The role of simulation methods in Macroeconomics

Alfonso Novales*
Departamento de Economía Cuantitativa
Universidad Complutense

September 2000

Abstract

After reviewing the reasons to use solution methods in macroeconomics, this survey paper discusses different aspects relative to a rigorous use of the numerical output of such methods. Special attention is paid to suggestions that have been made to incorporate parameter uncertainty. Finally, the need to test for usually maintained assumptions, such as rationality of expectations, is emphasized.

Key words: Numerical solution methods, rational expectations, calibration.

JEL classification: C8, C9, E3, E6, O4.

---

*Departamento de Economía Cuantitativa, Universidad Complutense, Somosaguas, 28223 Madrid. E-mail: eccua09@sis.ucm.es

1. Introduction

Widespread use of dynamic, stochastic, model economies has led to the need of using numerical methods to characterize the properties of a given theoretical economy. Even though model simulation has become almost standard in some research areas in economics, there is still some misunderstanding regarding a correct use of numerical methods. The non-specialist has sometimes the impression that simulation is more a caprice of the researcher than a real need. Solution methods seem difficult to understand and replicate. Besides, assignment of numerical values to key structural parameters is thought from the outside to be an arbitrary decision, which totally conditions the results. That way, the argument goes, any model can conceivable be consistent with a too wide variety of properties: what is true for some parameter values, can easily be shown to be false for some others. Lastly, fundamental skepticism comes from considering whether results characterized by simulation can ever be compared to properties we learn about through a formal mathematical proof.

This paper explains that numerical solution methods are needed to characterize the type of models which are increasingly considered to be appropriate for many purposes, and specifically, for policy analysis. Without entering into technical details characterizing the main solution methods, which have been adequately covered somewhere else[1], we discuss some of the main issues concerning implementation of solution methods, such as calibrating parameter values, and producing and interpreting the implied results. Finally, we argue that the use of numerical methods has considerably enlarged the class of questions we address when analyzing dynamic, stochastic economic models[2].

In section 2 we explain why numerical solutions are the only possibility to analyze a wide class of interesting economic models. In section 3 we discuss how the results obtained through numerical simulation should be presented. Section 4 deals with model specification and calibration. It discusses the limitations in actual practice, and describes interesting suggestions that have recently been made by some authors. In section 5 we describe the Bayesian approach to simulation. Section 6 reviews different approaches for evaluation of calibrated models. Section

---

[1]See the January 1990 issue of the Journal of Business and Economic Statistics, or the volume: Computational methods for the Study of Dynamic Economies, cited in the references, as examples.

[2]The crossed discussion in the papers by F.Kydland and E.C.Prescott, L.P.Hansen and J.J.Heckman, and C.A.Sims in the Winter 1996 of the Journal of Economic Perspectives is strongly recommended.

7 emphasizes the importance of guaranteeing stability of the implied solution. Section 8 discusses some statistical issues having to do with the analysis of the results, while section 9 points out the relevance of characterizing the transition between steady-states. Section 10 argues that researches have in fact, added interesting and important questions to their analysis since being able to simulate theoretical models. The paper closes with some conclusions.

2. Why do we need to simulate?

The influential work of a number of economists during the seventies[3] showed the need to formulate macroeconomic questions in a dynamic, stochastic setup, to avoid the important bias that could be introduced by addressing economic questions in a more limited framework. Agents in an economy (consumers, firms, even governments), were viewed to make their decisions by optimizing specific objective functions, taking into account that the consequences of their actions: 1) are uncertain a priori, 2) will be noticed through a number of periods, 2) influence some other variables, that will in turn feedback into the economy.

Economists turned then their attention to mathematical methods to solve dynamic, stochastic optimization problems [Bellman´s dynamic programming, Pontryagin´s optimal control principle, or the work by Kushner on stochastic control]. Linear-quadratic optimization problems, those in which the objective function is quadratic and restrictions are linear, will produce a quadratic Lagrangian and hence, first order conditions will be linear in state and decision variables. State variables are predetermined each period, being either past decision variables or variables exogenous to the decision-maker. In a model in which different economic agents solve their own optimization problem, variables which are decisions for one agent may be state variables for other agents in the economy. In a deterministic setup, first order conditions, together with budget constraints and the assumed mechanism for price formation will form a linear system each period, with as many equations as decision variables, providing the optimal values for the decision variables as a function of the values of the state variables. It is necessary, however, to check that transversality conditions hold when they are necessary for optimality, as it is the case in most economic models.

Under uncertainty, we enter a higher level of difficulty: the system of first order conditions is no longer a *complete system*, since it involves expectations of some

---

[3]Brock and Mirman (1972), Lucas (1976), Sargent (1979) and the collections of papers by Phelps et al. (1970), and Lucas and Sargent (1981) among many others.

functions of state and decision variables, as well as their realized values. If the expectations formation mechanism is assumed to be endogenous, the system will not be complete. In linear-quadratic problems, state and decision variables enter into the first order conditions in a linear manner, so that we have either a variable or its expectations, possibly at different horizons. However, in this special case, separation of control and estimation (*the certainty-equivalence principle*) applies. That allows for application of standard deterministic optimization methods, imposing afterwards conditional expectations where needed [see Sargent (1979)] to solve for the current value of each decision variable as a function of: past values of decision and state variables and expectations of future exogenous variables. It is important that in those solutions we should not have current expectations of future decision variables, which are also endogenous.

This is relatively simple in a partial equilibrium framework, where, *e.g.*, a price-taking firm maximizing expected present value of profits will make its employment decisions each period as a function of the number of workers hired in previous periods, and its expectations about the future evolution of salaries and output prices. Then, all we need is to compute analytical expressions for expectations of the state variables. Perceptions of the optimizing agent (be a consumer, firm or the government) on the stochastic structure of the process governing the behavior of state variables is all we need to eliminate those expectations and compute optimal values for each decision variable. This means, in particular, two well known principles: 1) it is the perceptions of rational economic agents on the future evolution of exogenous variables that matters, rather than their actual probability characteristics, and 2) optimal decision rules for private agents will depend on the structure and parameter values used in the policy rules, which gives rise to the Lucas critique [Lucas (1976)].

The previous considerations do not essentially change if we consider rational expectations *reduced-form linear models* that include expectations, at different points in time, of future state and decision variables, even if they have not been explicitly obtained as solution to dynamic, stochastic control problems. Methods to analytically solve these models can be seen in Whiteman (1983).

Of course, one can proceed to develop a full, general equilibrium model, with each agent solving a specific optimization problem in which, on top of the full set of first order conditions for each problem, constraints and market-clearing conditions are imposed. Consideration of equilibrium conditions will give us a complete system which should conceivably allow us to solve for equilibrium paths for endogenous variables. However, doing that, we enter an additional level of dif-

ficulty: general equilibrium considerations preclude us from assuming any specific perception about the stochastic process governing future prices to be substituted into the expectations expressions. We now need to simultaneously solve for the optimal values of the decision variables of each agent in the economy, as well as for equilibrium prices. Once prices are endogenous, the budget constraints that enter into the optimization problems will no longer be linear, since they will involve cross products of endogenous prices and decision variables. Losing the linear-quadratic setup, we can no longer invoke the certainty-equivalence principle, *i.e.*, the separation of estimation and control and, consequently, there is no hope of applying the methods in Sargent(1979). This is a much more complicated setup.

In the simple growth model in which the representative agent faces a known, but endogenously determined rate of return on his/her savings when maximizing expected time-discounted utility, we will have: $R_t = \frac{\partial U/\partial C_t}{\beta \ E_t(\partial U/\partial C_{t+1})}$, where R_t denotes the gross equilibrium rate of return, which is announced at time t, to be paid at time $t+1$ on one-period investments. We could similarly obtain longer-horizon equilibrium rates of return. The consumer saves to the point where his/her marginal utility of consumption (output minus savings) is equal to the cross product of the discounted value of R_t and the expected level of marginal utility at the time the savings returns will be received. This equilibrium condition could be used to obtain equilibrium prices each period if we made any ad-hoc assumption on the stochastic behavior of future consumption or on the nature of expectations, be they adaptive or of any kind. Together with the rest of equilibrium conditions, that would allow us to solve for the remaining endogenous quantities and prices: savings, output and interest rates, obtaining a specific stochastic process for each of them. However, rationality of expectations is contrary to arbitrary assumptions on the expectations formation mechanism, which must be fully consistent with the structural model.

