ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 67/68 – 2002

How Deep Are the Deep Parameters?

Filippo ALTISSIMO, Stefano SIVIERO,
Daniele TERLIZZESE *

ABSTRACT. – Policy evaluation based on the estimation of dynamic stochastic general equilibrium models rests on the assumption that a representative agent can be identified, whose behavioural parameters are policy-independent. Building on earlier work by GEWEKE [1985], this paper shows that in the presence of agents' heterogenity the representative agent, if any, is not structural, as its estimated behavioural parameters are not policy-independent. The paper identifies two sources of non-structurality and provides an example illustrating its effects.

Les paramètres fondamentaux, sont-ils vraiment fondamentaux ?

RÉSUMÉ. – Les analyses fondées sur l'estimation des modèles d'équilibre général dynamiques stochastiques supposent qu'il soit possible d'identifier un agent représentatif characterisé par des paramétres de comportement indépendants des politiques menées. Dans le même esprit que GEWEKE [1985], nous montrons qu'en cas d'hétérogéneité des agents un agent représentatif eventuel n'est généralement pas structurel, ses paramètres de comportement variant selon la politique suivi. Nous identifions deux causes a ce manque de structuralité et nous en illustrons les effets.

* F. ALTISSIMO : Bank of Italy, Research Department;
S. SIVIERO : Bank of Italy, Research Department;
D. TERLIZZESE : Bank of Italy, Research Department.
We thank the late A. ANDO, C. FAVERO, C. GIANNINI, M. LIPPI, I. VISCO, the participants to seminars at University of Rome "La Sapienza" and at CIDE and two anonymous referees for helpful comments. The usual disclaimers apply.

1 Introduction

If one were to name the single most entrenched commandment of the young scholar approaching macroeconomic policy evaluation there is little question that the mantra "Thou shalt beware of the *Lucas critique*" would spring to mind. Indeed, since *Lucas* published his rightly famous critique (LUCAS [1976]), the usefulness of traditional 'structural' econometric models as a means for macroeconomic policy selection has been seriously questioned. It scarcely needs to be recalled what is the content of the *Lucas critique*: traditional econometric models *–ie*, models embodying decision rules with constant coefficients– fail to recognise that, due to the need to anticipate the future course of policy variables, the coefficients of (rational and forward looking) agents' decision rules depend on the parameters that govern the policy stochastic process, as well as on the primitive (or 'deep') parameters that characterise tastes and technology. As a result, traditional 'structural' models do not capture the actual structural parameters and the estimated coefficients are subject to variability in the presence of shifts in the policy rules.

While the stringent logic of the critique undoubtedly contributed to its success,[1] it can be argued that had the critique been merely a negative one it would have been dismissed as paralysing, and consequently neglected or downgraded, as SIMS [1982] suggested, to the rank of a cautionary footnote. Instead, much of the strength of the *Lucas critique* lies in its constructive content, hinted at in Lucas' original paper and more fully developed in later works by HANSEN and SARGENT [1980, 1981] and SARGENT [1981]. Indeed, the critique does not leave the practitioner at sea, as it comes with a 'recipe' for proper (econometric policy evaluation) behaviour: *(a)* solve the agents' optimisation problem and derive the explicit expressions of their decision rule coefficients as a function of deep and policy parameters; *(b)* estimate the coefficients of the decision rules together with the coefficients of the policy process, disentangling the dependence on deep parameters from that on policy parameters; *(c)* recompute the coefficients of the decision rules, taking into account the change in policy parameters, while keeping the deep parameters unchanged. This prescription, as conceptually simple as it is technically demanding, would guarantee that the simulated response to a policy shock takes into account the purposeful response of private agents, thus improving on the naïve prediction of the traditional approach.

Recognising the importance of the constructive side of the *Lucas critique*, however, is like being kind to be cruel. *Lucas*' recipe can be applied by macroeconometricians only if a representative agent (RA henceforth) is

1. In truth, the sharp logic of the critique is not as compelling as it looks at first sight. On the one hand, general rational expectation models are plagued by indeterminacy of the equilibria, so that Lucas-proof constant parameter optimal decision rules can be found that do not violate the rationality of expectations (FARMER [1991]). On the other hand, opposite to forward-looking, backward-looking behaviour might prevail in practice, as agents might simply adopt a "wait, see, react" strategy. Most importantly, whichever behaviour does in fact prevail can be statistically tested, and the *Lucas critique* is generically refutable (HENDRY [1988]; FAVERO and HENDRY [1992]).

warranted, since a single aggregate time series would not allow many idiosyncratic deep parameters to be recovered.[2] But the strength of the recipe – which we shall label as the representative agent *cum* rational expectations (*RARE*) approach– hinges on the possibility of finding an adequate representative agent. Which is where the troubles begin.

In a very remarkable paper KIRMAN [1992] gave an impressive list of the pitfalls presented by the notion of RA.

First, *Debreu, Mantel* and *Sonneschein*'s theorem can be taken as an "impossibility theorem": in general equilibrium, the RA does not (in general) exist. This follows from the simple fact that, as aggregate excess demand need only satisfy continuity, homogeneity and *Walras*' law, it will not in general satisfy the weak axiom of revealed preferences; therefore there will be no utility function that generates the given aggregate excess demand. Even when confined to the special cases in which an RA can be found, examples can be constructed where the RA does not represent, in welfare terms, the agents whose aggregate actions it reproduces.

Most importantly, the RA might be 'non-structural'. More explicitly in presence of agents's heterogeneity, the RA that is appropriate (if it exists) before a given policy shock could well be different from the RA that will be appropriate after the shock (GEWEKE [1985]); clearly this would imply that the response of the aggregate economy to the shock would be misrepresented by the response of the RA recovered from data that do not include the shock. But this squarely contradicts the presumption underlying the $RARE$ approach and implies that it suffers from the same logical difficulty that was originally imputed to the traditional approach.

Geweke's point has been happily ignored.[3] Possibly, this is because (*a*) his argument is not really general, as it is shown by way of example; (*b*) moreover, the example chosen by *Geweke* to highlight the non-structurality of the RA does not belong to the class of models to which the *Lucas critique* was originally directed, as it did not require future expectations regarding policy variables; (*c*) no explicit interpretation of the lack of structurality of the RA is provided.

In this paper we try to remedy to those shortcomings.

