On the Microfoundations of Dynamic Macroeconomics

by

Mario Forni*
Marco Lippi**

Febbraio 1998

* Università degli Studi di Modena
Dipartimento di Economia Politica
Viale Berengario, 51
41100 Modena (Italy)
e-mail: fornì@unimo.it

** Università degli Studi di Roma
Dipartimento di Scienze Economiche
Via Cesalpino, 14
00161 Roma (Italy)
e-mail: lippi@giannutri.caspur.it
Abstract

We survey a number of important results concerning aggregation of dynamic, stochastic relations. We do not aim at a comprehensive review; instead, we focus heavily on the results collected in Forni and Lippi (1997). We argue that the representative-agent assumption is misleading and the microfoundation of dynamic macroeconomics should be based on explicit modeling of heterogeneity across agents. An unpleasant aspect of this modeling strategy is that macroeconomic implications of micro theory are difficult to obtain. However, difficulties are reduced by large number results. Moreover, puzzling implications of existing theories could be reconciled with empirical evidence on macro data.

Keywords: Aggregation, heterogeneity, representative agent, linear stochastic process, dynamic factor model.

1. Introduction

This paper reviews some important results on aggregation of linear stochastic dynamic models employed in macroeconomics. A comprehensive discussion is beyond our purposes, so that some interesting contributions will not be mentioned. Rather, we report mainly results by ourselves, illustrated in detail in Forni and Lippi (1997). Rigour is not our primary concern here. Exposition is based on simple examples, with a hint to generalizations, so that a non-specialist reader may get an idea of the problems, the difficulties and extant results. Nonetheless, we assume that the reader is acquainted with elementary theory of ARIMA stochastic processes and the notation based on the lag operator $L$.

The framework in which our aggregation problem arises is Standard Macroeconomics of aggregate consumption, income, investment, employment. This field has been dominated in the last two decades by two, not necessarily conflicting, approaches. The first is New Classical Macroeconomics, which is characterized by the strong prescription that models linking observable variables must be derived from microeconomic first principles, and in particular that the dynamics of such models must be a consequence of intertemporal optimization, rather than ad hoc superposition to a static maximization scheme. The second is based on VAR and Structural VAR models, in which estimation of a relatively theory-free statistical model comes first, while theory enters at the identification stage.\footnote{For a vast treatment of models based on dynamic objective functions and rational expectations see Hansen and Sargent (1991). A simple presentation of the main topics can be found in Sargent (1987). On Structural VAR models, see Bernanke (1996), Shaikh and Watson (1988), Evans (1988), Giannini (1992).}

We shall focus mainly on the former approach, but the latter is also discussed briefly in Section 7. We will claim that irrespective of which approach is taken up, interpretation of dynamic macroequations encounters a very serious aggregation problem: when heterogeneity of agents is allowed, an important change of dynamic shape is likely to occur between micro and macro equations. Basic properties of the micro model do not hold in general for the macromodel: for instance, micro cointegration does not imply macro cointegration, lack of Granger causality in the micro model does not imply the same property on the macro level, static microequations may be transformed by aggregation into dynamic macroequations.

As a consequence, overidentifying restrictions produced by the theory cannot be tested directly using aggregate data. This is unpleasant, since the difficulties involved in formulating a macro model are increased. Aggregate implications of micro theory can only be found by explicit modeling of heterogeneity, which is likely to require a lot of additional information with respect to the traditional strategy. On the other hand, there is also a pleasant implication: existing models which are at odds with aggregate data under the representative-agent assumption could be reconciled with empirical evidence.

The paper is organized as follows. In Section 2 we present a most simple example of the micro-macro effect that we want to illustrate in this paper. We have a microequation linking an independent to a dependent variable. The microequation is static, while the independent variable is autocorrelated; if heterogeneity in the microparameters is allowed, the microequation linking the aggregate dependent variable to the aggregate independent variable is dynamic. The example in Section 2 is highly artificial. In Section 3 a more realistic model for the independent variables is discussed. The model is based on the distinction between common and idiosyncratic components. When the number of individuals is huge, common components survive aggregation whereas idiosyncratic components are washed away. Recent empirical work is reported, in which it is shown that major macroeconomic variables are driven by several common components. In Section 4 we give a definition of micro and macroeconomic models that is general enough to accommodate most of the models employed in the literature. We show that the microparameters are analytic functions of the deep microparameters. As a consequence, restrictions on the microparameters either hold on the whole microparameter space or hold only on a subset of zero Lebesgue measure. An application of this Alternative Principle is given in Section 5, where we show that cointegration of micromodels does not imply macro cointegration, apart from negligible subsets of the microparameters space. In Section 6 we review an important case in which aggregation has a positive effect. Permanent-income theory of consumption under rational expectations is at odds with aggregate empirical evidence when the representative agent is assumed. Two well-known inconsistencies are "excess sensitivity" and "excess smoothness". However, allowing for some heterogeneity and assuming limited information, aggregate data can be reconciled with the theory. In Section 7 we show some unpleasant consequences of aggregation on Granger non-causality and the interpretation of VAR models. Section 8 concludes.

2. An Example: A Dynamic Macroequation with a Static Microequation

Let us begin with a short review of the theoretical steps involved in the formulation of a typical "microfounded" partial-equilibrium macroeconomic model.

