

$$\begin{cases} \log y_t = \alpha_0 \log x_t + \alpha_1 \log y_{t-1} + u_t & \text{(B.1)} \\ z_t = \beta_0 x_t + \beta_1 y_t + v_t & \text{(B.2)} \end{cases}$$

with $(u_t, v_t)' \sim$ IIN $(\underline{0}, \Sigma)$. For simplicity we require Σ to be a diagonal matrix, and use σ_u^2 and σ_v^2 to denote the variances of u_t and v_t, respectively.

The optimization problems we are interested in always consist in moving the model from a given (stochastic) steady-state growth path to a different one. We therefore require the model to be homogenous of degree one: this simply amounts to imposing the constraint $\alpha_0 + \alpha_1 = 1$.

The main advantage of this model is that, although nonlinear, it has a closed-form solution: this enables us to make an analytical evaluation of the bias of the deterministic simulation, which is a key factor in explaining the differences in the performance of alternative optimization algorithms.

The reduced form is:

$$\begin{cases} y_t = \exp[\alpha_0 \log x_t + \alpha_1 \log y_{t-1} + u_t] & \text{(B.3)} \\ z_t = \beta_0 x_t + \beta_1 \exp[\alpha_0 \log x_t + \alpha_1 \log y_{t-1} + u_t] + v_t & \text{(B.4)} \end{cases}$$

The deterministic one-step ahead predictors, obtained by substituting u_t and v_t with their expected values, are:

$$\begin{cases} y_t^d = \exp[\alpha_0 \log x_t + \alpha_1 \log y_{t-1}] & \text{(B.5)} \\ z_t^d = \beta_0 x_t + \beta_1 \exp[\alpha_0 \log x_t + \alpha_1 \log y_{t-1}] & \text{(B.6)} \end{cases}$$

The closed-form one-step ahead predictors can be obtained by taking the expected value of the reduced form:

$$\begin{cases} y_t^c = E\,(y_t \mid x_t,\ y_{t-1}) = y_t^d \cdot \exp\,(\sigma_u^2/2) & \text{(B.7)} \\ z_t^c = E\,(z_t \mid x_t,\ y_{t-1}) = & \\ \quad = z_t^d + \beta_1 \exp[\alpha_0 \log x_t + \alpha_1 \log y_{t-1}] \cdot [\exp(\sigma_u^2/2) - 1] & \text{(B.8)} \end{cases}$$

Note that the expressions above would not be true in the presence of parameter uncertainty.

The one step ahead bias of the deterministic predictors can be easily evaluated as the difference between the deterministic and the closed-form solution:

$$\begin{cases} B_{y_t}^d = y_t^d\,[1-\exp(\sigma_u^2/2] & \text{(B.9)} \\ B_{z_t}^d = \beta_1 \cdot B_{y_t}^d & \text{(B.10)} \end{cases}$$

This measure of bias is not unit-free; a scale-freee measure of bias, that shows what portion of the mean-squared prediction error of the deterministic simulation can be attributed to the deterministic bias, is:

$$\psi_{z_t}^d = (B_{z_t}^d)^2/[\text{variance of } z_t] =$$

$$= [1-\exp(\sigma_u^2/2)]^2/[\exp(\sigma_u^2)(\exp(\sigma_u^2)-1)] \qquad \text{(B.11)}$$

It should be noted that the absolute bias depends on all the parameters as well as on the initial conditions, while the relative bias does not.

The formulae above have been obtained under the assumption that, in each period t, the value of y_{t-1} is known. However, in computing an optimal path over a given time horizon, e.g. [1, T], only the value of y_0 is known. That is, all the formulae above have to be computed conditional on y_0 (not on y_{t-1}).

The relevant formulae are those of equations [13]-[14]-[15]-[16] in the text.