In the first part of the work, we move some steps in interpreting the nature of non-structurality (partially addressing shortcoming (*a*) and (*c*), Sections 2 and 3). First, abstracting from estimation issues, non-structurality arises when aggregating common exogenous variables (typically, policy variables) whose coefficients in the individual decision rule are functions which combine deep and policy parameters, and derives from the non-aggregability of those functions. Noteworthy, we argue that this condition will be satisfied in almost all the models to which the *Lucas critique* applies, thus supporting the generality

2. It should be stressed that this paper is concerned with macroeconomics. Many of the problems that we shall highlight would not arise if the focus of the analysis were the explanation of individual behaviour, using individual rather than aggregate data. A survey of results on aggregation and microfoundations of dynamic macroeconomics can be found in FORNI and LIPPI [1997].

3. To consider just one representative example, the recent macroeconomic textbook by TURNOVSKY [1995] – largely centred around the RA device – while quoting *Kirman* paper as an example of critical reaction to the RA paradigm, does not mention the possible non-structurality of the RA, which is instead very much stressed by *Kirman*.

of *Geweke*'s argument. Secondly, focusing on the estimation issues allows us to identify a second source of non-structurality, logically distinct from the non-aggregability mentioned above, resulting from imposing on the aggregate data a misspecified model, patterned on the model valid for the individual agents.

In the final part of the work, we show that non-structurality arises in a model that is a prototype of those supported by the $RARE$ approach (addressing shortcoming *(b)*). We set up an 'experimental world' populated by heterogeneous, infinitely living, rational economic agents, who make their decisions taking into account an exogenous (stochastic) process governing a (pay-off relevant) policy variable, as well as a random idiosyncratic shock. Following the spirit of the $RARE$ approach, we then recover from the parameters of an aggregate decision rule obtained by constrained maximum likelihood estimation the (supposedly) deep parameters of the representative agent (the agent whose choice would be equal to the average of the individual agents' choices) and use that knowledge to predict the response of the economy to a policy regime shift. The comparison shows that the RA response can be different from the true one. The reason for the difference lies, as anticipated, in the non-structurality of the *RA*: the deep parameters recovered before the policy shock are different from those recovered afterwards. To put it differently, the deep parameters of the RA are not deep at all, since they are a function of the policy parameters.

2 Non-Structurality Resulting from Non-Aggregability

The traditional approach to policy evaluation, in its simplest possible form, relies on running a regression $y_t = fx_t + u_t$ of an (aggregate) action (y) on a policy control (x) and, given the estimated value of the coefficient f, assess the effect on y of a given change in x (u is a random error). The *Lucas critique* to this approach stresses that the relationship between y and x is not structural, as in turn it emerges from the solution to an intertemporal decision problem, in which the conditional expectation of a pay-off function (U) is maximised subject to the stochastic process governing the policy variable:

$$\max_{\{y\}} E_0 \sum_{t=0}^{\infty} U(y_t, x_t, \varepsilon_t; \beta) \quad (1)$$
$$\text{s.to } x_t = g(x_{t-1}, \eta_t; \rho),$$

where: E_0 is the expectation operator conditional on the initial information, (β and ρ are the parameters, ε and η are the random components, of the pay-off function and of the policy process, respectively).

A solution to (1) can often be expressed (after linearisation, and neglecting dynamics) as $y_t = f(\beta, \rho)x_t$, *ie*, with the 'coefficient' f being a function of the 'deep' parameters β and ρ. Hence, the correct response of y to a change in x or, more precisely, to a change in ρ generating a unitary change in x, is not simply f (as would be implied by the traditional approach) but rather $(f + \frac{\partial f}{\partial \rho}x)$.

The solution proposed by *Lucas* (what we have labelled the $RARE$ approach) is to first solve problem (1), to establish how ρ and β enter the function f; this determines a cross-equation constraint involving the regression of y on x and that of x on its own past (more generally, the policy process), and the econometrician should proceed in estimating simultaneously these two equations, imposing the constraint. This would allow to recover the 'deep' value of β, so that the correct policy evaluation can be performed recomputing $f(\beta,\rho)$ for each alternative value of ρ.

Suppose now that there are N heterogeneous agents (each with the same pay-off function, up to a different value of β), solving the above problem, yielding $y_t^i = f(\beta_i,\rho)x_t$, $i = 1,2,...,N$. The average aggregate action, $y_t = \frac{1}{N}\sum_{i=1}^N y_t^i$, is interpreted as resulting from the decision of a RA, also assumed to have the same pay-off function (and therefore computing the same decision rule), whose "deep" parameter β^{RA} is such that $y_t = f(\beta^{RA},\rho)x_t$. Thus, β^{RA} can be found as the solution to:

$$f\left(\beta^{RA},\rho\right) = \frac{1}{N}\sum_{i=1}^{N} f\left(\beta_i,\rho\right), \tag{2}$$

for a given value of ρ. Here and in the following we shall assume that the hypotheses of the implicit function theorem are satisfied, so that implicit equations like (2) above always have a solution.

Once a value β^{RA} has been determined, according to the $RARE$ approach we could predict how the economy would react to a policy shift from ρ_0 to ρ_1 (say) by changing the appropriate coefficient in the RA decision rule from $f\left(\beta^{RA},\rho_0\right)$ to $f\left(\beta^{RA},\rho_1\right)$. For this change in the value of the coefficient to be correct it must be that the value of β^{RA} that solves equation (2) for ρ_0 also solves it for ρ_1 or, more generally, that β^{RA} is not a function of ρ. Equivalently, it must be that β^{RA} is also a solution to:

$$f_\rho\left(\beta^{RA},\rho\right) = \frac{1}{N}\sum_{i=1}^{N} f_\rho\left(\beta_i,\rho\right), \tag{3}$$

where f_ρ is the partial derivative with respect to ρ. Indeed, if β^{RA} is not a function of ρ we can simply take a derivative with respect to ρ on both sides of (2) to get (3). By applying the implicit function theorem to (2) and using (3) we immediately obtain that the derivative of β^{RA} with respect to ρ is identically zero.

However, it is readily apparent that the system (2) and (3) will in general have no solution, since it is a system of two equations and one unknown. More precisely, the system (2) and (3) will have a unique solution for every ρ if and only if the function $f(\beta,\rho)$ is *separable, ie,* is of the form:

$$f(\beta,\rho) = h_1(\rho) + h_2(\rho)g(\beta) \tag{4}$$

for some functions h_1, h_2 and g. A proof of this claim can be found in ALTISSIMO *et al.* [1999]. Note that this characterisation is reminiscent of the conditions proposed by GORMAN [1953] for proper aggregation of individual demand functions.

Given that any aggregate decision rule which is potentially subject to the *Lucas critique* must involve at least one coefficient that can be expressed as a function of deep and policy parameters, this result provides a general necessary and sufficient condition for the viability of the $RARE$ approach.