It is clear that we face a very serious problem. If we believe that economic agents take decisions under uncertainty, being aware of the fact that their actions influence their own feasible sets in the future, we will need to specify stochastic, dynamic models. If we want to obtain their implications under the assumption that prices are consequence of market interactions among agents and expectations are endogenous (rational), an analytical solution will not exist except for very special cases [McCallum (1989) and Marcet (1994)]. But it is precisely this type of model that we want to use to analyze the consequences of alternative economic policies. These models cannot be analyzed except by numerical simulation, which is therefore not an option alternative to some others, but rather, the only way to

fully analyze a broad class of very relevant model economies[4]. The discussion on whether it is convenient to simulate economic models is largely spurious.

Several points are worthwhile making before we move into a deeper discussion: 1) *numerical methods are not specific to equilibrium models.* Their need is motivated by the appearance of expectations of nonlinear functions of state and decision variables, together with the assumption of rationality [see Danthine and Donaldson (1992)], 2) when the equivalence between the competitive equilibrium and the central planning resource allocation mechanisms holds, it is generally helpful to obtain the numerical solution to the equilibrium model. Extended use of this property leads sometimes to the perception that *numerical methods are specific to models where the second welfare theorem holds, which is not the case,* 3) *analysis by numerical methods* can be used to deal with agents heterogeneity, so *it is not restricted to the representative agent framework,* 4) the previous comments apply to any optimizing behavior, so *numerical methods are not specific to Macroeconomics,* 5) by the same token, *simulation methods and numerical solutions are not specific to real business cycle theories:* even if stochastic shocks affect policy rules, or agents' preferences, and not technology, numerical solutions may provide the only analysis feasible in a model with endogenous expectations.

3. What do we get out of model simulation?

A theoretical dynamic, stochastic model can be seen as imposing a set of restrictions on the probability distribution of the vector of relevant variables. Such restrictions emerge from: *a)* the analytical structure of the model: functional specifications for technology, preferences, productive and human capital obsolescence, information sets for each agent, endowments, etc., *b)* parameter values,

---

[4]The competition among solution methods which has sometimes been considered in the literature is subject to some logical limitations, since there is nothing like a best solution method. Solution methods impose different approximations which allow for a numerical solution to be obtained, and for characterizing the main properties of a model. Approximations are needed because the solution to the original, non-linear model (a set of nonlinear difference equations) can not be obtained. There is then a trade-off between making approximations so that the solution approach as simple as possible, but not so many assumptions so as to produce a numerical solution whose properties on relevant issues may significantly differ from those from the original model. Again, we will never be able to quantify those differences because, to do so, we would need the solution to the original model in the first place, so this is a delicate issue that may call for investing in robust solution methods, even if they are computationally more demanding, or more complicated to implement.

and *c*) the multivariate probability distribution of the vector stochastic process for the exogenous perturbations.

The analytical *solution to the model* is the probability distribution of that vector stochastic process. In the case of a stochastic, general equilibrium model, that solution is the *dynamic, stochastic equilibrium of the model.* Some, but not all, of the characteristics of that probability distribution can then be obtained from the analytical solution. However, we have argued above that, very often, such a solution cannot be obtained. Then, *simulation* methods provide us with an approximation, in the form of a frequency distribution for that vector stochastic process.

Computing a *numerical solution* to a set of equations summarizing the main properties of a model economy is just the first step in model simulation. A *numerical solution* is a set of time series, one for each relevant variable in the model economy, satisfying each period all the conditions in the model. *Simulation* is a procedure by which a *numerical solution* is found for each specific time series realization of the vector stochastic process of the exogenous perturbations in the economy. As we will explain below, sometimes parameter values are also changed across different numerical solutions. By reproducing a large number of sample realizations, we can approximate arbitrarily well the probability distribution of the vector stochastic process of relevant variables.

We often view economies as being in their *stochastic steady state*[5], characterized as stable fluctuations around a *deterministic steady state*, which may or may not exhibit growth. In exogenous growth models, the deterministic steady state arises by setting all random perturbations to zero at all periods, and imposing specific constant rates of growth for the endogenous variables. Another, more interesting case, is that of endogenous growth models. At a difference from the former, in these models the obtained time series will not be stationary even after correcting for the deterministic trend. Even though most of the analysis of endogenous growth models has so far been performed just in steady-state, so only long-run statements have usually been made, a numerical solution can also be found for these models. [Novales et al. (1999)]. As explained in section 9, there is no reason why the analysis should be restricted to steady-state. Once we have the tools for solving non-linear, dynamic, stochastic models, we are equipped to also analyze what happens to an economy in transition to its new steady state

---

[5]We use the *steady-state* denomination indistinctly from the *balanced growth* path, even though they are not equivalent concepts. The reader should use one or the other, according to the model he/she has in mind.

from a given initial situation. This is crucial in policy evaluation exercises, since a policy intervention will generally take an economy outside steady-state. By characterizing the transition, we can compute the welfare consequences of the policy intervention, and not limit our analysis just to the long-run consequences, *i.e.*, once the new steady-state has been reached (if ever), which could well be misleading.

Once a specific sample realization for the vector of states and decisions has been obtained, we can then summarize the properties of their joint distribution in the form of standard statistics: sample means, standard deviations, coefficients of variation, simple and partial autocorrelation functions, correlation coefficients, regression coefficients, cross correlation functions, vector autoregressive representations (VAR), impulse responses in a subset of variables, decompositions of variance, spectral density matrices, etc.. For each of these point statistics we will obtain as many realizations as numerical solutions we get for the model, i.e., as many as sample realizations we draw from the probability distribution for the exogenous random perturbations in our simulations.

The realizations of the statistic across a large number of simulations configure its empirical frequency distribution which can be taken as an approximation to its true, unknown density function. If we are interested in a price-elasticity of demand, or on the relative volatility of two variables, we can report not only its mean value in a given model and for given parameter values, but rather, its full empirical distribution. Simulation results should be reported by providing the full information generated in the analysis, *i.e.,* the empirical distribution for each of the statistics of interest. However, it should be born in mind that such an output is consequence of the aggregate of: *a*) a given structure for the theoretical model, *b*) given parameter values, *c*) a given probability distribution of the vector of structural shocks.

In this view, that density function emerges as a consequence of the sampling error associated to the structural shocks impinging on the economy, which leads to simulating the theoretical model under a large number of realizations drawn from theoretical probability distributions for exogenous shocks. Alternative views are discussed later on in this paper.

This raises the possibility of a very rich statistical analysis. As an example, the variance of the empirical distribution of a regression coefficient, estimated by least-squares with each realization of the set of equilibrium time series, will coincide with the variance reported by the least-squares theory when a single realization is available, only if the assumptions underlying the estimation method

hold, which will not usually be the case. As another example, there is no reason why the empirical distribution of an estimated statistic might not be bimodal, even if we sample from well-behaved, Gaussian probability distributions for the random perturbations in the model[6].

4. Calibrating a theoretical model

To simulate a model, we need first to assign numerical values to its structural parameters. Then simulation allows us to characterize the model's properties, which the researcher will want to compare with their analogue, computed from actual data. Before that, he/she will have selected a set of such characteristics as relevant for his/her analysis. There is, hence, some sense in which the artificial economy is estimated and tested, since after the mentioned comparison, we will conclude whether or not the model is adequate for the issue in mind.

4.1. What is a reasonable model?

Selection of an appropriate theoretical model is not an aspect specific to simulation, but rather it is common to any research strategy. Even though it includes a great variety of approaches to the confrontation of theory with data, those who follow the methodology described above share some criteria about the properties that a model must have, which can be traced back to work by Friedman (1953) on the concepts of simplicity and realism, and later exposed by R.J.Lucas ["Methods and Problems in Business Cycle Theory", 1980]:

"...One of the functions of economic theory is to provide fully articulated, artificial economic systems that can serve as laboratories in which policies that would be prohibitively expensive to experiment with in actual economies can be tested out at much lower cost..." [...] "...Insistence in the realism of an economic model subverts its potential usefulness in thinking about reality. Any model that is well articulated to give clear answers to the questions we put to it will necessarily be artificial, abstract, patently unreal..." [...] "...Not all well-articulated models will be equally useful. Though we are interested in models because we believe they

---

[6]An alternative strategy when working with a real business cycle model with a single productivity shock, consists on simulating the model using Solow residuals estimated with data from a real economy are as realization for the productivity shock. Then, each statistic will take a single value and we will not have an empirical distribution. In that case, the possibility of sampling error is not considered.

may help us to understand matters about which we are currently ignorant, we need to test them as useful imitations of reality, by subjecting them to shocks for which we are fairly certain how actual economies would react. The more dimensions on which the model mimics the answers actual economies give to simple questions, the more we trust its answers to harder questions. This is the sense in which more realism in a model is clearly preferred to less".