A few remarks are in order:

i) The deterministic bias is due to the nonlinear relationship between the variables and the disturbances. In

this model it is the nonlinear relationship between u_t and y_t that produces the deterministic bias.

ii) Given that the only component responsible for the bias is the stochastic one in the first equation, the relative bias for z_t should decrease with the contribution of u_t with respect to that of v_t: it is easy to verify that $\psi^d_{z_t}$ tends to zero as (σ^2_v/σ^2_u) tends to infinity. Given this feature, we eliminated the error term from the second equation.

iii) The bias for z_t increases (in absolute value) with σ_u and β_1^2; the β_1 parameter measures the weight that the first nonlinear equation has in determining the value of z_t.

iv) From the expression relative to $\psi^d_{z_t}$, it appears that a large weighted bias is obtained for large values of σ^2_u. Given that the differences between the optimal paths obtained with the alternative algorithms discussed in Appendix A critically depend on the size of the bias, a reasonable experimental design would appear to consist in choosing very large values for σ^2_u. On the other hand we found that, with a high variance of u_t, the stochastic component of the first equation completely overshadows the deterministic structure of the model: in particular, the model never comes close to a steady-state growth path, as the growth rates of y_t and z_t always tend to behave randomly, driven by the u_t process. There thus appears to be a trade-off as to the choice of the σ^2_u parameters.

Given the role played by the various parameters in determining the properties of this model, we chose to generate several different data sets, with the following parameter values:

$\alpha_0 = 0.2, 0.5, 0.8$ $\quad (\alpha_1 = 0.8, 0.5, 0.2)$

$\beta_0 = 1$ $\quad \beta_1 = 0.1, 0.5, 0.9$

$\sigma^2_u = 0.01, 0.0001, 0.000001$ $\quad \sigma^2_v = 0$

The initial conditions are: $y_0 = 1000$, $x_1 = 1000$

One hundred data points have been generated for each possible parameter combination (giving 27 data sets), with the exogenous variable growing at 0.5 per cent per period; the same seed has been used to generate the stochastic disturbances in the 27 cases.

Table B.1 shows the main characteristics of three representative cases.

The first one is that corresponding to the largest bias

in our sample: as stated above (remark (iv)), a larger bias is obtained in exchange for a high volatility in the growth rates of the z_t variable. This case corresponds to the results that are detailed in the text.

The third case is the one corresponding to the smallest bias: it generates a z_t sample that grows at rates very close to the theoretical 0.5 per cent per period.

Table B.1

Characteristics of three representative data sets

	σ_u^2	α_1	β_1	bias in period 81	bias in period 100	weighted bias
Maximum bias	0.01	0.2	0.9	-6.7149	-7.3824	$2.5*10^{-3}$
Intermediate bias	0.0001	0.5	0.5	-0.03707	-0.04076	$2.5*10^{-5}$
Minimum bias	0.0000001	0.8	0.1	-0.00073	-0.00080	$2.5*10^{-7}$

Note: Figures are given for the bias in periods 81 and 100 since this is the time range for the optimization exercises.

Appendix C: Details on experimental design: Empirical measurement of deterministic bias

A key factor in determining the relative performance of the optimization algorithms compared in this paper is obviously the accuracy of the stochastic simulation as an approximation of the true (closed-form) solution. In particular, the performance of the algorithm proposed by Hall and Stephenson depends fundamentally on the correct measurement of the bias of the deterministic simulation.

Before comparing the relative performance of the optimization algorithms we therefore measured the deterministic bias with different numbers of replication and different techniques. Having done so, we chose what seemed to be a reasonable design for the simulation stage of the optimization algorithms.

There is obviously a trade-off between the accuracy of the approximation to the deterministic bias and the solution time; one way to improve the performance of the stochastic simulation dramatically without an enormous increase in the number of replications consists in making use of variance-reduction techniques.

In particular we use the so-called method of antithetic variates, first proposed by Calzolari (1979); this method considerably improves the degree of approximation to the true (closed-form) deterministic bias.

Here we present the comparisons relative to the data set having the following characteristics: $\alpha_1 = 0.2$, $\beta_1 = 0.9$, $\sigma_u^2 = 0.01$. A set of one hundred errors is generated having zero mean and variance equal to 0.01. The m replications (m = 100, 1,000, 10,000) used to compute the stochastic simulations are then extracted from this set using the method of McCarthy (McCarthy, 1972).

For a generic nonlinear model whose reduced form is given by:

$$y_t = g(x_t, u_t; \theta) \tag{C.1}$$

the stochastic simulation of size m is given by:

$$y_t^{s(m)} = \frac{1}{m} \sum_{i=1}^{m} g(x_t, u_t^{(i)}; \theta) \tag{C.2}$$

where the i index refers to the replication number.