A question naturally arises here: Is separability likely to hold in the type of models that are potentially affected by the *Lucas critique, ie,* models that result in decision rules whose coefficients are a function of both deep and policy parameters? Typically, a problem such as (1) delivers an *Euler* equation –linear or linearised– whose solution involves eliminating the future expected values of the policy variable. This in turn inextricably bounds together the subjective evaluation of the future and the objective law of motion of the policy variable, so that the non-separability of deep and policy parameters will result. To take just one example, suppose the (possibly linearised) *Euler* equation is of the form:

$$y_t^i = \beta_i E_t(y_{t+1}^i) + x_t + u_{i,t}, \tag{5}$$

and suppose also that $x_t = \rho x_{t-1} + \eta_t$, with $\beta,\rho < 1$ and $u_{i,t}$ is an idiosyncratic white noise. Then, the bounded solution for y_t^i will take the form:

$$y_t^i = f(\beta_i,\rho)x_t = \frac{1}{1-\beta_i\rho}x_t + u_{i,t}. \tag{6}$$

Clearly, $f(\beta,\rho)$ is a non-separable function. A similar conclusion would hold for any autoregressive policy process and in general in all cases in which the agent must discount future policies in order to solve his forward looking problem. This means that the non-structurality of the RA, far from being a non-generic *curiosum,* is almost sure to arise in models that are prone to the *Lucas critique* as soon as agents' heterogeneity is allowed for.

While going some way towards generalising *Geweke*'s point, it should be acknowledged that our conclusions rest on a rather special set-up: *(i)* the solution of the agents' problem is assumed to be linear – or conveniently linearisable – in the policy variable; *(ii)* there is no dynamics, and as a result the problems of estimation have been sidestepped. The latter point will be taken up again in the example studied in Section 3.

As to the issue of linearity, the question is to whether results concerning the non-structurality of the RA in a linear framework are of any general relevance and, more specifically, are robust to non-linear generalisations. There are a number of reasons to conclude that they are indeed both.

First and most importantly, if there are aggregation problems in a linear framework, they can only become worse in a non-linear one, according to the most obvious of Jensen's inequality-type of arguments: as the sum over i of a non-linear function of a variable y_t^i, $i = 1,2,...,N$, is different from the same non-linear function of the sum over i of the variables y_i, there is no way at the aggregate level –that is for the RA– to preserve the same functional form valid at the individual level.[4] In fact, moving away from a linear-quadratic setting, the agents' optimisation problem would in general not even have a closed form solution. In these circumstances, either we linearise the original problem around the steady state –thereby returning the linear-quadratic world– or we try to estimate –by a generalised method of moments or by simulation[5]– Euler conditions that, being non-linear, suffer from problem of the Jensen's inequality type mentioned above.

Secondly, the linear-quadratic setting encompasses some interesting models that have been proposed to study consumption and permanent income, the dynamic demand for factors of production and many other issues. Even if issues involving risk aversion cannot be seriously addressed in that framework, there is nevertheless a large class of problems that can be explored fruitfully under the simplifying assumption of linearity.

Thirdly, in keeping with the methodological principle of choosing the 'battleground' most favorable to the target of our criticisms, the linear-quadratic framework is the obvious choice, given the extensive use that contributors to the $RARE$ approach have made of it. So there are a number of reasons to conclude that the results concerning the non-structurality of the RA in a linear framework are of general relevance and, more specifically, are robust to non-linear generalisations.

3 Non-Structurality Resulting from Misspecified Dynamics

The analytical results presented in the previous Section, abstracting from the dynamic component of the agents' decision rules, allowed us to set aside estimation issues, and to isolate just one source of non-structurality affecting the $RARE$ approach. However, estimation issues need to be confronted, if anything because the approach advocates the use of sophisticated estimation

4. There is, of course, the possibility of approximate aggregation. It might be argued that the RA is indeed non-structural, but only slightly so. Given that the RA can be grossly non-structural in a linear framework, we find this argument unconvincing.

5. It is possible to estimate the deep and policy parameters of stochastic general equilibrium models by a generalized method of moments or by simulation, using either the simulated MLE or simulated method of moments or indirect inference. The GMM method exploits orthogonality conditions originating from the maximization problem of the agent to create a quadratic function which has to be minimized with respect to the parameters. The estimation by simulation is based on the possibility of simulating the desired model, so as to use the simulated data to generate proper moment conditions to be used in the estimation. A general survey of the results of estimation by simulation is given by GOURIEROUX and MONFORT [1996].

techniques. Indeed, we shall argue (along with FORNI and LIPPI [1997]) that when dynamics is allowed for, the implementation of the $RARE$ approach suffers from one additional source of non-structurality, resulting from the misspecification of the estimated (aggregate) model.

The implementation of the representative agent approach rests on the requirement that the dynamic structure of the aggregate model must be the same as that of individual decision rules. However, in presence of neglected agents' heterogeneity this, as it is well known, forces a dynamically misspecified model onto aggregate data. In turn, this will induce a dependence of the estimated parameters of the misspecified model on all the parameters of the true underlying model; in particular, a dependence on the policy parameters. This kind of non-structurality, resulting from dynamic misspecification, is logically different from that identified in the previous Section, resulting from non-aggregability. In practice, both kinds will be simultaneously affecting the results of empirical estimation of the deep parameters but, except in special cases, it will not be easy to disentangle them.

Consider a modified version of the example presented in the previous Section, where we assume that the agent faces the same first order condition as in (5) but the policy variable x_t is modelled as $x_t = \mu + \eta_t$, and $u_{i,t}$ is an idiosyncratic autoregressive process, modelled as $u_{i,t} = \alpha_i u_{i,t-1} + \varepsilon_{i,t}$ with $\alpha_i < 1$ and $\varepsilon_{i,t}$ is a zero mean idiosyncratic shock. The decision rule of agent i is:

$$(1-\alpha_i L)y_t^i = (1-\alpha_i L)(\frac{\mu}{1-\beta_i} + \eta_t) + \frac{1}{1-\beta_i\alpha_i}\varepsilon_{i,t},$$

which is an ARMA(1,1) process.

The assumptions imply different dynamic responses of the agents to the same policy shock. Hence, the dynamic behaviour of the aggregate differs from each of the individual ones. Assuming for simplicity the presence of only two agents in this economy, the dynamic specification of the aggregate, $y_t = y_{1,t} + y_{2,t}$, is:

$$\begin{aligned}(1-\alpha_1 L)(1-\alpha_2 L)y_t &= (1-\alpha_1 L)(1-\alpha_2 L)(\frac{\mu(2-\beta_1-\beta_2)}{(1-\beta_1)(1-\beta_2)} + 2\eta_t)\\ &\quad + (1-\alpha_2 L)\frac{1}{1-\beta_1\alpha_1}\varepsilon_{1,t} + (1-\alpha_1 L)\frac{1}{1-\beta_2\alpha_2}\varepsilon_{2,t}.\end{aligned}$$

Thus, if $\alpha_1 \neq \alpha_2$, the aggregate dynamic is more complicated than the individual ones, following an ARMA(2,2). By imposing that the dynamic specification of the aggregate model be the same as that of the individual ones results in dynamic misspecification. This form of mispecification is different from that arising from the non-separability of the coefficients of the policy variable.