(i) A quadratic intertemporal optimization problem for an economic agent is set up and solved; the result is a linear equation linking the dependent variable \( Y_t \) to the expected future values of the independent variable \( x_t \) and possibly the lagged values of both variables;

(ii) A stochastic linear dynamic equation for \( x_t \) is specified. The example

\[ x_t = u_t + \alpha_1 u_{t-1}, \]

with \( u_t \) white noise, will be useful to fix ideas. Assuming rational expectations, equation (1) is used to replace the expected future values of \( x_t \) with present and past values, leading to an equation like

\[ y_t = \alpha_2 x_{t-1} + \beta x_t + \epsilon_t, \]

where only one-period lags have been included for simplicity and \( \epsilon_t \) is a white-noise residual.

(iii) The agent following the micro model (1)-(2) is assumed to be "representative", meaning that model (1)-(2) can be interpreted as a macro model and therefore can be employed directly for estimation and testing with macro data.

In this paper we depart thoroughly from this microfoundation paradigm. The crucial difference is that we drop the representative agent assumption in step (iii) and assume instead that agents are heterogeneous. We do not place special emphasis on the optimization and the rational expectations steps. Rather, we start by introducing heterogeneity into the micro equations (1) and (2):

\[ y_t = a_1 x_{t-1} + \beta x_t + \alpha_2 x_{t-1} + \epsilon_t, \]
\[ x_t = \beta u_t + \alpha_1 u_{t-1}. \]

Then we focus on the following questions: What is the macroequation linking the aggregate variables, i.e. \( Y_t = \sum_i y_{it} \) and \( X_t = \sum_i x_{it} \)? Can we suppose that it is obtained by simply averaging over the coefficients of (3)? Under which conditions the dynamic properties of the micro models hold true in the macro model?

In general the features of the micro models are not preserved at the macro level. The following example should be sufficient to give the reader an idea of the complications arising. Let us assume for simplicity \( b_i = c_i = 0 \) and \( e_i = 0 \) for all \( i \), so that we are left with the static, exact micro equations

\[ y_i = a_2 x_{it}. \]

Let us assume also that the independent variables of different agents are orthogonal at any lead and lag, i.e. \( \text{cov}(x_{it}, x_{i't}) = 0 \) for \( i \neq i' \) and any integer \( k \), and \( \text{var}(x_{it}) = 1 \) for any \( i \). Finally, let us assume that there are only two agents (or two groups of identical agents), i.e. \( Y_t = \sum_i y_{it} \) and \( X_t = \sum_i x_{it} \). Now consider the static regression of \( Y_t \) on \( X_t \):

\[ Y_t = AX_t + \Omega_t. \]

A simple calculation shows that

\[ A = a_1 (1 + a_1^2) + a_2 (1 + a_2^2) \]
\[ \text{var}(\Omega_t) = (a_1 - A)^2 (1 + a_1^2) + (a_2 - A)^2 (1 + a_2^2) \]
\[ \text{cov}(\Omega_t, \Omega_{t-1}) = (a_1 - A)^2 \sigma_1 + (a_2 - A)^2 \sigma_2 \]

Hence \( \text{var}(\Omega_t) = 0 \) if and only if \( a_1 = a_2 \), while \( \text{cov}(\Omega_t, \Omega_{t-1}) = 0 \) if and only if either \( a_1 = \epsilon_2 \) or \( a_2 = \epsilon_2 \). Thus if the behavioral coefficients are heterogeneous, and there is some autocorrelation in individual incomes, a researcher estimating (5) will find a non-zero residual (whereas no residual is present in the first equation of the micromodel (4)) and detect autocorrelation in \( \Omega_t \). If our researcher shares the common representative-agent attitude, he will conclude that the true relationship between the micro counterparts of \( Y_t \) and \( X_t \) is a dynamic equation like (2) and therefore model (4) must be rejected.

We can stop here the analysis of this very simple example. The message is that aggregation of dynamic multivariate models can produce non-trivial changes in
the dynamic shape of each single equation. Starting with a static microequation, like the first of (4), we may end up with a dynamic equation provided that some heterogeneity among the agents is allowed.²

3. Independent Variables: Common and Idiosyncratic Components

The examples in Section 2 are partial equilibrium models, in which one or several variables are taken as given both from the agents and the researcher. For general equilibrium models we refer to Forni and Lippi (1997, Chapter 7). Here we shall deal only with partial equilibrium models, which are sufficient for our illustrative purposes.

In the latter models, agents may differ for two reasons: because their independent variables are different, and because their responses to their independent variables are different. Let us begin by modeling the differences in independent variables. In the example of the previous section we have assumed that the variables $z_{it}$ corresponding to different individuals, are orthogonal at all leads and lags. Although convenient to simplify calculations, this assumption is far from being realistic. Incomes of different agents, wages faced by different firms, although different, are correlated both simultaneously and with lags. A natural way in which both differences and correlation across individual variables can be represented is the dynamic factor model (Sargent and Sims, 1977; Geweke, 1977).