Using antithetic techniques we have:

$$y_t^{s(m)A} = \frac{1}{2} \left[\frac{1}{m} \sum_{i=1}^{m} g(x_t, u_t^{(i)}; \theta) + \frac{1}{m} \sum_{i=1}^{m} g(x_t, -u_t^{(i)}; \theta) \right] \tag{C.3}$$

Chart C.1

Accuracy of measurement of the deterministic bias. Comparison of the results obtained with stochastic simulations with 100, 1,000 and 10,000 replications; standard method and antithetic technique

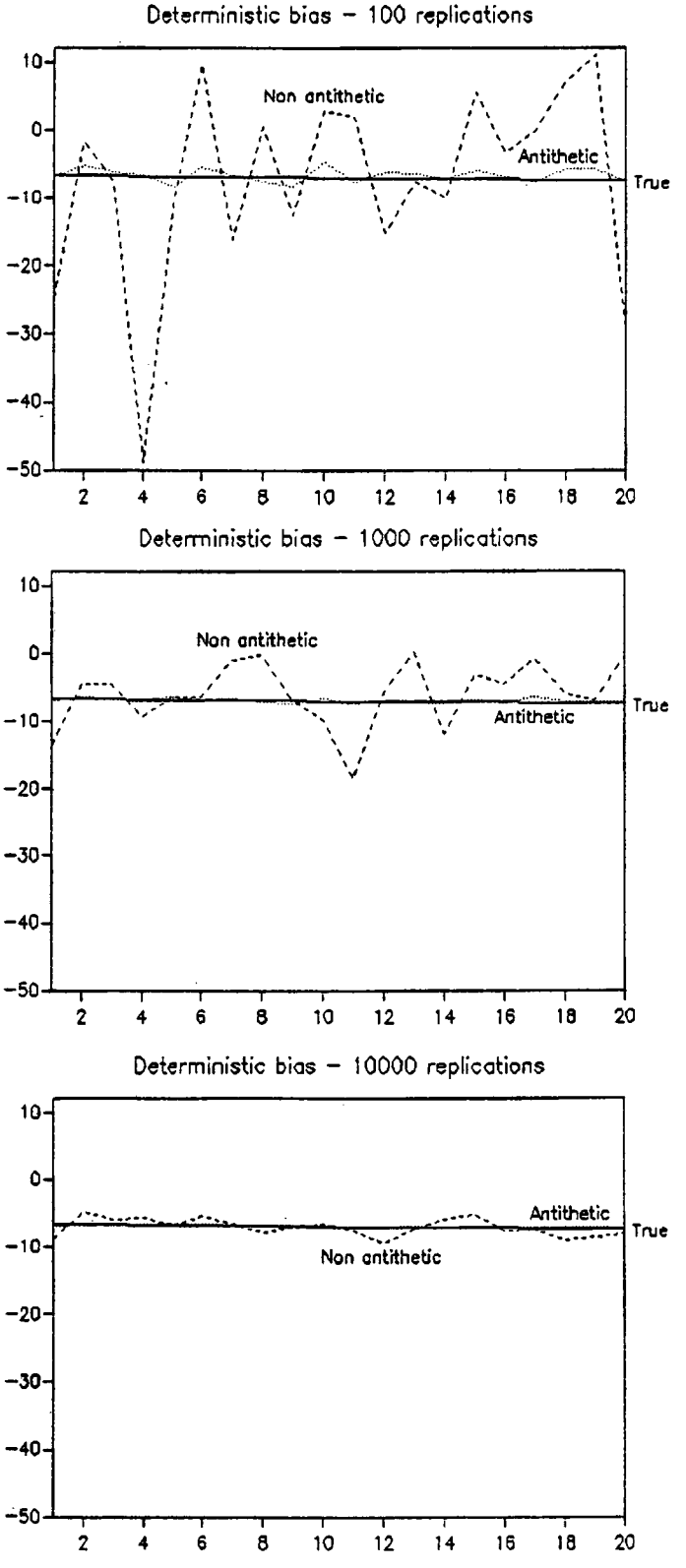


Chart C.1 shows the results. It can be seen that using antithetic techniques (which implies doubling the number of replications) leads to a dramatic improvement in the approximation to the true (closed-form) deterministic bias. For instance, while antithetic techniques give reasonable results with m = 100, the standard stochastic simulation sometimes gives a very poor measure of the deterministic bias. Moreover, antithetic simulations with m = 1,000 compare favourably with standard simulations with m = 10,000.