Generalising our example, we can write the result of the intertemporal maximising behaviour of the i agent, directly in the case of a linear-quadratic problem or after proper linearisation, as an ARMAX model:

$$\Phi_i(L)\, y_t^i = \Psi_i(L)\, \varepsilon_{i,t} + \Theta_i(L)\, x_{t-1}, \tag{7}$$

where, in keeping with the notation of the previous example, y_t^i is the action of agent i in time t, $\varepsilon_{i,t}$ is the idiosyncratic shock component, x_t is the policy variable, and Φ, Ψ and Θ are lag polynomials, respectively of (finite) order p, q and l, and they are all function of policy parameters ρ and individual deep parameters β_i. Note that equation (7) implicitly defines a set of non linear cross equation restriction as the same deep and policy parameters are arguments of all lag polynomials.

Assuming a finite number of agents N in the economy, then the DGP that generates the aggregate data will diverge substantially from each DGP underlying individual decision rules, as the sum of ARMAX processes is still an ARMAX but of different order lag. In particular, if the individual decision rule is given by (7) then the aggregate (by average) over N agents is:

$$\begin{aligned} &\prod_{i=1}^{N} \Phi_i\,(L)\, y_t \\ &= \frac{1}{N} \sum_{i=1}^{N} \prod_{j \neq i} \Phi_j\,(L) \left(\Psi_i\,(L)\, \varepsilon_{it} + \Theta_i\,(L)\, x_{t-1}\right), \end{aligned} \tag{8}$$

which is still an ARMAX process but of order $(N \times p, (N-1) \times p + q, (N-1) \times p + l)$.

For simplicity, we consider only the case of a finite number of agents in order to assure that the process of the aggregate dynamic belongs to the same class of model as the agents' one even if the aggregate is characterised by a much articulated dynamic. We do not want to address the extreme case of the aggregation of an infinite cross-section of linear stationary ARMAX model. In that case the claim of dynamic misspecification of the aggregate model would be even stronger. In fact, it is even possible that the aggregate process does not belong to the same class of models of the individual processes and can exhibit peculiar dynamic properties (the aggregate process can exhibit long-memory while the individual processes are all of the short-memory type, as pointed out by GRANGER [1980] and LIPPI and ZAFFARONI [1999]).

When all the roots of the autoregressive polynomials Φ are equal across agents, then the aggregate model is no longer dynamically misspecified and reduces to:

$$\begin{aligned} &\Phi\,(L)\, y_t \\ &= \frac{1}{N} \sum_{i=1}^{N} \left(\Psi_i\,(L)\, \varepsilon_{it} + \Theta_i\,(L)\, x_{t-1}\right). \end{aligned}$$

This still does not guarantee that the estimated "deep" parameters be policy-invariant, for two reasons. First, the coefficients of the policy variable are in general non separable, as discussed above. Secondly, the cross-equation restrictions involving the coefficients of the MA terms and of the policy variables in the individual models are in general not valid in the aggregate model.

It is worth stressing that the difficulty with aggregation resulting from the presence of idiosyncratic dynamic terms will almost invariably be present, as the latter are an intrinsic component of the decision rule of forward-looking rational agents, as long as the problem is genuinely intertemporal –*ie*, it cannot be reduced to a sequence of unrelated static problems– which in turn is a precondition for the parameters of the policy process to show up in the coefficients of the decision rule: in a word, for the *Lucas critique* to apply. We can then conclude that whenever a representative agent decision rule mirroring the individual one is estimated on the aggregate data, the resulting model will be misspecified.

The misspecification implies that the RA is non-structural. More precisely, the estimated parameters of equation (7) will converge to their pseudo-true values under the actual DPG, given by equation (8). As a result, the pseudo-true values of the estimate will be a function of all the parameters of the aggregate DGP (8). Hence, the estimated parameters that the approach identifies as structural are not independent of policy shifts: once again, the deep parameters are in fact not deep at all, so that any prediction about the aggregate reaction to a change in policy is bound to be biased.[6]

The non-structurality of the RA induced by misspecification is then a pervasive characteristic of the $RARE$ approach when implemented in 'real life' circumstances, as it neglects the fact that the aggregate DGP does not mimic any individual decision rule. In addition it should not be forgotten, of course, that the other cause of non-structurality –which was shown in Section 2 to be the non-separability of the function of deep and policy parameters representing (at least one of) the coefficients of the individual decision rule– is still at work, even in a stochastic world.

For completeness, it should also be mentioned that heterogeneity of agents implies in general that the non-linear cross equation restrictions that hold at the micro level are not necessarily consistent with the aggregate data. By imposing these constrains in estimation we introduce a further source of misspecification, which is added to the previous one. More explicitly, imposing on the aggregate data a model with the same structure (functional form, variables, lags and errors stochastic process) implies that we are restricting the analysis to a class of models that does not include the data generating process of the aggregate data; if we also impose cross equation restrictions which are not actually satisfied, this will narrow the search to an even smaller class of models. This, in a way, produces a model which is even more misspecified.

Moreover, it is interesting to assess whether the said non-structurality is quantitatively important. In this remainder of the paper we shall present numerical results that confirm the non-structurality of the estimated 'deep' parameters of the RA.

This we shall do by setting up an 'experimental world' populated by heterogeneous, infinitely living, rational economic agents, who make their decisions by optimising over the infinite future the (rationally) expected value of their objective function and where policy regime shifts are publicly announced. In this 'ideal' world the *Lucas critique* unquestionably applies and his 'recipe' to

6. It is possible to show that in the general case of linear decision rule, with a large number of agents, the pseudo-true value of the estimates of the representative agent model are a function both of the policy parameters and of a measure of the heterogeneity of the agents, given by the variance of the distribution of the taste parameters across agents.

solve the problem (the $RARE$ approach) promises to be most helpful. Yet we will show that such approach suffers from severe shortcomings, of the same logical nature of those plaguing the traditional approach that it was meant to replace.