To fix ideas suppose that the independent variable is income and that $y_{it}$ is income of agent $i$. A very simple dynamic factor model for $y_{it}$ is

$$y_{it} = b_i U_t + \xi_{it},$$

where we assume that:

(a) $U_t$ and $\xi_{it}$ are stationary stochastic variables;
(b) $\xi_{it}$ is orthogonal to $U_{it-k}$ for any integer $k$;
(c) $\xi_{it}$ is orthogonal to $\xi_{it-j}$ for any $j \neq i$ and any integer $k$. Indeed, this model is so simple that it could be confused with a static factor model; notice however that, although the response of the variables to $U_t$ is static, the orthogonality conditions are dynamic. The variable $U_t$, the common factor or common shock, is a source of variation affecting all micro incomes, even though with different impacts as measured by the coefficient $b_i$. The term $b_i U_t$ will be called the common component of $y_{it}$. Changes in $U_t$ can represent for instance fiscal or monetary policy changes, as well as changes in productivity or labor supply that affect the whole economy. By contrast, each of the variables $\xi_{it}$, the idiosyncratic component, represents events affecting only one individual, like health or luck, and are therefore uncorrelated to one another at any lead and lag.

Model (6) can be generalized by introducing more than one common shock and add lags in the response of $y_{it}$ to the common shocks, thus obtaining

$$y_{it} = b_{it1}(L)U_{it1} + b_{it2}(L)U_{it2} + \cdots + b_{itk}(L)U_{itk} + \xi_{it},$$

where $\xi_{it}$ is orthogonal to $U_{it-k}$ for any integer $k$ and any $s = 1, k$. Notice that we are not assuming that the variables $U_{it}$, or the variables $\xi_{it}$, are white noises. However, if the $U_{it}$'s are assumed to be stationary, model (7) can be rewritten as

$$y_{it} = a_{it1}(L)U_{it1} + a_{it2}(L)U_{it2} + \cdots + a_{itk}(L)U_{itk} + \xi_{it},$$

where $(u_{t1}, u_{t2}, \ldots, u_{tk})$ is an orthogonal white-noise vector, i.e. $\text{var}(u_{it}) = 1$ for any $i$, $\text{cov}(u_{it}, u_{jt-k}) = 0$ for $i \neq j$ and any integer $k$.

Moreover, since empirical incomes are non-stationary while income changes are stationary, then either we interpret the variables $y_{it}$ in (8) as deviations from a deterministic trend, or we modify (8) in such a way that the stationary RHS drives the changes of income:

$$(1-L)y_{it} = a_{it1}(L)U_{it1} + a_{it2}(L)U_{it2} + \cdots + a_{itk}(L)U_{itk} + \xi_{it}.$$³

Factor models like (8) or (9) have been recently employed in macroeconomic literature as parsimonious representations when the number of variables is large with respect to the number of available observations over time.³ With a small number of factors, models (8) or (9) can provide a considerable reduction of the number of parameters to be estimated as compared to an unrestricted VAR model, in which each of the variables is regressed on itself and lagged values of the others. Despite parsimony, model (9) is flexible enough to allow for a substantial amount of heterogeneity.

Models (8) and (9) have a very important property when the number of individuals is huge. Let us fix ideas on (8). Formally, it is convenient to assume that there exists a countable infinity of agents. We assume also that $\text{var}(\xi_{it})$ is bounded, i.e. there exists a real $\Lambda$ such that $\text{var}(\xi_{it}) \leq \Lambda$ for any $i$. Finally, for simplicity, we set $a_{it1}(L) = b_i(L) \neq 0$ for any $i$ and $s = 1, k$. It is easily seen that when $n$ tends to infinity the variance of the aggregated common component

$$n(b(L)\xi_{it1} + \cdots + b_k(L)\xi_{itk})$$

tends to infinity as $n^2$, whereas the variance of the aggregated idiosyncratic component

$$\sum_{i=1}^n \xi_{it}$$

cannot tend to infinity faster than $n$. Thus, if pre-capita income $y_{it} = Y_{it}/n$ is considered, the idiosyncratic component disappears as $n$ gets larger.⁴ For large $n$ we have approximately

$$y_{it} = b_1(L)y_{it1} + b_2(L)y_{it2} + \cdots + b_k(L)y_{itk},$$

or, more in general,

$$y_{it} = a_1(L)y_{it1} + a_2(L)y_{it2} + \cdots + a_k(L)y_{itk},$$

³ The large-numbers effect is well-known for static models, particularly in the finance literature (see e.g. Chamberlain and Rothschild, 1983). Some important implications for aggregation and macroeconomics are discussed in Granger (1987, 1990). Forni and Lippi (1997, Chapter 1) provide conditions under which the result can be extended to the dynamic case.

⁴ See for instance Qian and Sargent (1993) and Forni and Reichlin (1995).
where \( \alpha_i(L) \) is the cross-sectional average of the \( \alpha_{it}(L) \).

The elimination of the idiosyncratic components has crucial consequences for both theory and empirical work. From a theoretical point of view, it provides a nice way to reconcile macroeconomics with micro heterogeneity. Individual variables corresponding to different agents may be almost orthogonal to one another owing to big idiosyncratic components as compared to the common components. Therefore, individual variables can be viewed as spanning a vector space with a huge number of dimensions. Nonetheless, this is perfectly consistent with a very basic idea of macroeconomic theory—that aggregate variables, rather than depending on all the corresponding microvariables, can be represented as driven by a relatively small number of macroeconomic sources of variation.

Regarding empirical work, the large-numbers effect can be exploited in order to estimate the model and to study the problem of how many independent common components drive the micro- and the macrovariables. On estimation we refer to Forni and Reichlin (1995) and Forni, Hallin, Lippi and Reichlin (1998).5 Regarding inference on the number of common components, some work can be found in Forni and Lippi (1997) and in both of the papers quoted above.