Generally speaking, antithetic techniques already perform quite well with m = 100, and the results improve only marginally by increasing the number of simulations. On the other hand, only 10,000 replications give reasonable results for the standard stochastic simulation case.

These results seem to justify the choice of 1,000 antithetic simulations as a reasonable compromise between accuracy and computing time. All the results presented in the paper are based on this experimental design.

REFERENCES

Brandsma, A.S. - Hallet, A.J.H. - Van der Windt, N. (1984), Optimal Control of Large Nonlinear Models: An Efficient Method of Optimal Policy Search Applied to the Dutch Economy, "Journal of Policy Modelling", vol. 5, pp. 253-270.

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

Chow, G.C. (1976), An Approach to Feedback Control of Nonlinear Econometric Systems, "Annals of Economic and Social Measurement", 5, 3, pp. 297-309.

Cividini, A. (1990), An Optimal Control Algorithm for Analyzing Econometric Models, in "Proceedings of the IFIP TC7 Conference on Modelling the Innovation: Communications, Automation and Information Systems", edited by M. Carnevale - M. Lucertini - S. Nicosia, Amsterdam, North-Holland.

_____ - Petersen, C.E. (1987), Implementig Optimal Ordering and Newton-Raphson Algorithms in the Modeleasy Software for Simulation of Econometric Models, Banca d'Italia, mimeo.

_____ - Siviero, S. (1992), Approximated Solution of Stochastic Optimal Control Problems for Nonlinear Models: A Comment on an Algorithm Recently Proposed by Hall and Stephenson, Banca d'Italia, mimeo.

Federal Reserve System (1986), Modeleasy Speakeasy User Guides, vol. III, mimeo.

Hall, S.G. - Stephenson, M.J. (1990a), An Algorithm for the Solution of Stochastic Optimal Control Problems for Large Nonlinear Econometric Models, "Journal of Applied Econometrics", 5, pp. 393-399.

_____ - _____ (1990b), Optimal Control of Stochastic Nonlinear Models, in "Economic Modelling at the Bank of England", edited by S.G.B. Henry - K.D. Patterson, London, Chapman and Hall.

Holbrook, R. (1974), A Practical Method for Controlling a Large Non Linear Stochastic Model, "Annals of Economic and Social Measurement", 1, pp.155-175.

Mariano, R.S. - Brown, B.W. (1983), Prediction Based Tests of Misspecification in Nonlinear Simultaneous Systems, in "Studies in Econometrics, Time Series and Multivariate Statistics", edited by S. Karlin - T. Amemiya - L. Goodman, New York, Academic Press.

_____ - _____ (1988), Predictors in Dynamic Nonlinear Models: Finite Sample Behavior, University of Pennsylvania, mimeo.

McCarthy, M.D. (1972), Some Notes on Generation of Pseudostructural Errors for Use in Stochastic Simulation Studies, in "Econometric Models of Cyclical Behavior. Appendix", edited by B.G. Hickman, New York, Columbia University Press.

Mitchell, D.W. (1979), Risk Aversion and Optimal Macro Policy, "Economic Journal", 89, pp. 913-918.

Petersen, C.E. - Cividini, A. (1989), Vectorization and Econometric Model Simulation, "Computer Science in Economics and Management", 2, pp. 103-117.