4 The Two Sources of Non Structurality Operating Together

4.1 The Set-Up of the Experiment

We consider N heterogeneous agents, namely firms, facing a standard, well known capital accumulation problem subject to idiosyncratic productivity shocks, taking as given the (stochastic) process governing the rental rate (HANSEN and SARGENT [1980]; see also INGRAM [1995], upon which our discussion of the model properties is largely based). For each agent the decision rule that solves the problem – which the assumptions made guarantee to have a closed form – is simulated, generating a time series of values for the (individual) capital stocks. Those time series are averaged across agents providing, together with the time series for the rental rate, the aggregate data that are the input for the estimation.[7] The latter is carried out following the $RARE$ by restricted maximum likelihood.

Before presenting the results of the experiment, and in order to introduce the notation, it is useful to recall briefly the main features of the underlying model. Let firm i, endowed with a linear-quadratic production function and subject to quadratic adjustment costs, choose the capital stock so as to maximise the present discounted value of its profits; *ie*, let it solve, under the appropriate transversality condition, the following problem:

$$\text{(9)} \quad \max_{k} E_0 \sum_{t=0}^{\infty} \beta_i^t \left[\left(\gamma_i + a_t^i + r_t \right) k_t^i - \frac{1}{2} \vartheta_i (k_t^i)^2 - \frac{1}{2} \delta_i (k_t^i - k_{t-1}^i)^2 \right],$$

where: E_0 is the expectation operator conditional on the initial information;

β, γ, ϑ and δ are the discount factor and the technology parameters, respectively;

r is the rental rate, exogenously given to the firm;

a is a productivity shock.

In the following, we shall assume δ, the parameter on which adjustment costs depend, to be the same for all firms; the rationale for this assumption will be made clear below. All other parameters and productivity shocks are assumed, in the most general case, to be idiosyncratic to firm i.[8] To complete

7. In line with *Lucas*' prescription, we perform the estimation on data that pertain to the same policy regime.

8. We have assumed that the only source of heterogeneity is different parameter values. This rules out another potentially very relevant source of heterogeneity, namely the possibility that individual payoff functional forms differ.

the description of the model, the stochastic processes underlying r and a have to be specified. The rental rate is the policy variable, for which the following stochastic process is assumed:

$$(10) \qquad r_t = \mu + \rho r_{t-1} + \varepsilon_t, \text{with } \varepsilon \sim NID(0, \sigma^2) \text{and } |\rho| < 1.$$

The process generating the idiosyncratic productivity shocks for firm i is also assumed to be an autoregressive process of order one:

$$(11) \qquad a_t^i = \eta_i a_{t-1}^i + \xi_{it}, \text{with } \xi_{it} \sim NID(0, \tau_i^2) \text{and } |\eta_i| < 1.$$

Furthermore we assume that all the stochastic components are uncorrelated, both serially and among themselves. Note that the linear-quadratic set-up of the model guarantees that a closed-form expression for the decision rule can be easily computed. Following INGRAM [1995], the *Euler* condition characterising firm i optimal behaviour is:

$$\delta \beta_i E_t k_{t+1}^i - [\delta(1+\beta_i) + \vartheta_i] k_t^i + \delta k_{t-1}^i + \gamma_i + a_t^i - r_t = 0$$

and the optimal decision rule for firm i can be written as:

$$(12) \qquad \begin{aligned} k_t^i = \lambda_i k_{t-1}^i - \frac{\lambda_i}{\delta} (\frac{r_t}{1 - \beta_i \lambda_i \rho} - \frac{a_t^i}{1 - \beta_i \lambda_i \eta_i} \\ - \frac{\gamma_i}{1 - \beta_i \lambda_i} + \frac{\beta_i \lambda_i \mu}{(1 - \beta_i \lambda_i \rho)(1 - \beta_i \lambda_i)}), \end{aligned}$$

where λ_i is the stable root of the following equation:

$$(1 - \lambda_{1i} L)(1 - \lambda_{2i} L) = 1 - (1 + \frac{1}{\beta_i} + \frac{\vartheta_i}{\delta \beta_i}) L + \frac{1}{\beta_i} L^2$$

and L is the lag operator.[9] Equations (12) and (10), one for each firm, together with equation (11), constitute our data generating process (DGP). Note that here the two sources of non structurality are playing together a role given that the coefficient of the policy variable is non separable and the degree of dynamic of the dependent variable is idiosyncratic, *ie*, λ_i and η_i are different across agents.

To provide a benchmark for the following analysis let us now briefly recall the steps needed to recover the deep parameters. To this end, after dropping the index i (*ie*, assuming that all agents are alike), standard algebraic manipulations yield the following VAR:

---

9. Equation (12) can easily show the main point of the *Lucas critique*: a change in r_t, if caused by a change in the policy parameter ρ, modifies the parameters of the decision rule, which therefore cannot be considered as structural. If, on the other hand, changes in r_t are caused by a particular realisation of the stochastic disturbance, then no change in the decision rule parameters is expected to take place (on the distinction between these two sources of change in policy see SIMS [1982]). The reduced form decision rule is therefore not equipped to make correct inferences concerning the outcome of a change in policy.

$$
\begin{aligned}
\begin{bmatrix} k_t \\ r_t \end{bmatrix} = & \begin{bmatrix} \lambda + \eta & \dfrac{\lambda(\eta - \rho)}{\delta(1 - \beta\lambda\rho)} \\ 0 & \rho \end{bmatrix} \begin{bmatrix} k_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} -\lambda\eta & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} k_{t-2} \\ r_{t-2} \end{bmatrix} \\
& + \begin{bmatrix} \dfrac{\lambda(1 - \eta)}{\delta(1 - \beta\lambda)}(\gamma - \dfrac{\beta\lambda\mu}{(1 - \beta\lambda\rho)}) - \dfrac{\lambda\mu}{\delta(1 - \beta\lambda\rho)} \\ \mu \end{bmatrix} \\
& + \begin{bmatrix} \dfrac{\lambda(\eta - \rho)}{\delta(1 - \beta\lambda\eta)} & -\dfrac{\lambda}{\delta(1 - \beta\lambda\rho)} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_t \\ \varepsilon_t \end{bmatrix}.
\end{aligned} \tag{13}
$$

More compactly, with obvious redefinition of terms:

$$
\begin{bmatrix} k_t \\ r_t \end{bmatrix} = C + A \begin{bmatrix} k_{t-1} \\ r_{t-1} \end{bmatrix} + B \begin{bmatrix} k_{t-2} \\ r_{t-2} \end{bmatrix} + V_t,
$$

where: $V_t \sim NID(0,\Omega)$ with:

$$
\begin{aligned}
\Omega_{11} &= \tau^2 \left(\frac{\lambda}{\delta} \frac{(\eta - \rho)}{(1 - \beta\lambda\eta)} \right)^2 + \sigma^2 \left(\frac{\lambda}{\delta} \frac{1}{(1 - \beta\lambda\rho)} \right)^2, \\
\Omega_{12} &= \Omega_{21} = \sigma^2 \left(-\frac{\lambda}{\delta} \frac{1}{(1 - \beta\lambda\rho)} \right), \\
\Omega_{22} &= \sigma^2.
\end{aligned}
$$