The researcher is not interested in verifying if the model is correct, since he/she knows from the beginning that it is not. He/she is satisfied with the fact that, through a theoretical re-specification process, a simple, stylized model can be found that captures an increasing number of data characteristics. An interesting iterative methodological process starts by which the researcher repeatedly moves from theoretical model to actual data and back until reaching convergence. Then, the degree of satisfaction with the limit reached must be evaluated, i.e., the number and importance of empirical aspects that the theory has been able to account for. Research should focus on characterizing robustness, *i.e.*, on how the answer to a question of interest changes with local or global changes in the structure of the model.

Questions of interest to the researcher can take different forms: is it possible to mimic a given empirical regularity using a particular model? and if so, how much of that regularity can be explained by impulses in a given shock? how does the stochastic process for the vector endogenous variables change if the stochastic process for the vector of exogenous variables is modified? is it possible to reduce a specific discrepancy between theoretical model and data by introducing a particular structural feature in the model? Some examples are: can we use the neoclassical growth model to explain the empirical observation, common to many countries, that the relative volatility consumption/output is less than one, while the relative volatility investment/output is well above one? To what extent can we reproduce the numerical values of those volatilities considering just technology shocks? Does the goodness of fit of the model significantly increase if we add shocks to preferences of individual consumers? What effects does it have on the economy to decide and announce a given time path for tax rates on income for the next four years, versus the possibility of maintaining continuous discretionality on them? What is the effect on price volatility in the neoclassical, monetary growth model (Sidrauski (1967)) if the monetary authority quits the monetary rule of $k\%$ annual growth for the money supply to implement a policy of controlling interest rates? Since the standard growth model predicts a high correlation between hours worked and productivity, which is contrary to empirical evidence, will incorporat-

ing a second sector, for human capital accumulation, improve upon this property of the model?

Many limitations that, in a first analysis, are associated with the methodology of analyzing models through simulation are more apparent than real. It is sometimes said that: *a*) almost any model is able to mimic the value of an statistic of interest in an actual economy, just by appropriately picking parameter values, and also that *b*) different models can be found that are able to account for the sample value a given statistic and, yet, have radically different implications relative to evaluating alternative economic policies. The first question has a simple answer and it is not, by any means, an important limitation: it is not true that there always exists a parameter vector that can make any model to replicate a given empirical regularity. Furthermore, selecting a model among other candidates just because it best explains the value of a particular statistic in actual data is a bad research strategy, very different from what we are proposing, since it might well be the case that such a model performs very badly relative to almost any other statistic of interest. A more correct strategy consists in selecting a model that mimics to a reasonable extent the values of a variety of statistics relevant to the empirical regularity under consideration

Relative to the second statement, it is true that there will generally be several models able to explain to a similar extent a given empirical regularity. That is just the reflection of the fundamental problem of *lack of identification* which arises when economic models are confronted with data. Resolution of this dilemma rests on augmenting the level of exigency, asking the model to also be able to explain additional empirical characteristics of the economy. In other words, lack of identification is, in many occasions, just a reflection of the fact that the loss function used to rank alternative models includes too short a list of arguments. It may also happen that there are different points in the parameter space able to produce a given data characteristic in a given model. Again, admissible parameter values are those that, imposed on the selected economic structure, have reasonable implications relative to aspects other than the one which is the central object of research.

4.2. What is calibration?

There is not a clear-cut definition of calibration, which makes the discussion intricate, unless a specific concept is chosen. From a purely technical point of view,

calibrating a model consists in associating numerical values to its parameters[7], so that a given numerical solution method can be used to generate time series sample realizations for its variables. Since it associates numerical values to parameters, there is some sense in which calibration is similar to estimation. Nevertheless, the relationship between calibration and inferential methods of classical statistics, *estimation* and *hypothesis testing*, is one of the least clarified aspects of numerical solution methods.

First, values for some structural parameters, typically the output elasticity of some production factors, the degree of risk-aversion or the elasticity of intertemporal substitution, are taken from micro data estimates or, from some casual empirical characteristics from the economy which is to be studied. After that, we use the fact that steady-state values of the variables in any dynamic model economy can be written as (nonlinear) functions of structural parameters. The standard calibration method, which started with Kydland and Prescott (1982) is rather informal, and uses those expressions to derive values for some of the remaining structural parameters so that steady-state levels for the most relevant variables match sample averages observed in actual time series data. In this computation the possibility that the data may not be stationarity needs to be taken into account. A number of such conditions is used smaller than the number of parameters in the model, so that some of them will remain free. Finally, a last subset of sample moments is used for evaluating the degree of fit of the model.

Hansen and Heckman (1996) justify the outlined use of sample averages, because: *a*) they are more robust than other statistics to the presence of zero-mean measurement errors, and *b*) steady-state relationships are generally robust against alternative specifications of short-run dynamics. Reproducing sample averages, calibration of a given set of sample moments may be consistent with a wide array of models, some of which may differ in their short-run implications. However, even without measurement errors, Sargent (1987) shows that correlations and cross-correlations can contain more information on a given model than sample averages, so that it is important to test the model by confronting the data in many dimensions, using a wide variety of statistics from the multivariate distribution of the variables of interest.

---

[7]Pagan (1994), Canova (1994) and Canova and Ortega (1996), take a more general view. For Canova (1994), calibrating is the full process which starts with the definition of the question to be analyzed, selecting a theoretical model, assigning parameter values, generating time series, characterizing its properties, and comfronting them with the similar characteristics in a real economy (empirical regularities), and it may even include economic policy analysis.

Once the model has been simulated, we will be able to use sample moments not used in calibration (degrees of freedom) to test it. For that, it is standard to compute a given set of moments in the simulated series (usually, relative volatilities and correlation coefficients), analyze their sensitivity to changes in the free parameters, and examine whether in some cases, they come close to the values of the same moments in a real economy. Evaluating the statistical significance of the distance between actual and simulated moments, the researcher is testing the adequacy of the model to represent important features of actual data, so there is also some sense in which calibration, in a broad sense that includes simulation, is a testing procedure.

4.3. Limitations in structural parameter calibration

Lacking in most cases a rich enough history of empirical estimates for structural parameters, the range of values that can be considered for a key parameter such as the elasticity of intertemporal substitution is too wide for most purposes, in spite of the fact that this parameter, by itself, conditions some of the most relevant statistical properties of a theoretical model.

Furthermore, a given empirical analysis will generally be different from any other in aspects which may be substantial, so that choosing parameter estimates from any one of them is not fully consistent, and produces significant parameter uncertainty which is almost always neglected in model evaluation. Besides, that some parameter values are chosen a priori implies a selection bias, since there are many empirical studies that could be used, so that different researchers can use different references for calibration,

Cross-section estimates are sometimes used for calibration, but it is not obvious which is their relationship with time series parameters. For instance, the marginal propensity to consume estimated in a cross section indicates the variation in consumption expenditures produced when we consider families of different income at a given time period, i.e., for given economic conditions. On the other hand, the marginal propensity to consume in time series refers to the variation on consumption expenditures which arises when income changes over the business cycle, which are two quite different things.

4.4. Limitations in calibrating exogenous stochastic processes

Calibrating exogenous stochastic processes is also necessary for model simulation, but it is hard to find information from a real economy concerning the stochastic

structure of technology shocks, shocks in preferences, errors of controlling money growth or tax revenues, or the correlations among them. Standard deviations of structural shocks are usually chosen so that the volatility of a key variable, such as output, or the ratios of volatilities of consumption, investment or hours worked to output, match that of a real economy.

1) Persistence properties in actual time series data is also used to calibrate some aspects of the model. In the simplest business cycle model, an $AR(1)$ model is assumed for productivity shocks, with the coefficient generally chosen so that the simulated output series exhibits a persistence similar to the GNP series in actual economies. Unfortunately, the strategy of replicating output persistence seems to condition most of the model´s properties, and the ability of the basic model to explain empirical business cycle regularities dramatically decreases when productivity shocks are assumed not to have persistence. Extreme care must be used when calibrating models so as not to achieve a spurious adjustment to data through *ad-hoc* assumptions.