In Forni and Lippi (1997, Chapter 2) data on incomes and wages relative to US states are employed to show that the number of common components in a representation of the form (9) for individual incomes is definitely bigger or equal to two. This evidence is at odds with some recent applied macroeconomic work on consumption, in which heterogeneity of incomes is introduced but the micro model contains only one common component (see Section 6). This outcome, i.e. more than one common component in individual major economic series is interesting both per se, and because, as we will see, many of the effects of aggregation on dynamic micro models depend on (1) heterogeneity of different agents’ behaviors, (2) at least two common components driving the independent variables.

Now let us come back to the mutual orthogonality assumption for the idiosyncratic components. The following two examples show that this condition rules out interesting economic models and should be relaxed in some way.

As a first example, consider an n-industry constant-returns economy, where production of industry \( i \) at time \( t \) is given by the equation

\[
y_{it} = a_{i1}y_{it-1} + a_{i2}y_{it-2} + \cdots + a_{in}y_{it-n} + u_{it} + \chi_{it},
\]

where \( u_{it} \) is a demand common shock, \( \chi_{it} \) is an idiosyncratic shock fulfilling the orthogonality assumption, i.e. \( \text{cov}(X_{it}, \chi_{it-k}) = 0 \) for \( i \neq j \) and any integer \( k \), \( \alpha_{it} \) is the quantity of the \( i \)-th product necessary as a means of production to produce one unit of the \( j \)-th product, the one-period lag on the RHS meanings that industry \( i \) is purely producing to replace means of productions employed by other industries in the previous period. Starting with (10) and inverting the input-output matrix we obtain an equation like (8):

\[
y_{it} = \alpha_i(L)u_{it} + \xi_{it},
\]

where

\[
\begin{pmatrix}
\xi_{1t} \\
\xi_{2t} \\
\vdots \\
\xi_{nt}
\end{pmatrix}
= (I + DL + D^2 L^2 + \cdots) \begin{pmatrix}
X_{1t} \\
X_{2t} \\
\vdots \\
X_{nt}
\end{pmatrix}
\]

and

\[
D \text{ being the matrix having } D_{ij} \text { in place } i, j. \text { However, unlike in (8), unless } D \text { is diagonal, the assumption } \xi_{it} \perp \xi_{it-k} \text { for } j \neq i \text { and any integer } k \text { is no longer valid.}
\]

The problem stems from the fact that shocks originated in sector \( i \), while affecting directly only \( y_{it} \), propagate through the system via the autoregressive linkages.

As a second example, consider the model

\[
y_{it} = a_i(L)u_{it} + b_i(L)v_{it} + c_i(L)x_{it},
\]

where

\[
\begin{pmatrix}
\xi_{1t} \\
\xi_{2t} \\
\vdots \\
\xi_{nt}
\end{pmatrix}
= (I + DL + D^2 L^2 + \cdots) \begin{pmatrix}
X_{1t} \\
X_{2t} \\
\vdots \\
X_{nt}
\end{pmatrix}
\]

Here the income of region \( i \) in nation \( j \) is driven by a world-wide shock \( u_{it} \), a national shock \( v_{it} \) and a local shock \( x_{it} \). A model like this is employed in Forni and Reichlin (1997) in order to study intercurrences between European regions and to compare them with US counties. The national shocks could be accomodated in model (8) as common shocks; however, if the number of nations in the model is large, parsimony would be lost. By contrast, absorbing the national component into the idiosyncratic term would violate orthogonality, since the variance-covariance matrix would be block-diagonal rather than diagonal.

Now, if the orthogonality assumption on the idiosyncratic terms is dropped in (8) or (9), the theoretical distinction between idiosyncratic and common components is lost and boundedness of \( \text{var}(\xi_{it}) \) is no longer sufficient to ensure that the per-capita idiosyncratic variance tends to zero. However, mutual orthogonality and bounded variance can be substituted by the following more general condition.

Let \( \Sigma_{\xi} \) be the variance-covariance matrix of the vector \( (\xi_{1t}, \xi_{2t}, \ldots, \xi_{nt}) \), and let \( \lambda^* \) be its maximum eigenvalue. If \( \lambda^* \) is bounded then the variance of

\[
\frac{1}{n} \sum_{t=1}^{n} \xi_{it}
\]

tends to zero as \( n \) tends to infinity. This is very easy to show. Indicating by \( w \) the \( n \)-dimensional vector with all its components equal to unity, we have

\[
\text{var} \left( \frac{1}{n} \sum_{t=1}^{n} \xi_{it} \right) = \frac{1}{n^2} \text{tr}^2 \Sigma_{\xi} \leq \frac{1}{n^2} \| w \|^2 \lambda^* = \frac{1}{n} \lambda^*.
\]

The bounded eigenvalue condition has been introduced in place of the traditional orthogonality assumption by Chamberlain (1983) and Chamberlain and Rothschild (1983), in a static context. The resulting model is named “approximate factor model”. A somewhat different assumption, based on dynamic eigenvalues (Brillinger, 1981), is introduced in Forni and Lippi (1988), where the “approximate dynamic factor model” is proposed and the representation theorems in Chamberlain and Rothschild (1983) are generalized to the dynamic framework.