Note that A_{11} and B_{11} must satisfy $\eta^2 - \eta A_{11} - B_{11} = 0$, so that, given the error structure, the following restriction – which in the words of *Sargent* is the hallmark of rational expectations – can be imposed: $\Omega_{22} A_{12} + \Omega_{21} (\eta - A_{22}) = 0$, where η is a solution to the quadratic equation above. This restriction should be taken into account at the estimation stage: this in turn requires adopting a maximum likelihood approach.[10] Note that in this example β, γ, δ and σ^2 are not separately identified, and thus the value of one of these parameters must be specified *a priori*. As mentioned above, in the following we will assume the parameter δ to be the same across all agents and known exactly *a priori*.

The structural parameters can be recovered as follows: δ is specified *a priori*; $\mu = C_2$, $\sigma^2 = \Omega_{22}$, $\rho = A_{22}$,

10. The approach outlined closely follows the original proposal by *Lucas* and *Sargent* and shows quite clearly the need to disentangle deep from policy parameters in estimating decision rules. An alternative – and indeed more frequently adopted – approach would be to estimate directly the *Euler* condition by GMM, to avoid the structural instability that would affect the decision rule (12) were the policy rule (10) to be modified in the sample period. In this respect note that, as we control the data generating process, we can avoid in-sample policy breaks. It is also important to realise that the problems of non-structurality that we shall be concerned with would appear essentially unchanged under the alternative procedure. In particular, recovering the aggregate decision rule from the aggregation of the individual *Euler* equations is not equivalent to the straightforward aggregation of the individual decision rules, as a non linear transformation is involved. In particular, the aggregation of individual first order conditions would involve weighting individual deep parameters with (time-varying) terms of the form k^i/k, that depend on the policy parameter (see (12)).

$$
\begin{aligned}
\eta &= A_{12} - \frac{\Omega_{22}}{\Omega_{21}} A_{22}, \\
\lambda &= A_{11} - \eta, \\
\beta &= \left[1 - \frac{\lambda}{\delta} \frac{(\eta - \rho)}{A_{12}}\right] \frac{1}{\lambda \rho}, \\
\gamma &= \frac{\beta \lambda \mu}{(1 - \beta \lambda \rho)} + \frac{\delta}{\lambda} \frac{(1 - \beta \lambda)}{(1 - \eta)} (C_1 + \frac{\lambda}{\delta} \frac{\mu}{(1 - \beta \lambda \rho)}), \\
\tau^2 &= \frac{\left[\Omega_{11} - \sigma^2 \left(\frac{\lambda}{\delta (1 - \beta \lambda \rho)}\right)^2\right]}{\left(\frac{\lambda}{\delta (1 - \beta \lambda \eta)}\right)^2}.
\end{aligned}
$$

With those values on hands, the coefficients of the VAR in eq. (13) can be modified to reflect a policymaker's intervention on (say) ρ. According to the $RARE$ approach the correct prediction of the effects of a change in r_t due to the said intervention would thus obtain.

In deriving the VAR (13) it was assumed that all agents are alike. If, instead, the world is populated by heterogeneous agents, and the VAR is estimated with aggregate (macro) data (as will be done in the next section), then the micro-founded VAR model (13) is dynamically misspecified. This implies that the residual V_t is not distributed as a white-noise vector, as suggested by the findings in Section 3. To the practical consequences of that misspecification we now turn.

5 The Numerical Results

We consider ten different firms with heterogeneous parameters. In Table 1 the assumed parameters are listed, together with their means and standard deviations. Assuming first a policy regime characterised by ρ and μ equal to 0.55 and 1 respectively, we let each of the firms compute its optimal investment decision in a stochastic environment, with shocks to both the idiosyncratic productivity and the interest rate processes,[11] according to (the idiosyncratic version of) equation (13). The aggregate stock of capital, k_t, is then computed as a simple average of the individual capital stocks.[12] With these data for the aggregate capital stock, and with the data for the rental rate, a constrained maximum likelihood estimate of the system (13) is performed, and the RA parameters are recovered from the coefficients of the aggregate decision rule. The procedure is repeated in a new policy regime, characterised by ρ equal to 0.605 (a 10 per cent shock). The new estimation only uses data from the *post*-shock regime

11. The variance of the both idiosyncratic and policy innovations were set equal to 0.01.
12. While different aggregation criteria could be considered that avoid some of the problems we shall highlight, aggregation by sum or average characterizes national account time series.

TABLE 1

Assumptions on the Parameters

	η	β	γ	λ	δ
Firm 1	0.610	0.950	5.5	0.724	0.500
Firm 2	0.650	0.900	8.0	0.693	0.500
Firm 3	0.690	0.850	10.5	0.660	0.500
Firm 4	0.730	0.800	13.0	0.626	0.500
Firm 5	0.770	0.750	15.5	0.591	0.500
Firm 6	0.810	0.700	18.0	0.556	0.500
Firm 7	0.850	0.650	20.5	0.519	0.500
Firm 8	0.890	0.600	23.0	0.481	0.500
Firm 9	0.930	0.550	25.5	0.443	0.500
Firm 10	0.970	0.500	28.0	0.405	0.500
Average	0.790	0.725	16.750	0.570	0.500
Standard Dev.	0.121	0.151	7.56	0.108	0.00

TABLE 2

Aggregate Reduced form Coefficients before and after the Policy Shock
(ρ raised from 0.55 to 0.605)

	A_{11}	A_{12}	B_{11}	C_1
Before the shock	1.381	0.352	−0.469	4.537
After the shock	1.383	0.275	−0.470	4.464
After the shock, using the estimates of deep parameters obtained before the shock	1.381	0.275	−0.469	4.487

TABLE 3

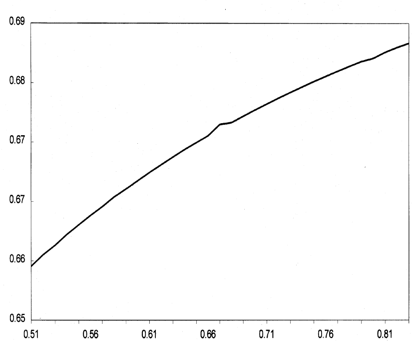
Aggregate Deep Parameters before and after the Policy Shock
(ρ raised from 0.55 to 0.605)