2) Generally, exogenous perturbations are assumed to be independent when simulating, even though we have already pointed out that the probability distribution of the vector stochastic process of exogenous shocks is a key aspect in determining the model's properties. Policy evaluation is generally performed under orthogonality assumptions among the perturbations to policy variables (exogenous shocks or control errors), or between these and state variables, and the sensitivity of results to these assumptions has barely been analyzed. The possible ability of the economic authority to intentionally perturb policy variables, and establish some correlations between policy induced perturbations and observed exogenous shocks adds a new dimension of great interest to the analysis, where asymmetric information may play a crucial role. This issue can only be analyzed with numerical procedures[8]. We will get back to this in section 11.

4.5. Calibration versus formal estimation

From a Bayesian viewpoint, estimation is the solution to a problem of minimizing (the expected value of) a given loss function. Different estimators are solutions to problems with different loss functions, so discussing their relative properties does not makes much sense without a reference to the corresponding loss functions.

Simulating a theoretical model requires assigning numerical values to struc-

---

[8]As another example, Cassou (1995) considers shocks to tax rates and public expenditures, to characterize the optimal correlation between them.

14

tural parameters, and the different ways of doing that assignment are, more or less formally, different estimation procedures. As pointed out by Hansen and Heckman (1996), it is somewhat paradoxical that agents in an economic model are assumed to make optimal decisions relative to some loss (or objective) function, while the researcher does not do the same. What is then an appropriate loss function? Calibration, based in choosing values for some structural parameters as a function of long-run sample averages, can be interpreted to the light of a particular loss function, which weighs heavily matching some long-run characteristics, relative to other data properties.

A quite different strategy seeks to use the simulated time series to estimate some or all structural parameters through a formal method like the *Generalized Method of Moments* (*GMM*) [Duffie and Singleton (1993), Christiano and Eichenbaum (1992)], *Maximum Likelihood* (*MV*) [for instance, McGrattan, Rogerson, Wright (1991))], or the *Simulated Method of Moments* (*SMM*) [Lee and Ingram (1991), García-Mila (1987), Valles(1997)]. These more standard econometric procedures choose values for all parameters: a) optimizing a given criterion (the likelihood of the data, given the model, in the case of *MV*), b) exploiting orthogonality conditions implied by the conditional moments involved in the optimality conditions (*GMM*), or c) minimizing the distance between simulated moments and moments computed from actual data (*SMM*).

Statistical evaluation of a loss function has several advantages: *a*) it avoids a possibly arbitrary election of parameter values, and *b*) it provides a measure of dispersion that can be used to evaluate the goodness of fit of model to data. It also has some disadvantages, since *a*) it needs a specific selection of moments to be used in fitting evaluation, *b*) there are finite sample biases, which may lead to spurious inference, *c*) the type of uncertainty which is imposed on the model by an estimation process does not necessarily reflect the uncertainty faced by a researcher when calibrating a vector of parameters, which is specified more appropriately through Bayesian methods.

5. A Bayesian approach to simulation

The generalized practice of deriving the implications of a model for a given parameterization, obtained from previous estimations or from beliefs on the structural characteristics of a real economy, disregards the existence of uncertainty on parameter values. Even though alternative values are often considered for some key parameters, estimation by calibration is considered to be exact. This practice is

too restrictive: precise numerical results are provided for a given set of selected statistics, but no measure of uncertainty is presented.

A given belief on structural parameter values could be incorporated in the form of a *prior* probability distribution on the parameter space. Parameter uncertainty could then translate into uncertainty on the value of a given statistic, in the form of a frequency distribution. Canova (1994, 1995) and Canova and de Nicolo (1995) consider actual data statistics as fixed numbers, while the uncertainty in simulated data is used to provide a measure of how well the model fits actual data. This uncertainty comes from the probability distribution of exogenous shocks, as well as from parameter variability. Rather than fixing their numerical values, empirical information on some parameters is used to build a probability distribution on the parameter space, and each simulation is computed with a different point drawn from that distribution. *"The characteristics of a model reproduced in research must always come accompanied by indicators of the degree of uncertainty they embed, which is just a consequence of uncertainty on the right model specification, in the sense described by Leamer (1978). To adequately represent that uncertainty, it is necessary to incorporate uncertainty on parameter values directly in the simulation exercise"* [Canova and de Nicolo (1995)].

De Jong, Ingram and Whiteman (1996) suggest calibrating by centering at a particular point estimate an arbitrarily diffuse Normal prior. After repeated simulations with different parameter vectors, we can compute either the size of the calibration tests, or the percentiles of the empirical distribution of the simulated statistic where the value of the statistic in actual data falls.

Actual and simulated data are used symmetrically, and one could either ask whether actual data could be generated by the model or viceversa, whether the simulated data is consistent with the distribution of the observed sample. The *confidence interval criterion*, and the *difference of means*, proposed by DeJong, Ingram and Whiteman (1996) measure the degree of overlap of the distributions of actual statistics and those obtained from the model. However, since empirical frequency distributions for the relevant statistics are, more often than not, asymmetric, with significant kurtosis, and even several local modes, their median should be used in computing these statistics. These authors show how this method can pinpoint aspects of the King, Plosser and Rebelo (1988) model that should be changed to improve its implications, in their confrontation with actual data.

6. Evaluating procedures for calibrated models[9]

Not even the need for model evaluation is uniformly accepted among calibrators. Interpreting calibration as estimation with no error forces an informal evaluation of the distance between actual and simulated statistics: once parameter values have been chosen, uncertainty comes only from the exogenous stochastic processes. Beyond that, the model establishes an exact relationship between endogenous variables and parameters, and we will not be able to conclude whether actual and simulated statistics are significantly different. Sometimes, alternative models are compared but, most often, only a subjective comparison is established between a few statistics, by essentially ad-hoc procedures, lacking a rigorous statistical foundation. *"No attempt is made to determine the true model. All models are abstractions and are, by definition, false"* [Kydland and Prescott (1982)]. This practice reflects an important position: *"...the trust a researcher has in an answer given by the model does not depend on a statistical measure of discrepancy, but on how much he believes in the economic theory used and in the measurement undertaken"* [Kydland and Prescott (1991)].

If we accept that the parameters in the model are estimated with sampling error, then it makes sense to use measures of dispersion for simulated statistics that reflect parameter uncertainty. A quadratic distance between the vector of statistics computed from actual data and those obtained by simulation will asymptotically follow a chi-square distribution with degrees of freedom equal to the number of overidentifying restrictions (number of statistics used in fitting, minus the number of estimated parameters) under the null hypothesis that, under the parameterization chosen, the model is true. That allows for a formal evaluation of the distance between the two vectors of statistics. The weighing matrix in the quadratic form should be the variance-covariance matrix of the vector of statistics being used, which can easily be estimated. This analysis is the basis for the *estimation by simulation* methodology, by which values for the free parameters are chosen iteratively.

It is specially interesting to compare statistics from the joint probability distribution of subsets of variables, like impulse response functions, *VAR* representations, cross-correlations or coherence functions. Examining autocorrelation functions, for instance, neglects possible cross effects. Furthermore, not being orthogonalized in any useful sense, they incorporate dynamical aspects which may be due

---

[9]The discussion of the Winter 1996 issue of the Journal of Economic Perspectives is specifically relevant to this section.

to joint fluctuations, not being necessarily specific of either one of the variables being considered.

Canova and Ortega (1996) classify the different approaches in model evaluation depending on whether they take into account the sample variability in actual data and the uncertainty in simulated data. Sampling error comes from the fact that we have a single realization of the vector stochastic process underlying the evolution of the real economy, while the latter is due to a less than perfect knowledge of the values of the parameters in the model. The error that any simulation method introduces, being an approximation of some kind to a theoretical model should also be taken into account, although it is usually ignored.