5 Forni and Reichlin (1996) propose to estimate the factors and the common components by using simple averages, Forni and Reichlin (1997) and Forni, Hallin, Lippi and Reichlin (1998) propose procedures based on principal components. An estimator which is a linear combination of the observable variables is developed by Stock and Watson (1997) for a static model with time-varying coefficients.
4. Micro and Macromodels

In the previous section we have established and commented upon a model for the independent variables. Here we give a general representation of micromodels linking dependent to independent microvariables. Let us begin by an example. Assume that agent $i$ determines the variable $y_{it}$ in the following way:

$$y_{it} = d_i E(x_{it}|I_{it-1}),$$

where $x_{it}$ is an independent variable, $d_i$ is a deep parameter, e.g., a parameter of the utility or of the production function of agent $i$, $E(|I_{it-1})$ being the expectation conditional on the information set available at time $t-1$. Lastly, assume that

$$x_{it} = u_t + \alpha u_{t-1} + \xi_t$$

where $u_t$ is a common white noise with unit variance, while $\xi_t$ is an order-one idiosyncratic moving average:

$$\xi_t = \eta_t + \beta_1 \eta_{t-1}.$$  

For simplicity we assume $\eta_t$ and $\xi_t$ orthogonal at all leads and lags. The model of agent $i$ contains four parameters: $d_i$, $\alpha$, $\beta_1$, and $\sigma_{\xi_i}^2$, determining both the independent and the dependent variable. Notice that the parameter $\alpha$ is not agent specific, i.e., is equal for all the agents. To solve for the conditional expectation appearing in (11) we have to make some assumption about the set $I_{it}$. Two simple alternatives are:

(A) Agent $i$ observes and employs separately both the common and the idiosyncratic component of $x_{it}$. If this is the case

$$y_{it} = d_i \alpha u_{t-1} + d_i \beta_1 \eta_{t-1},$$

$$x_{it} = u_t + \alpha u_{t-1} + \eta_t + \beta_1 \eta_{t-1},$$

(B) Agent $i$ observes and employs only the variable $x_{it}$, with no distinction between the components of $y_{it}$. If this is the case, in order to determine $E(x_{it}|I_{it-1})$ we must resort to the univariate Wold representation of $x_{it}$

$$x_{it} = u_t + \alpha u_{t-1} + \eta_t + \beta_1 \eta_{t-1} + \delta_t \epsilon_{t-1},$$

where $\delta_t$ and $\sigma_{\epsilon_t}^2 = \text{var}(\epsilon_{t-1})$ are determined by the equations

$$\text{var}(x_{it}) = \sigma_{\epsilon_t}^2 (1 + \delta_t)^2 = 1 + \alpha^2 + \sigma_{\epsilon_t}^2 (1 + \beta_1^2)$$

$$\text{cov}(x_{it}, x_{it-1}) = \sigma_{\epsilon_t}^2 \delta_t = \alpha + \sigma_{\epsilon_t}^2 \beta_1.$$  

(14)

This system will provide two reciprocal values for $\delta_t$, the one smaller than unity in modulus being the solution. Thus we end up with the equations

$$y_{it} = d_i \delta_t \epsilon_{t-1},$$

$$x_{it} = \epsilon_{t-1} + \delta_t \epsilon_{t-1}.$$  

We can stop here with this example. What we want to retain is that in both cases we end up with a couple of linear dynamic equations linking $y_{it}$ and $x_{it}$ to the individual and common shocks $u_t$ and $\eta_t$, whose coefficients are functions of the common and individual parameters $d_i$, $\alpha$, $\beta_1$, $\sigma_{\epsilon_t}^2$. Under assumption (A) above such functions are elementary manipulations, whereas under assumption (B) they imply taking the roots of the algebraic system of equations (14).

Now, these features of our example are quite general. Even though the micromodels employed as microfoundations of dynamic macromodels can be much more complicated than the example, the general procedure already outlined at the beginning of Section 2 can now be described more precisely:

(a) One starts with intertemporal objective functions and dynamic equations for the independent variables, both depending on deep microparameters.

(b) Then one must solve algebraic equations resulting from the intertemporal optimization and possibly from the procedure necessary to obtain the Wold representation of the independent variables (as in the example, case (B)). The coefficients of such algebraic equations are simple functions of the deep microparameters (as in (14)).

(c) The final result is a system of linear dynamic equations, like (12) or (15), linking the variables of interest to the micro shocks, whose coefficients result from algebraic combinations of the deep parameters and the roots of the algebraic equations just mentioned in (b).

Thus, sticking for simplicity to the case of two variables $y_{it}$ and $x_{it}$, a micromodel can be defined as

(i) A set $\Gamma \subset \mathbb{R}^r \times \mathbb{R}$. This is the admissibility region for the microparameters, where $c$ is the number of common microparameters while $s$ is the number of individual microparameters. We can assume that $\Gamma$ is open and connected.

(ii) A function associating with any element of $\Gamma$ a system $S$ of linear dynamic equations linking the variables $y_{it}$ and $\eta_{it}$ to the common and idiosyncratic shocks. The coefficients of such systems $S$ are real functions defined on $\Gamma$, continuous on $\Gamma$, and analytic on $\Gamma$ with the exception of a subset of Lebesgue measure zero.