	β	λ	η	γ
Before the shock	0.662	0.599	0.781	14.477
After the shock	0.669	0.603	0.779	14.126

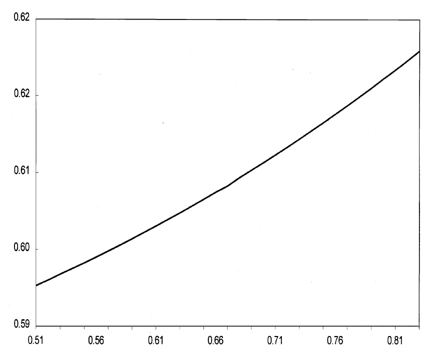
In both cases we use a very large sample (90,000 observations) to virtually eliminate small sample bias that can induce spurious dependence of the ‘deep’ parameters from the policy one. Given the large sample size the estimates are a close approximation to the probability limit of the parameters of the (possibly misspecified) estimated model under the true DGP – that is, they can be interpreted as pseudo-true values of the corresponding parameters. The standard deviation of the parameters implied by the maximum likelihood optimisation are indeed negligible and the variability of the parameters observed in this large sample exercise is indeed significant.

The first rows of Tables 2 and 3 present, respectively, the estimated coefficients and the implied ‘deep’ parameters. The second rows of the two Tables present the corresponding values obtained in the *post*-shock regime. Given the nature of pseudo-true values of these estimates, the logic of the $RARE$ approach would be called into question were the parameters to change, as they would no longer have any claim to being structural (or deep). Equivalently, troubles for the $RARE$ approach would be signalled by sizeable differences between these ‘actual’ *post*-shock decision rule coefficients and the coefficients –shown in the third row of Table 2– recomputed by taking into account the policy change and assuming all RA parameters to be

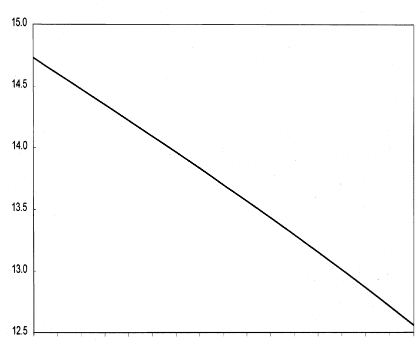
FIGURE 1
Parameters Estimated as Function of the Policy (Large sample)



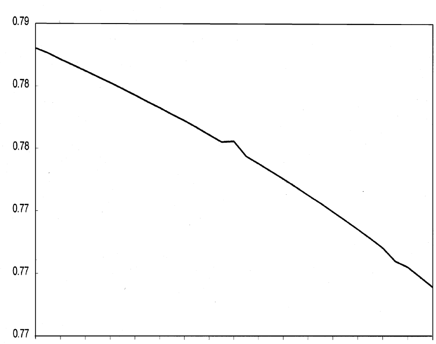
Deep parameter β



Deep parameter λ



Deep parameter γ



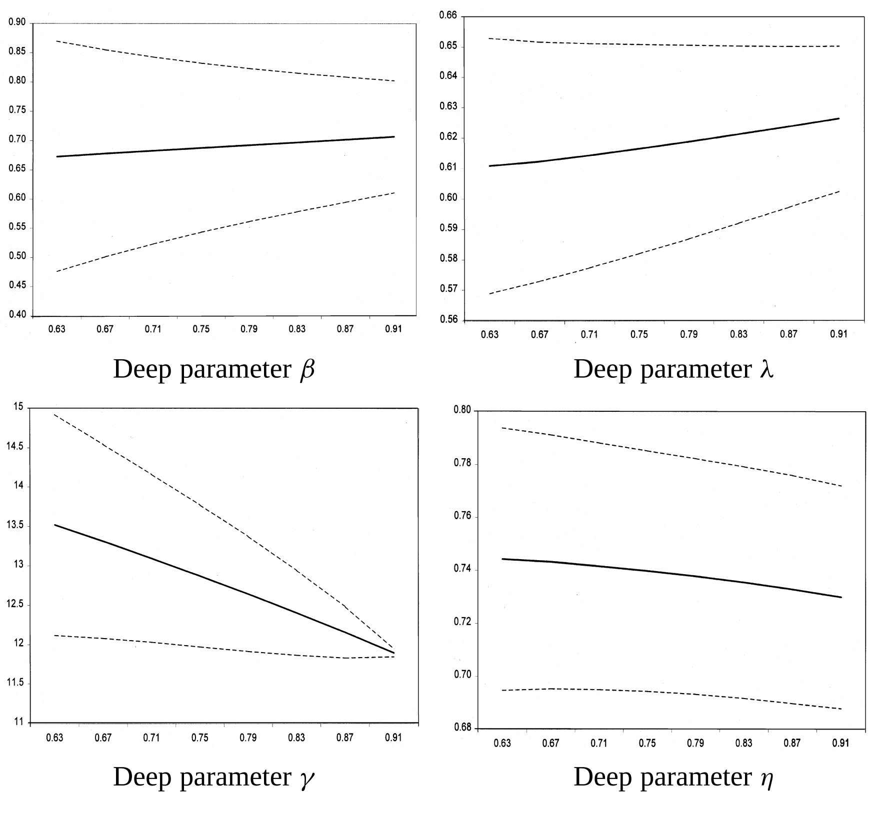
Deep parameter η

unchanged. As it is apparent from the tables, neither the ‘inferred’ decision rule coefficients (third row of Table 2) nor the *ante*-shock RA parameters (first row of Table 3) are equal to, respectively, the actual decision rule coefficients (second row of Table 2) and the *post*-shock RA parameters (second row of Table 3).

The non-structurality of the RA parameters is graphically exemplified in Figure 1, where the functions relating the pseudo-true value of the deep parameters $\{\beta,\lambda,\eta,\gamma\}$ to ρ are plotted, with each corresponding parameter being the estimate computed using data generated consistently with the corresponding value[13] of ρ. The graph highlights the non-linear dependency of the supposedly deep parameters on the policy parameter.

In order to check the estimation results and the appropriateness of the sample size, we performed the same experiment considering a single agent economy with parameter values equal to the averages used in the heterogenous case and with the same sample size. In this case the variation of the deep parameters is insignificant (of the order of one percent of the one found in the heterogenous case).

FIGURE 2
Parameters Estimated as Function of the Policy (Small sample)



Deep parameter β — Deep parameter λ

Deep parameter γ — Deep parameter η

13. To draw the graphs, we simulated and estimated the model for forty different values of the policy parameter ρ on the interval $(0.51, 0.83)$ using the same underlying set of innovations.