According to these authors, different evaluation strategies are:

1) informal procedures that compare, by mere inspection, the sample numerical estimation of a given statistic with actual data with that obtained for fixed parameter values and a realization of the exogenous shocks, as in Kydland and Prescott (1982). If these are not considered to be fixed, then the average of the empirical distribution obtained for that statistic with a large number of realizations of the shocks is used,

2) neither the sample variability in the data, nor the uncertainty in simulated data are taken into account, but rather, attention is placed in the statistical properties of the error u_t which needs to be added to the model to reproduce certain statistics, since the model is just an approximation to the stochastic data generating process (DGP) [Watson (1993)]. The best fit possible between model and data is attained when the variance of the error u_t, is minimized. Two measures of lack of fit are suggested by comparison: one, computes an R^2-measure for each variable in the model. Besides, this indicator can be employed over a range of frequencies, *e.g.*, those of a business cycle. The fitted time series for relevant variables can also be used to test how the model explains specific historical episodes,

3) sample variability in actual data and, sometimes, also the uncertainty in parameter estimation is considered to obtain a distance between model and data [Christiano and Eichenbaum (1992), Cechetti et al. (1993), and Féve and Langot (1994)]. Ortega (1996) takes the spectral density matrix for actual data as an empirical regularity which the model should replicate. The *average* distance between simulated and actual spectral density matrices, computed over a large set of simulations, is used to compute the *fit test* and the *comparison test.* Again business cycle frequencies could just be considered, and a matrix of weighs could be used to assign different importance to variables and frequencies. Experiments in Ortega (1996) are interesting because they show that two models which are

similar in their representation in the time domain may have very different properties at specific frequencies. The same procedure could be used to test the extent to which two theoretical models are different from each other, or to compare the properties of a same model under two different parameterizations,

4) uncertainty in simulated data is used to obtain a distance between model and data. Generally, the vector stochastic process for exogenous shocks is considered to be random, and the parameters constant, as in Gregory and Smith (1994), Cogley and Nason (1994). If we accept that the realizations for the exogenous shocks come from a known distribution for the stochastic processes, we can compute dispersion measures for simulated statistics. This approach centers on exogenous uncertainty, more than on uncertainty on parameter values. Sampling variability of simulated data can be used to evaluate the distance between statistics computed with actual and simulated data, computing the *quantil* of the empirical distribution of the simulated statistic corresponding to the numerical value obtained from actual data. The theoretical model is assumed to be the true *DGP*, so that its evaluation is based on the size of this calibration test. Alternatively, as we have already mentioned above, some key parameters (risk aversion, output elasticity of capital in the aggregate technology, etc.) can also considered to be stochastic (Canova (1994, 1995)) and calibration is made with a probability distribution for each parameter, rather than with a single parameter value.

5) sample variability in actual data and uncertainty in simulated data are taken into account for model evaluation as described in section 5. The Bayesian approach of De Jong, Ingram and Whiteman (1996) considers the processes for exogenous shocks as fixed, taking into account the variability in parameter values, focusing on the overlap between the distributions of the actual and simulated vector of statistics.

7. Stability

Stability is an important characteristic of a numerical solution, although it is frequently neglected. The analysis of stability is different in a stochastic, dynamic model than in its deterministic version. It is also different in nature in endogenous than in exogenous growth models. Besides, with the exception of the special cases that lead to linear models, we will be facing the stability of a non-linear, possibly stochastic system, for which general analytical conditions are unknown. In this respect, all we can do is build the best linear approximation to the model, and discuss stability in the approximated model. Using stability conditions makes an

important difference in terms of the behavior of the paths generated as solutions to the model.

Since the approximation must be made around some specific reference (usually the steady state of detrended variables in exogenous growth models), conclusions can only be local. This is important, since although that local analysis may be enough to characterize properties of fluctuations around steady state, it might not be enough when analyzing the effects of a policy intervention that takes the economy far from steady-state.

Some numerical solution methods are better equipped than others to handle stability. Lack of stability in the solution trajectories would show in the parameterized expectations approach as a difficulty to converge in the expectations function. In the log-linear approximation proposed by Uhlig (1999), for instance, the set of difference equations representing the linear approximation to the original model is solved by taking care of the stable and unstable roots, very much as it was shown in simple partial equilibrium contexts in Sargent (1979). Solving through an eigenvalue-eigenvector decomposition as in Sims (2000) and Novales et al. (1999), equilibrium realizations are obtained from the same equations that would be used by any other method, but stability conditions are added to the model. These come as orthogonality conditions between the vector of relevant variables and the eigenvectors associated to the unstable eigenvalues in the vector linear approximation to the original model. The approximated model is just used to compute stability conditions, which are hence only approximate for the nonlinear original problem. However, in obtaining the time series that solve the model, the original model is used, with no approximation involved. Stability conditions of this kind amount to relationships between forecast errors for the functions whose expectations appear in the model and exogenous structural perturbations, and they can be equivalently represented in the form of relationships between decision and state variables. Stability analysis provides also useful information on the degree of determination of the model. In some cases, there is a local indeterminacy in that there is more than one possible trajectory leading to the steady state. In some other cases, there is global indeterminacy, in that there is multiplicity of steady states. Computing the steady state of a model amounts to solving a set of nonlinear equations. Even though the system is complete, nonlinearity can produce lack of solutions, a single solution or a multiplicity of them. These will generally be local characteristics, arising for just some regions of the parameter space, that the researcher can characterize by the appropriate methods. The eigenvalue-eigenvector decomposition mentioned provides this in-

formation by producing either as many stability conditions as needed to solve for all the expectations and expectations errors in the model, or a smaller number of them.

A method that deals explicitly with stability conditions may be appropriate to provide initial conditions for a method involving less approximation error, as the parameterized expectations approach, for which convergence is often troublesome, unless appropriate initial conditions are chosen. Hence, there are several reasons why an explicit study of stability is useful.

The problem becomes harder when dealing with endogenous growth models. The implied time series have a deterministic trend which can easily be taken care of, very much as in exogenous growth models, but they also have a unit root. Transitory shocks will have permanent effects even after detrending, i.e., after discounting for the deterministic steady state rate of growth. So, there is a fundamental lack of stationarity which cannot be handled by just normalizing variables. That version of the model can generally be solved, but it will provide time series realizations for ratios of variables, not for their levels. But it is the levels of the relevant variables that are needed for welfare analysis of the kind used in policy evaluation, so this type of analysis may be hard to address in that framework. Novales et al. (1999) and Novales and Ruiz (2000) show that the model in normalized variables can however be successfully used to estimate stability conditions, since the same conditions apply to the original model, in levels of the variables. That allows for generating time series realizations for the relevant variables, which can then be used for policy analysis.

8. Some statistical issues

- *Non-parametric* Statistics is very much underutilized in economic data analysis, and it can be particularly useful when evaluating the results obtained from numerical solutions to macroeconomic models. Questions are usually brought up to the extent to which a given statistic (output variance, capital/labor ratio, etc.) behaves similarly under different parameterizations, different versions of a model economy, or in relation with its value in actual data. In most cases, a somewhat informal comparison of mean values across simulations for the statistic is used to discuss this issue. This is amazing, when non-parametric methods like Kolmogorov-Smirnov test, or chi-squared parametric tests for equality of distributions are available, which use all the information in the empirical distribution of the statistic of interest. The

use of more information will lead to a gain in efficiency, and more powerful tests. Furthermore, evaluating differences to the light of the comparison of the mean and variance of a given statistic across simulations, implicitly assumes the empirical distribution of the statistic to be Gaussian, when it is rather unlikely that Normality of the exogenous shocks will be preserved by a non-linear model. This renders non-parametric tests even more interesting, since they are distribution-free.

- The convenience of using an average statistic to compare models is often questionable[10]. Often that distribution is not symmetric, and can even have more than one mode, so that it is not at all appropriate to mechanically rely on comparing average values. It is much more appropriate to compute measures of distance between the empirical distributions emerging from two different models or two different parameterizations of a same model economy.

- Two more issues: a) homogeneity tests between the (unknown) theoretical value of a given statistic under two parameterizations, or two different models, should be one-tailed. In most cases, there are theoretical reasons to believe that a given parameter vector, or feature of a model, or policy rule, will be more likely to account for a given stylized fact. Following rough Bayesian recommendations, the researcher will be better off by initially establishing such a point of view before simulating, to then check whether the resulting average across simulations accords to his/her initial belief. If the prior belief is not corroborated, he/she will most likely prefer to run two-tail tests, b) the possible existence of extreme values of the statistics being studied, due to a (or a few) sample realization in the tail of the distribution is usually neglected. Fat-tailed empirical distributions are often obtained for a given statistic, and notorious deviations from Normality, like the possible existence of more than one local mode are common. Again, comparison of the whole empirical distributions obtained in two different models or parameterizations would avoid these biases.