In the example above we have one common parameter, namely $\alpha$, so that $c = 1$, and three individual parameters, namely $d_i$, $\beta_1$, and $\sigma_{\epsilon_t}^2$, so that $s = 3$. A possible definition for $\Gamma$ is given by the constraints $1 > \alpha > -1$, $1 > \beta_1 > -1$, $\sigma_{\epsilon_t}^2 > 0$. Moreover, in case (A) the coefficients of (12) are elementary functions of the deep parameters, while in case (B), equations (15) contain both deep parameters and $\delta_t$, which is a function of deep microparameters through system (14). To see why in both cases the coefficients of system $S$ are analytic we must recall that the roots of algebraic equations are analytic as functions of their coefficients, with the exception of those values of the coefficients corresponding to multiple roots. It is reasonable to require that multiple roots may occur only for a negligible subset of $\Gamma$. Lastly, notice that, according to the above definition, in our example we have
two different micromodels, depending on whether we make assumption (A) or (B) on the information set $I_{n-1}$.

Once the micromodel has been defined we must define the macromodel. Let us firstly go back to our example. By aggregation over $n$ agents we have

$$Y_t = \alpha u_{t-1} + \sum_{i=1}^{n} d_i + \sum_{i=1}^{n} d_i \beta u_{t-1}$$

$$X_t = n (u_t + \alpha u_{t-1}) + \sum_{i=1}^{n} (\theta_i + \beta u_{t-1})$$

in case (A), while in case (B)

$$Y_t = \sum_{i=1}^{n} d_i \delta_i (u_t + \alpha u_{t-1}) + \sum_{i=1}^{n} d_i \delta_i (u_t + \beta u_{t-1})$$

$$X_t = n (u_t + \alpha u_{t-1}) + \sum_{i=1}^{n} (\theta_i + \beta u_{t-1})$$

The natural definition of the macromodel corresponding to $n$ agents and a given micromodel is:

(I) Assuming for simplicity that $\Gamma = \mathbb{R}^c \times \mathbb{R}^d$, the set

$$\Gamma_n = \mathbb{R}^c \times \mathbb{R}^d \times \cdots \times \mathbb{R}^c \times \mathbb{R}^d,$$
There are two important cases in which a c fulfilling (17) exists:
(a) If \( \delta = \delta_j \) for any \( i \) and \( j \). In this case \( c = \delta \).
(b) If there exists a \( \tau \) such that \( \beta_t = \sigma_i \), for any \( i \). In this case \( c = \sum \delta_i \alpha_n / \sum \alpha_n \).

The economic meaning of the above conditions is simple. If \( x_t \) is consumption and \( y_{it} \) is income, condition (a) means that all agents share the same long-run propensity to consume. On the other hand, when condition (b) holds, the shocks \( w_{it} \) and \( w_{it} \) are 'redundant', this meaning that all incomes are driven by the shock \( w_{it} + \gamma w_{it} \). This in turn implies that all individual incomes are pairwise cointegrated.

More generally, if a c fulfilling (17) exists then
\[
\sum_{i=1}^{n} \beta_i \delta_i = \sum_{i=1}^{n} \delta_i \alpha_n \beta_i,
\]
which has conditions (a) and (b) as particular cases. Since we have assumed five individual microparameters and no common microparameters, the set \( \Gamma_n \) is an open connected subset of \( \mathbb{R}^6 \). Since (18) describes an algebraic hypersurface in \( \mathbb{R}^6 \), the subset of \( \Gamma_n \) fulfilling (18) is negligible.

This result is fairly obvious. We have three parameters (\( a_i \) and \( \delta_i \) play no role) that are locally free to vary with respect to one another (\( \Gamma \) is an open set). Then we take all possible combinations of points of \( \Gamma \), and therefore a thick subset of \( \mathbb{R}^6 \). As an easy consequence the points of \( \Gamma_n \) that fulfill (18) form a negligible subset. However, a less obvious result is the following theorem which is based on the Alternative Principle mentioned in Section 4. Let us modify the example by assuming that \( a_i, \beta_i, \sigma_i, \sigma_j \) and \( \delta_i \) are not deep parameters themselves but functions of the deep parameters, i.e., functions defined on \( \Gamma \), analytic on \( \Gamma \) with the exception of a negligible subset. Then the following results hold:
(i) Each of the conditions (a) and (b) either holds for the whole \( \Gamma \) or for a negligible subset of \( \Gamma \).
(ii) If there exists a point of \( \Gamma \) such that neither (a) nor (b) holds, then \( Y_t \) and \( X_t \) are not cointegrated.
(iii) If (a) or (b) holds for the whole \( \Gamma \) then \( Y_t \) and \( X_t \) are cointegrated for any point of \( \Gamma_n \).

Some remarks are in order. First, the aggregation effect summarized in the theorem above, statement (a), requires that there are at least two non-redundant common permanent shocks driving the microvariables (condition (c) can hold only for a negligible subset of \( \Gamma \)). On the other hand, as we have reported in Section 3, we have strong evidence that several non-redundant permanent common shocks drive the microvariables corresponding to major macrovariables. Second, we have already observed that the construction of \( \Gamma_n \) consists in taking all possible combinations of agents picked up from \( \Gamma \). We impose no restriction whatever on the agents of a population. Moreover, many of our results refer to subsets of zero Lebesgue measure of \( \Gamma_n \). This means that we assume a state of profound ignorance about the distribution of the microparameters in empirical populations. If informations were available, these might lead to restrictions on the set \( \Gamma_n \). However, such restrictions would not necessarily imply fulfilment of (18).