To conclude our investigation, we ask if the variability found with the large sample is still detectable when only sample sizes typical of macroeconometric applications are available (250 data points). We performed a Monte carlo exercise generating with the heterogeneous agents model 1,000 replications of a sample of 250 observations; for each sample, restricted maximum likelihood was performed. Figure 2 shows the average estimate, for the whole set of 1,000 replications, of the deep parameters $\{\beta, \lambda, \eta, \gamma\}$ as a function of ρ (solid line), plus or minus one standard deviation (dotted lines). The results suggest that, at least in the case considered, the econometrician would not be able to detect the effects of heterogeneity on the estimated deep parameters. This does not imply that the latter are negligible, but only that they might be difficult to identify when limited information is available.

Finally, it is also worth recalling that the estimates of the aggregate decision rule coefficients frequently implied extreme values for some of the 'deep' parameters of the RA – outside the convex combination of the individual deep parameters or even outside the range of admissible values (in these circumstances, in spite of the linearity of the individual decision rules, there would be no representative agent). The presence of non-linear cross equation restrictions implies that the deep parameters are recovered through non-linear transformations of the reduced form coefficient estimates. The non-linear transformation can easily induce large changes in the variance of the estimated deep parameters. It is therefore possible that even if the econometrician is able to recover quite precise estimates of the reduced form parameters, a non-linear transformation of the latter can present a very large variance, and it is not infrequent to recover bizarre 'deep' parameters.

6 Conclusions

The device of interpreting macroeconomic phenomena as corresponding to the optimising behaviour of a (large) RA is –as a cursory look at the most recent macro textbooks will confirm– a corner-stone of modern, micro-founded macroeconomic theory. From an empirical point of view, the RA device finds its support in the possibility of recovering (aggregate) deep parameters from aggregate data.[14]

Our results, which builds on arguments outlined in GEWEKE [1985] and KIRMAN [1992], clearly imply that neglecting the possible heterogeneity of the agents and imposing a representative agent model on aggregate macroeconomic data can result in a similar but conceptually different source of non-struturality of the policy relevant parameters. This form of non-struc-

14. Critics of this program are usually not taken very seriously. To quote just one example, while KIRMAN [1992] original wording was hardly ambiguous "*it is clear that the representative agent deserves a decent burial, as an approach to economic analysis that is not only primitive, but fundamentally erroneous*", TURNOVSKY [1995] refers to it by writing: "*Kirman seems to suggest that we should [abandon the representative agent model]*", and hastens to add: "although that view seems extreme".

turality is related to the underlying neglected heterogeneity and on the differential effect of policy changes on the cross-section of agents.

This leaves economists in a very uncomfortable position. If we cannot interpret aggregate data with the conceptual framework provided by microeconomic theory, should we limit ourselves to take note of broad aggregate correlations and hope that they will persist?

As is often the case in economics, a ‘corner’ solution is not the optimal one. Neither ascribing ‘*legitimacy only to models that are exact aggregation of agents who optimize subject to constraints*’ (HAHN and SOLOW [1995]), nor resorting to pure time series analysis is likely to be the appropriate way out. Rather, to quote again HAHN and SOLOW, we should “*pay attention to microfoundations in the sense that [our macro models] are suggested by or analogous to or loosely abstracted from the micro models*”. As a consequence “*econometricians who are working in view of studying macroeconomic policy must be satisfied with impure procedures: they are evidently not justified in deforming the reality of complex phenomena in order to force it to fit into overly simplified specifications*” (MALINVAUD [1981]). ▼

• References

ALTISSIMO F., SIVIERO S., TERLIZZESE D. (1999). – « How Deep are the Deep Parameters? », Bank of Italy *Discussion Paper*, n°354.

FARMER R.E.A. (1991). – « The *Lucas Critique*, Policy Invariance and Multiple Equilibria », *Review of Economic Studies*, p. 321-32.

FAVERO C., HENDRY D.F. (1992). – « Testing the *Lucas critique*: A Review », *Econometric Reviews*, p. 265-66.

FORNI M., LIPPI M. (1997). – *Aggregation and the Microfoundations of Dynamic Macroeconomics*, Oxford, Clarendon Press.

GEWEKE J. (1985). – « Macroeconomic Modelling and the Theory of the Representative Agent », *American Economic Review Papers and Proceedings*, p. 206-10.

GORMAN W. M. (1953). – « Community Preference Fields », *Econometrica*, 21, p. 63-80.

GOURIEROUX C., MONFORT A. (1996). – *Simulation Based Methods in Econometrics*, Oxford, Oxford University Press.

GRANGER C. (1980). – « Long Memory Relationships and the Aggregation of Dynamic Models », *Journal of Econometrics*, 14, p. 227-238.

HAHN F.H., SOLOW R. (1995). – *A Critical Essay on Modern Macroeconomic Theory*, Oxford, Blackwell.

HANSEN L.P, SARGENT T. (1980). – « Formulating and Estimating Dynamic Linear Rational Expectation Models », *Journal of Economic Dynamics and Control*, p. 7-46.

HANSEN L.P., SARGENT T. (1981). « Linear Rational Expectation Models for Dynamically Interrelated Variables », in R. E. Lucas and T. Sargent (eds.), *Rational Expectation and Econometrics Practice*, London, Allen and Unwin.

HENDRY D.F. (1988). – « Testing Feedback vs. Feedforward Econometric Formulations », *Oxford Economic Papers*, p. 132-49.

KIRMAN A. (1992). – « Whom or What Does the Representative Individual Represent », *Journal of Economic Perspectives*, Vol. 6, p. 117-36.

INGRAM B.F. (1995). – « Recent Advances in Solving and Estimating Dynamic Macroeconomic Models », in K.D. Hoover (ed.), *Macroeconometrics*, Boston, Kluwer.

LIPPI M., ZAFFARONI P. (1999). – « Aggregation of Simple Linear Dynamics: Exact Asymptotic results », STICERD-LSE Econometrics *Discussion Paper* 350.

LUCAS R.E. (1976). – « Econometric Policy Evaluation: A Critique », in K. Brunner and A. Meltzer (eds.), *The Phillips Curve and Labor Market*, Amsterdam, North-Holland.

MALINVAUD E. (1981). – « Econometrics Faced with the Need of Macroeconomic Policy », *Econometrica*, p. 1363-75.

SARGENT T. (1981). – « Interpreting Economic Time Series », *Journal of Political Economy*, p. 213-48.

SIMS C. (1982). – « Policy Analysis with Econometric Models», *Brookings Papers on Economic Activity*, p. 106-152.

TURNOVSKY S.J. (1995). – Method of Macroeconomic Dynamics, Cambridge MA, MIT Press.