- A theoretical model should not be expected to produce a time series for output, say, that matches the actual pattern observed in the *US* economy,

---

[10]A further, even if obvious, difficulty which is too often seen in work dealing with numerical solutions is that the standard deviation of that distribution is brought up to discuss whether or not the difference between two averages is significant, forgetting about the fact that it is the standard deviation of the mean which should be used to that effect.

for example. Most often, the adequacy of a model should be viewed as the degree to which the restrictions it introduces on the multivariate probability distribution of the set of relevant variables are also observed in the analogous probability distribution, estimated from actual data. However, there are cases in which the researcher is specifically interested in reproducing the shape of a given time series of actual data. An example is Watson (1993), in which the goodness of fit of the model is evaluated by computing how large a noise would have to be added to the implied time series, so that the same fluctuations as in actual data are obtained. We must be aware of the difference between these two approaches as well as the fact that, most often, we are in the former situation.

- Filtering numerical solutions is commonplace. This is specially surprising when working with models which display (exogenous or endogenous) steady-state growth, since the average rate of growth produced by the model should be one of the important features to mimic or be taken into account when calibrating the model. Even the solution to models that display no growth is filtered when comparing with actual data. It may be clear the need to filter the latter to make it comparable with the time series from the no-growth model, but it is much less clear the need to do so with a non-zero growth model.

- Numerical solutions open possibilities that have yet rarely been explored when characterizing a model's properties. Since the model can be simulated under any multivariate probability distribution for the set of exogenous shocks, the researcher can always characterize the properties of the model which are implied by each of the shocks impinging on the economy. To that end, it is enough to set variances for all other shocks as well as their correlations with the former shock, all equal to zero. We could find out that a given feature is due to a specific shock or, alternatively, that it can be produced by two of the exogenous shocks in the economy, although one of them produces it to a much lesser extent. What is true under a shock on preferences could be false when a single shock in productivity affects the economy. Clearly, the consideration of demand versus supply shocks is one of the interesting differences to be established. Policy analysis like Poole (1970), concluding what instrument the monetary authority should use as a function of what is the more important source of shocks in the economy, becomes a natural issue to discuss when numerical solutions are obtained.

We will get back to this issue in the next section.

- Along this line, the researcher can provide not only a given parameterization that is able to explain a given stylized fact but, rather, a whole continuum of combinations of parameters with that ability. As a simple example, it is interesting not only to know that a model can account for the relative volatility of investment to output in US data with a given set of variances for the productivity shock and the shock to money growth but also, to characterize the line representing combinations of both variances that share that property. In fact, it is hard to believe that a given model parameterization which is able to reproduce a given stylized fact should be taken as a satisfactory answer. At least, the whole curve describing the (maybe implicit) functional relationship between the statistic considered and the value of each of the relevant parameters should be exhibited in the paper.

- But what I consider the more important limitation in the way research based on numerical solutions is conducted is the fact that in most cases, the model considered incorporates the assumption that agents form their expectations rationally and yet, it is very infrequent that the numerical solution is tested for rationality. This type of test should clearly be a requirement before any model's characteristic are displayed since, if rationality was rejected, it is less than clear that the researcher should advance much further in presenting results. There are obviously several dimensions along which we can test for rationality: most solution approaches allow for a time series for each of the expectations in the model to be obtained, once we have time series realizations for all the relevant variables. These time series allow for the realized value for the nonlinear function inside the conditional expectation to be computed each period, so that an expectation error can also be obtained. The resulting time series for rational expectations errors should be autocorrelation-free, and uncorrelated with any variable in the information set at the time the expectation was formed. This second fact is the basis for the denHaan-Marcet test [Den Haan and Marcet (1994)], which in spite of being a significant addition to the validation of numerical solutions, is most often forgotten.

9. The analysis of the transition between steady states

One of the most interesting analysis that emerge naturally, once the model has been numerically solved, is that of the characteristics of the transition to its steady-state. This may arise either because the economy is initially outside steady-state, or because some structural change is introduced (it could be a policy intervention) altering the steady-state. In the absence of endogenous growth, and under stability, the economy will follow a path converging on average to its deterministic steady state, even though it could experience short-term fluctuations along the convergence process.

This type of analysis is crucial, among other things, for evaluating the possible effects of changes in policy rules, i.e., of policy interventions. Without characterizing the transition, the only evaluations possible are those based on the levels that the relevant variables achieve in the respective steady-states, before and after the policy intervention. One would like to obtain analytically the temporal trajectories that the relevant variables follow in their transition towards steady-state, and use them to aggregate over time the levels achieved by the policy-maker objective function, but the structure of the model will generally make that impossible, as already described in section 2. Then, only numerical evaluation is possible, although it can approximate arbitrarily well a missing analytical result. So long as the empirical distributions (or the posterior distributions, in a Bayesian analysis) of the main statistics have low dispersion, we will be able to make precise statements on the general characteristics of the alternative transition process. It is interesting to make such estimates to consolidate the welfare gains or losses along the transition with those attained in steady-state, to obtain a rigorous evaluation of the effects of a change in policy.

It is specially interesting, and somewhat discomforting, to see that in a high proportion of cases in which this type of analysis has been performed, the welfare comparison of alternative policies tends to be contrary to the conclusion that would be achieved by focusing just on steady-state. As a consequence, the optimal choice of policy will finally rely on aspects such as a) the speed of convergence to steady-state, b) the rate of time discount, c) the concavity of preferences, and so on. It cannot surprise that the conclusions of that research will often be different for economic structures that differ in values of these parameters, even though initially, they were not directly related to the policy issue under discussion. Normative policy analysis, taking explicitly into account these crucial dynamic issues, is an exciting research area for the future.

10. Have we changed the type of questions we ask?

Maybe the most important contribution of numerical solution methods is that they have, in fact, changed the way we analyze models. We now ask new questions and compare models in dimensions that were not even considered a few years ago. Policy analysis is also viewed from a new, richer perspective, which explains that policy design is currently the most active research area in macroeconomics. Some of these, by now familiar, questions are related to statistical characteristics of the model, like the volatility of a given variable, or cross-correlations between some variables, that would have been impossible to characterize analytically, even if such a solution would exist for the levels of the variables themselves. Some other times, the simulation process itself and the numerical observations it produces, suggest interesting model features that might have never shown up if an analytical answer was available.

10.1. Should economic policy be cyclical or counter-cyclical?

Some questions can only be answered through the numerical solution of a dynamic, stochastic model economy: suppose that fluctuations in the expenditures/output ratio can be interpreted as controlled deviations around a pre-announced target level. Should they then be correlated with exogenous supply shocks?

This can be discussed by solving the model under different correlations and computing the frequency distribution of welfare (or average welfare, if preferred) over a large number of realizations). The optimal correlation, i.e., the one that implies higher welfare, can be obtained. The welfare effect will generally depend on how the Government chooses to finance its budget, so we will have a welfare effect for each correlation level and each possible tax or combinations of tax rates (on consumption, capital or labor income, etc.) to be adjusted. This would have clear implications on the optimal way to conduct policy.

The previous analysis makes sense even if we believe that the random deviation of the expenditures/output ratio is beyond the control of the economic authority, since there will still be a welfare-maximizing correlation. Using actual data, we could separately identify supply and fiscal shocks, possibly through an structural *VAR* type of analysis. Estimating the empirical correlation between both perturbations, and knowing the taxes which are most often being adjusted, we can figure out the extent to which that correlation is close to the level that the model predicts as optimum. In the same vein, given the estimated correlation in actual data between a supply shock and the fiscal shock, interpreted now as a

control error in public expenditures, we will be able to discuss which tax should be adjusted over the business cycle so that the Government budget constraint holds every period, and welfare is highest.

10.2. The type of policy conclusions we reach

From the previous examples, it can be seen that the conclusions to policy analysis will often be of the sort: " *...if the main shocks in the economy are supply shocks, then it is better to implement a monetary policy aimed to maintain a given growth rate of money, while leaving interest rates to be determined in the market, the opposite is true if randomness enters mainly through the demand side*" or: " *...if the elasticity of intertemporal substitution is above a critical value, then it is better to adjust labor income taxes over the cycle while maintaining capital income taxes roughly stable, while the opposite is true if the elasticity of intertemporal substitution of consumption is below that value*".