To conclude this section, let us go back to the example and introduce the following modification:
\[
y_{it} = \alpha_i \beta_i + \beta_i w_{it} + (1 - \alpha_i) \beta_i w_{it},
\]
\[
x_{it} = \alpha_i \beta_i + \beta_i w_{it} + (1 - \alpha_i) \beta_i w_{it}.
\]
Here \( \delta_i, \beta_i, \sigma_i, \sigma_j \) and \( \delta_i \) are the microparameters. Now the microvariables \( x_{it} \) and \( x_{it} \) are no longer cointegrated, apart from a negligible subset of \( \Gamma \). However, this does not imply the impossibility of macro cointegration. The condition is
\[
\sum_{i=1}^{6n} \delta_i \sum_{i=1}^{n} \beta_i = \sum_{i=1}^{6n} \beta_i \sum_{i=1}^{n} \alpha_i,
\]
which describes a \((6n-1)\)-dimensional subset of \( \Gamma_6 \), which is \(6n\)-dimensional. Thus, unless there is a good reason to assume that condition (a) or (b) holds, macro cointegration is not more likely when microvariables are cointegrated than when micro cointegration does not occur.

6. Aggregation is not Necessarily Bad: The Case of Consumption
There is a very important case in recent literature in which aggregation effects contribute to reconciling theory and empirical evidence. Assume that labour income obeys the equation
\[
\Delta x_t = \sigma(L) \Delta z_t,
\]
where \( \sigma(L) = 1 + \alpha_1 L + \alpha_2 L^2 + \cdots \), and \( z_t \) is a white noise. According to the Life-Cycle Permanent-Income theory in its simplest version consumption should obey
\[
\Delta c_t = \sigma(L) \Delta z_t,
\]
where \( \beta = 1/(1 + \gamma) \), \( \gamma \) being a risk-free interest rate. This is a famous result by Hall (1978). There are three remarkable features in Hall's result.
(1) Consumption changes follow a white noise process uncorrelated with labour income changes at time \( t - k \), for any \( k > 0 \).
(2) As noticed by Deaton (1987), if labour income is persistent according to the measure proposed by Cochrane (1988), i.e., \( \sigma(1)^2 / \sigma(L) \), and \( z_t \) is a white noise, then consumption is more volatile than income.
(3) Consumption and income changes are driven by the same white noise, so that the vector \( (\Delta x_t, \Delta z_t) \) has rank one as a stochastic vector.
(4) As shown by Campbell (1987), total consumption and income, defined as the sum of labour income and asset returns, are cointegrated with cointegrating vector \( (1, -1) \). The first three features of Hall's model are at odds with aggregate empirical evidence. Consumption changes are positively autocorrelated and correlated with past values of income: this fact is known as "excess sensitivity" (Flavin 1981). The variance of consumption changes is much smaller than the variance of income changes; indeed, this is what Friedman's (1956) original permanent-income theory
was designed to explain. At the same time, there is evidence indicating that income is persistent. This problem is known as “excess smoothness” of consumption or “Deaton’s paradox”. Lastly, consumption changes form a rank two stochastic vector, some evidence on this is presented in Lippi and Forni (1977, Chapter 13).

Regarding (4), evidence is not clear-cut. However, cointegration between aggregate income and consumption is accepted by several authors. Let us show how a very simple heterogeneity assumption can solve the excess smoothness and sensitivity problems. Suppose that income of agent i evolves according to

$$\Delta x_{it} = (1 + aL)u_{it} + (1 + bL)x_{it}, \quad (19)$$

which is a simplification of the example used in Section 4. We suppose that the idiosyncratic shocks $\chi_{it}$ are orthogonal to one another and that they share the same variance: $\sigma^2 = \sigma^2_{\chi}$. Notice that there are two common parameters, $a$ and $b$, no individual parameters and that different incomes differ only for the idiosyncratic term $\chi_{it}$. Then assume as in Section 4, Assumption (B), that agent i observes only $\Delta x_{it}$, rather than its components. There exist a real $d$, $|d| \leq 1$, and a white noise $\eta_{it}$ such that

$$\Delta x_{it} = (1 + dL)\eta_{it} = (1 + aL)u_{it} + (1 + bL)x_{it}$$

(notice that since $a$, $b$ and $\sigma^2$ are common, the coefficient $d$ does not depend on i). Lastly, apply Hall’s theory to agent i to obtain

$$\Delta x_{it} = (1 + dL)\eta_{it} = (1 + aL)u_{it} + (1 + bL)x_{it}$$

$$\Delta \alpha_t = \frac{1 + d\beta}{1 + dL}(1 + aL)u_t + (1 + bL)x_t.$$ 

Indicating by $x_t$ and $c_t$ per-capita magnitudes we get

$$\Delta x_t = (1 + aL)u_t + (1 + bL)x_t, \quad (20)$$

In general $d \neq a$, so that $\Delta \alpha_t$ is not a white noise and is correlated with the past of $\Delta x_t$. Moreover, it is not difficult to see that, if $d$ is negative and $a$ is positive, then we have both a persistent income and a smooth consumption.

What about cointegration? For brevity, we do not introduce explicitly assets and total income in the model. However, the discussion in the previous Section will be sufficient to convince the reader that, despite heterogeneity, cointegration is retained in the macro model, consistently with empirical evidence. This because in this model, by property (4), all individual long-run propensities to consume are equal to 1.

The possibility that heterogeneity and incomplete information might explain the excess sensitivity and/or the excess smoothness puzzles, has been noted in Lippi (1987).