RECENTLY PUBLISHED «TEMI» (*)

n. 153 — *Dinamica retributiva e differenziali salariali,* di A. GAVOSTO - P. SESTITO (luglio 1991).

n. 154 — *Interessi reali, sistema impositivo ed effetto Sylos Labini,* di P. VAGLIASINDI (luglio 1991).

n. 155 — *Trasformazione delle scadenze e margine d'interesse degli istituti di credito mobiliare,* di P. SABBATINI (luglio 1991).

n. 156 — *Gli effetti della quotazione internazionale: il caso delle azioni italiane a Londra,* di F. PANETTA (agosto 1991).

n. 157 — *Grandi e piccole imprese negli anni ottanta: la ristrutturazione dell'industria in un'analisi di dati di bilancio,* di L. F. SIGNORINI (agosto 1991).

n. 158 — *Demand and Supply Shocks in Industrial Output,* by A. GAVOSTO - G. PELLEGRINI (novembre 1991).

n. 159 — *I* futures *e le opzioni sui titoli di Stato. Un'analisi del mercato e delle prospettive in Italia,* di A. SCALIA - L. TORNETTA (novembre 1991).

n. 160 — *Earnings Uncertainty and Precautionary Saving,* by L. GUISO - T. JAPPELLI - D. TERLIZZESE (febbraio 1992).

n. 161 — *Migrazioni in Europa: andamenti, prospettive, indicazioni di politica economica,* di G. GOMEL - S. REBECCHINI (febbraio 1992).

n. 162 — *Monetary Aggregates and Monetary Policy Coordination on the Way to Economic and Monetary Union: the Role of Cross-Border Deposits,* by P. GIUCCA - A. LEVY (febbraio 1992).

n. 163 — *Cross-Border Deposits and Monetary Aggregates in the Transition to EMU,* by I. ANGELONI - C. COTTARELLI - A. LEVY (marzo 1992).

n. 164 — *Young Households' Saving and the Life Cycle of Opportunities. Evidence from Japan and Italy,* by A. ANDO - L. GUISO - D. TERLIZZESE (marzo 1992).

n. 165 — *Bequests and Saving for Retirement. What Impels the Accumulation of Wealth?,* by F. BARCA - L. CANNARI - L. GUISO (marzo 1992).

n. 166 — *The Microeconomics and Macroeconomics of the Permanent Income Hypothesis,* by A. DEATON (marzo 1992).

n. 167 — *Why is Italy's Saving Rate so High?,* by L. GUISO - T. JAPPELLI - D. TERLIZZESE (aprile 1992).

n. 168 — *Waiting for EMU: Living with Monetary Policy Asymmetries in the EMS,* by L. BINI SMAGHI (aprile 1992).

n. 169 — *Income and Saving in Italy: a Reconstruction,* by G. MAROTTA - P. PAGLIANO - N. ROSSI (giugno 1992).

n. 170 — *Finance and Development: The Case of Southern Italy,* by R. FAINI - G. GALLI - C. GIANNINI (giugno 1992).

n. 171 — *Generational Accounting: The Case of Italy,* by D. FRANCO - J. GOKHALE - L. GUISO - L. J. KOTLIKOFF - N. SARTOR (giugno 1992).

n. 172 — *Mancate interviste e distorsione degli stimatori,* di L. CANNARI - G. D'ALESSIO (giugno 1992).

n. 173 — *Inflazione attesa, tassi reali e la struttura per scadenza dei tassi d'interesse,* di R. CESARI (luglio 1992).

n. 174 — *Economie di scala e di diversificazione nell'industria bancaria: il ruolo dell'eterogeneità tra imprese,* di G. PARIGI - P. SESTITO - U. VIVIANI (luglio 1992).

n. 175 — *L'economia italiana nella prospettiva europea: terziario protetto e dinamica dei redditi nominali,* di F. BARCA - I. VISCO (luglio 1992).

n. 176 — *Struttura finanziaria e riforme nei paesi dell'Europa centrale e orientale,* di P. CATTE - C. MASTROPASQUA (settembre 1992).

n. 177 — *Creazione e distruzione di posti di lavoro in Italia,* di B. CONTINI - A. GAVOSTO - R. REVELLI - P. SESTITO (settembre 1992).

n. 178 — *Saving and Government Deficits (1951-1990),* by N. ROSSI - I. VISCO (settembre 1992).

(*) Requests for copies should be sent to:
Banca d'Italia - Servizio Studi - Divisione Biblioteca e Pubblicazioni - Via Nazionale, 91 - 00184 Rome.

BANCA D'ITALIA - CENTRO STAMPA