We are increasingly going to reach such contingent conclusions. Some researchers view such relativity as a weakness of economic analysis, suggesting that it would better to discuss policy in simpler models, even if they miss some interesting feature, since they allow for neater conclusions. The opposite is, however, more likely to be true. Economists may have been too ambitious in attempting to reach statements that would always be true, regardless of the type of economy being studied. In characterizing optimal policy as a function of the structure of the economy (source of shocks, structural parameter values, etc..) we are borrowing from Bayesian statisticians, who sometimes aim at providing their readers with a sort of catalogue, that defines the mapping between input (the structure of the economy) and output (the specification of optimal policy). Did we really believe that a similar kind of policy would be optimum for a variety of widely different economies and for any conceivable policy environment?

10.3. Heterogeneous agents models

Considering an environment with heterogeneous agents is of utmost interest for almost any issue in economics, from characterizing how markets work, to establishing a ranking among alternative economic policies: to analyze the impact of possible liquidity constraints we want to consider a setup where some agents are restricted at a given period, while some other agents are not, adding to the model the conditions that define the flow between restricted and unrestricted agents from one period to the next. Something similar could be said about characterizing the

effects of asymmetric information: we then want to consider different types of agents, each one having access to different information sets, which might depend on their past investment on information. The distribution of information across agents then becomes endogenous.

Unfortunately, this dramatically increases complexity, since it implies following the time evolution of the distribution of the vector of state variables among all individual agents. For instance, in an economy with a single cumulative good, in which agents differ in their initial endowment, it will be crucial to know the distribution of income across the population at the beginning of each decision period. In an economy where the state variable is the productivity of each worker, the number of individuals with each level of productivity at a given period will be a key variable in determining prices and quantities. Agents with different levels of the state variables will follow a different behavior, so that the information on aggregate variables will not be enough to anticipate the evolution of the economy, being also necessary to know the distribution of the aggregate level of the state variable across the population. Calibrating these models requires an initial distribution, which must be chosen so that it matches the analogue distribution observed in actual data.

The time evolution of the distribution of a state is a key factor determining the equilibrium dynamics, and a very important characteristic of the model. Computing an equilibrium will require finding a fixed point each period, not only in the space of prices, but also in the space of distributions of the state variables subject to heterogeneity: such distributions become additional states. To the standard conditions in representative agent models, consistency conditions among the optimality conditions of individual agents and the behavior of aggregate variables must be added when characterizing the competitive equilibrium [Rios-Rull (1995)].

An additional difficulty arises from the fact that, with a rich enough variety of agents, there will be a high probability of corner solutions, i.e., solutions in which, in each period, some agent has a zero quantity of some good. This further complicates obtaining the numerical solution, since testing for the possibility of a corner solution can quickly become an extraordinarily complicated process. To alleviate somewhat the curse of dimensionality, it has been standard to consider situations in which the distribution of states among the population does not affect relative prices (like real salaries, or the real rate of return on capital). Diaz-Gimenez et al. (1992), and Diaz-Gimenez (1997) consider a situation in which the Government commits to maintaining a constant inflation rate, chosen before hand. In other cases, it is assumed that a set of economic policies is implemented

guaranteeing that relative prices will stay constant. Policy rules can then only depend on the aggregate state of the economy, but not on its distribution.

Currently, heterogeneous agent models are being proposed to try to explain some empirical observations that cannot be explained in the representative agent framework, like the equity-premium puzzle. Besides, such an attempt opens again questions which are new in spirit to economic analysis, like the cyclical behavior of income distribution, the relevance of learning processes, and so on. Using these models is important because the answer to some traditional questions may turn out not to be robust to the consideration of heterogeneous agents economies: for instance, Imrohoroglu (1992) shows that the cost of cyclical fluctuations in an economy with liquidity constraints can be at least three times as large as the similar cost in an economy with perfect insurance. A promising use of these models in asset pricing can be seen in Marcet and Singleton (1999) and in Heaton and Lucas (1996).

Rios-Rull (1992, 1994a, 1994b, 1996) has succeeded in approximating overlapping generation models to actual economies, considering that any single individual can live a large number of periods. The impact of demographic changes in aspects like capital accumulation and social insurance can then be analyzed in an adequate framework. Other business cycle issues like the volatility of hours worked, that we have mentioned at different points in this survey, can be nicely studied in such a model, and Rios-Rull (1992) has found different volatility across workers of different age as an equilibrium characteristic. How such a result mixes with the current structure of population to produce a specific result for volatility of aggregate hours is an illustration of the central point of this promising research agenda.

11. Conclusions

Macroeconomic analysis has come a long way since the optimizing behavior of economic agents was explicitly incorporated in models attempting to explain fluctuations and growth. Initial work dealt with dynamic but deterministic models, which admitted analytical solutions, at least for the case of homogeneous agents. Then, it became clear the need to use stochastic control techniques to try to explain fluctuations, and the linear quadratic setup, again allowing for an analytical solution, became the standard. Later on, the requirement to work in a general equilibrium environment led to models with no analytical solution. For a number of years, a variety of numerical methods to simulate model economies has been

proposed and reviewed in the literature. Currently, our problem is not how to solve a model but, rather, how to rigorously use simulation methods to characterize the implications of a theoretical economy. This is the issue we have been dealing with in this review paper.

The limitations in learning about model economies through simulation do not emerge from the solution methods used, but from the way they are implemented and, more importantly, from a poor statistical analysis of the implied results. In particular, the fact that simulation provides us with a full frequency distribution for each statistic of interest is generally left unexploited. Furthermore, the researcher should explicitly acknowledge parameter uncertainty when numerically solving a model. Summarizing his/her beliefs in the form of a probability distribution over a subset of parameters will translate into a frequency distribution for a given statistic very different from the one obtained when simulating from fixed parameter values. A further limitation comes from not testing the solution across two dimensions: first, with a few exceptions, which can be known in advance, the solution should be stable. This is often guaranteed in almost any method used to solve exogenous growth models, but it will become an important issue as more attention is being paid to analyze endogenous growth models outside their balanced growth paths. Secondly, testing the numerical solution for the restrictions implied by the usually maintained assumption of rational expectations should be a requirement in any research work involving simulation. Unfortunately, such test is almost never carried out.

We have started by explaining why numerical solutions are needed to analyze a wide class of interesting model economies, describing how the results obtained through numerical simulation should be presented. After defining calibration, we have discussed its limitations in actual practice, reviewing some interesting suggestions that have recently been made, based in a Bayesian approach to simulation. We have examined different approaches for evaluation of calibrated models. We have moved to the discussion of some statistical issues having to do with the analysis of the results, emphasizing the importance of guaranteeing stability of the implied solution. Finally, we have argued that the ability to simulate model economies has opened interesting questions and important research lines for macroeconomic analysis.

12. References

References

[1] Brock, W.A., and L. Mirman, 1972, "Optimal Economic Growth and Uncertainty: The Discounted Case", Journal of Economic Theory, 479-513.

[2] Canova, F., 1994, "Statistical Inference in Calibrated Models", Journal of Applied Econometrics, 9, S123-S144.

[3] Canova, F., 1995, "Sensitivity Analysis and Model Evaluation in Simulated Dynamic General Equilibrium Economies", International Economic Review, 36, 477-501.

[4] Canova, F. and G. de Nicoló, 1995, "The Equity Premium and the Risk Free Rate: A Cross Country, Cross Maturity Examination", CEPR working paper 1119.

[5] Canova, F. and E. Ortega, 1996, "Testing Calibrated General Equilibrium Models", manuscript.

[6] Cassou, S., 1995, "Optimal Tax Rules in a Dynamic Stochastic Economy with Capital", Journal of Economic Dynamics and Control, 19, 1165-1197.

[7] Cechetti, S.G., Lam, P. and N. Mark, 1993, "The Equity Premium and the Risk Free Rate: Matching Moments", Journal of Monetary Economics, 31, 21-45.

[8] Cogley, T. and J.M. Nason, 1994, "Testing the Implications of Long-run Neutrality for Monetary Business Cycle Models", Journal of Applied Econometrics, 9, S37-S170.

[9] Christiano, L.J., "Solving the Growth Model by Linear Quadratic Approximation and by Value Function Iteration", Journal of Business and Economic Statistics.

[10] Christiano, L.J. and M. Eichenbaum, 1992, "Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations", American Economic Review, 82, 430-450.

[11] Danthine, J.P. and J.B. Donaldson, 1995 "Non-Walrasian Economies", in Cooley,T.F.(ed.), "Frontiers of Business Cycle Research", Princeton, N.J., Princeton U. Press.

[12] DeJong, D.N., B.Ingram and C.H.Whiteman, 1996, "A Bayesian Approach to Calibration", Journal of Business and Economic Statistics, 14, 1, 1-9.

[13] Den Haan, W. and A. Marcet, 1994, "Accuracy in Simulations", Review of Economic Studies, 61, 3-17.

[14] Diaz-Giménez, J.,.Prescott, E.C., Fitzgerald, T. and F. Alvarez, 1992, "Banking in Computable General Equilibrium Economies." Journal of Economic Dynamics and Control 16: 533-559.

[15] Diaz-Giménez, J., 1997, "Uninsured Idiosyncratic Risk, Liquidity Constraints and Aggregate Fluctuations." Economic Theory 10: 463-82.

[16] Duffie, D. and K. Singleton, 1993, "Simulated Moments Estimation of Markov Models of Asset Prices", Econometrica, 61, 929-950.

[17] Eichenbaum, M., 1991, "Real Business Cycle Theory: Wisdom or Whimsy,", Journal of Economic Dynamics and Control, 15, 607-626.

[18] Féve, P. and F. Langot, 1994, "The RBC Models through Statistical Inference: An Application with french data", Journal of Applied Econometrics", 9, S11-S37.

[19] Friedman, M., 1953, " Essays in Positive Economics" , University of Chicago Press.

[20] Garcia-Mila, T., 1987, "Government Purchases and Real Output: an Empirical Analysis and Equilibrium Model with public Capital", Ph.D. dissertation, manuscript, University of Minnesota.

[21] Gregory, A.W., and G.W. Smith, 1993, "Calibration in Macroeconomics" in Maddala, G.S. (ed.) Handbook of Statistics, vol.11, Amsterdam, North Holland.

[22] Gregory, A.W., and G.W. Smith, 1994, "Calibration as Testing: Inference in Simulated Macro Models", Journal of Business and Economic Statistics, 9, 293-303.

[23] Hansen, L.P. and J.J. Heckman, 1996, "The Empirical Foundations of Calibration", The Journal of Economic Perspectives, 10, 87-104.

[24] Heaton, J., and D.J. Lucas, 1996, "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing", Journal of Political Economy, 104, 443-487.

[25] Imrohoroglu, A., 1992, "The Welfare Cost of Inflation Under Imperfect Insurance", Journal of Economic Dynamics and Control, 16, 79-91.

[26] King, R., Plosser, C., and S. Rebelo, 1988, "Production, Growth and Business Cycles: I", Journal of Monetary Economics, 21, 195-232.

[27] King, R., Plosser, C., and S. Rebelo, 1988, "Production, Growth and Business Cycles: II", Journal of Monetary Economics, 21, 309-342.

[28] Kydland, F. and E.C. Prescott, 1982, "Time to Build and Aggregate Fluctuations", Econometrica, 50, 1345-1370.

[29] Kydland, F. and E.C. Prescott, 1991, "The Econometrics of the General Equilibrium Approach to Business Cycles", The Scandinavian Journal of Economics, 93, 161-178.

[30] Kydland, F. and E.C. Prescott, 1996, "The Computational Experiment: An Econometric Tool", The Journal of Economic Perspectives, 10, 69-86.

[31] Leamer, E.E., 1978, "Specification Searches: Ad-hoc inference with Nonexperimental Data", New York: John Wiley & Sons.

[32] Lee, B.S. and B.F. Ingram, 1991, "Simulation Estimation of Time-series models", Journal of Econometrics, 47, 197-205.

[33] Lucas, R.E., Jr., 1980, "Methods and Problems in Business Cycle Theory", Journal of Money, Credit and Banking, 12, 696-715; also in Lucas, R. E., ed., "Studies in Business Cycle Theory", Cambridge, Mass.: Massachusetts Institute of Technology Press, 1981, 271-296.

[34] Lucas, R.E., Jr., 1976, "Econometric Policy Evaluation: A Critique" in vol.1, Carnegie-Rochester Series on Public Policy, Karl Brunner and Allan Meltzer (eds.), North Holland, 19-46.

[35] Lucas, R.E., Jr. 1987, "Models of Business Cycles", Basil Blackwell, Oxford, U.K.

[36] Lucas, R.E., Jr. and T.J.Sargent, 1981, "Rational Expectations and Econometric Practice", The University of Minnesota Press, Minneapolis.

[37] Marcet, A., 1994, "Simulation Analysis of Stochastic Dynamic Models: Applications to Theory and Econometrics", en Sims, C., ed., "Advances in Econometrics: Sixth World Congress of the Econometric Society", Cambridge, Cambridge University Press.

[38] Marcet, A. and K.J. Singleton, 1999, "Equilibrium Asset Prices and Savings of Heterogeneous Agents in the Presence of Incomplete Markets and Portfolio Constraints", Macroeconomic Dynamics.

[39] Marcet, A. and G. Lorenzoni, 1999, "The Parameterized Expectations Approach: Some Practical Issues", in "Computational Methods for the Study of Dynamic Economies", (Marimón, R. and A. Scott, eds.), Oxford U. Press, U.K., 143-172.

[40] McCallum, B.T., 1989, "Real Business Cycle Models", in Barro, Robert J., ed. Modern Business Cycle Theory, Cambridge, Mass.: Harvard University Press, 16-50.

[41] McGrattan, E., R. Rogerson and R. Wright, 1991, "Estimating the Stochastic Growth Model with Household Production", Federal Reserve Bank of Minneapolis.

[42] Novales, A., Domínguez, E., Pérez, J., and J. Ruiz, 1999, "Solving Nonlinear Rational Expectations Models by Eigenvalue-eigenvector Decompositions", in "Computational Methods for the Study of Dynamic Economies", (Marimón, R. and A. Scott, eds.), Oxford U. Press, U.K., 62-95.

[43] Novales, A. and J. Ruiz, 2000, "Dynamic Laffer Curves", manuscript, U. Complutense, Madrid.

[44] Ortega, E., 1996, "Assessing the Fit of Simulated Multivariate Dynamic Models", manuscript, Department of Economics, European University Institute.

[45] Pagan, A., 1994, "Calibration and Econometric Research", Journal of Applied Econometrics, 9, S1-S10.

[46] Phelps, E.S., ed., 1970, "Microeconomic Foundations of Employment and Inflation Theory", New York: Norton.

[47] Poole, W., 1970, "Optimal choice of Monetary Policy Instruments in a Simple Stochastic Macro Model", Quarterly Journal of Economics 84, 197-216.

[48] Rios-Rull, J.V., 1992, "Business-cycle Behavior of Life-cycle Economies with Incomplete Markets". Cuadernos Económicos del ICE, 51, 173-196.

[49] Rios-Rull, J.V., 1994a, "Population Changes and Capital Accumulation: the Aging of the Baby Boom", manuscript, University of Pennsylvania.

[50] Rios-Rull, J.V., 1994b, "On the Quantitative Importance of Market Completeness", Journal of Monetary Economics, 34, 463-496.

[51] Rios-Rull,J.V., 1995, "Models with Heterogeneous Agents" in Cooley,T.F.(ed.), "Frontiers of Business Cycle Research, Princeton", N.J.. Princeton U. Press.

[52] Rios-Rull, J.V., 1996, " Life Cycle Economies and Aggregate Fluctuations", Review of Economic Studies, 63, 465-490.

[53] Sargent, T., 1979, "Macroeconomic Theory", New York, Academic Press.

[54] Sargent, T., 1987, "Dynamic Macroeconomic Theory", Harvard University Press.

[55] Sidrauski, M., 1967, "Rational Choice and Patterns of Growth in a Monetary Economy", American Economic Review: Papers and Proceedings 51, 534-544.

[56] Sims, C.A., 1996, Macroeconomics and Methodology", The Journal of Economic Perspectives, 10, 105-120.

[57] Sims, C.A., 2000, "Solving Linear Rational Expectations Models", Journal of Computational Economics, forthcoming.

[58] Uhlig, H., 1999, "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily", in "Computational Methods for the Study of Dynamic Economies", (Marimón, R. and A. Scott, eds.), Oxford U. Press, U.K., 30-62.

[59] Vallés, J., 1997, "Aggregate Investment in a Business Cycle Model with Adjustment Costs", Journal of Economic Dynamics and Control, 21, 7, 1181-1198.

[60] Watson, M., 1993, "Measures of Fit for Calibrated Models" Journal of Political Economy, 1011-41.

[61] Whiteman, C., 1983, "Linear Rational Expectations Models: a user´s guide", U. of Minnesota Press, Minneapolis.