---

8 In Pischke (1995) an important step forward has been obtained by using information coming from micro data. Pischke found that model (19), estimated for a USA panel of household data, provided a positive $a$, a negative $b$, and a $\sigma^2_{\chi}$ much bigger than $\sigma^2_{\chi}$, so that negative autocorrelation prevails in individual incomes. As a consequence $d$ is negative, so that per-capita consumption changes, according to model (20), should exhibit positive autocorrelation, positive correlation with past income change and low variance, consistently with empirical evidence.

Although extremely interesting, model (20) is a rank-one vector. For that matter, only one common shock is present in the micromodel, so that both the macrovariables are driven by the same shock. On the other hand, we know that only one common shock in individual incomes is unrealistic. Thus let us consider a micromodel for incomes with two common shocks:

$$\Delta x_{it} = (1 + a_{1I}L)u_{it} + (1 + a_{2I}L)\eta_{it} + (1 + b_{1I}L)x_{it}. \quad (21)$$

Again, we employ the univariate representation

$$\Delta x_{it} = (1 + d_{1I}L)\eta_{it} = (1 + a_{1I}L)u_{it} + (1 + a_{2I}L)\eta_{it} + (1 + b_{1I}L)x_{it}$$

to obtain individual consumption

$$\Delta \alpha_t = (1 + d_{1I}L)\eta_{it} = \frac{1}{n} \sum_{i=1}^{n} (1 + d_{1I}L)(1 + a_{1I}L)u_{it} + (1 + a_{2I}L)\eta_{it} + (1 + b_{1I}L)x_{it}.$$ 

Taking per-capita magnitudes

$$\Delta x_t = (1 + a_{1I}L)u_t + (1 + a_{2I}L)\eta_{it}$$

$$\Delta \alpha_t = \frac{1}{n} \sum_{i=1}^{n} (1 + d_{1I}L)(1 + a_{1I}L)u_{it} + (1 + a_{2I}L)\eta_{it} + (1 + b_{1I}L)x_{it}, \quad (22)$$

where $\bar{a}_i = \sum_i a_{ii}/n$. The rank of ($\Delta x_t$, $\Delta \alpha_t$) is less than two if and only if

$$\det \begin{pmatrix} 1 + a_{1I} & 1 + a_{2I} \\ \sum_{i=1}^{n} d_{1I} + a_{1I} & \sum_{i=1}^{n} d_{1I} + a_{2I} \\ \sum_{i=1}^{n} 1 + d_{1I} & \sum_{i=1}^{n} 1 + d_{1I} \\ 1 + a_{1I} & 1 + a_{2I} \\ \sum_{i=1}^{n} 1 + d_{1I} + a_{1I} & \sum_{i=1}^{n} 1 + d_{1I} + a_{2I} \end{pmatrix} = 0. \quad (23)$$

It is not difficult to get convinced that if the coefficients $a_{1I}$, $a_{2I}$, $b_I$ and $\sigma^2_{\chi}$ have sufficient freedom of variation with respect to one another, then (23) will hold only for a negligible subset of $\Gamma_n$. Formally, this is an application of the Alternative Principle: since (23) is an algebraic equations between the coefficients of (22), and since such coefficients are analytic functions defined on $\Gamma_n$, apart from a negligible subset, then if (23) does not hold for a point $\pi$ of $\Gamma_n$ it holds only on a negligible
subset of \( \Gamma_n \). On the other hand, under mild heterogeneity conditions, a point \( \beta \in \Gamma_n \), such that (23) does not hold is very easy to find.

In conclusion, allowing for heterogeneity of agents in a common-idiosyncratic micromodel, more than one common shock, and limited information of individual agents, only cointegration is implied at the macro level by Hall's model, whereas the puzzling implications are avoided. Even though several alternative solutions for these puzzles have been proposed, the one outlined above has the advantage of an explicit consideration of heterogeneity and aggregation.

7. Wold and autoregressive representations of the aggregate vector

Dealing with cointegration and with the consumption model we had only to consider representation (16) of the aggregate vector, in which the LHS we have the aggregate variables and on the RHS the common microshocks. Now, in order to outline some further results we need the Wold representation of the aggregate vector. Unlike (16), which is structural or semi-structural and has on the RHS a number of shocks usually bigger than the number of variables on the LHS, the Wold representation has the form

\[
Y_t = A_{11}(p, L)U_{11} + A_{12}(p, L)U_{21} \\
X_t = A_{21}(p, L)U_{11} + A_{22}(p, L)U_{21},
\]

(24)

where:

(1) \( A_{11}(p, L) \) is a power series in \( L \) whose coefficients are functions of \( p \in \Gamma_n \). Moreover, \( A_{11}(p, 0) = A_{12}(p, 0) = 0 \).

(2) \( (U_{11}, U_{21}) \) is a vector white noise. Comparison of (24) to (16) shows that \( U_{11} \) and \( U_{21} \) depend on \( p \); however we do not need to further complicate notation.

(3) \( A_{11}(p, L)A_{22}(p, L) - A_{12}(p, L)A_{21}(p, L) \) does not vanish in the open circle \( |s| < 1 \).

Forni and Lippi (1997, Chapter 10) show that the coefficients of \( A_{ij}(p, L) \) are continuous on \( \Gamma_n \) and analytic on \( \Gamma_n \) with the exception of a negligible subset. Therefore the Alternative Principle, stated in Section 3, can be applied. Furthermore, it is easily seen that the same Principle can also be applied to the coefficients of the power series \( B_{ij}(p, L) \) in the autoregressive representation obtained by inverting (24), i.e.

\[
B_{11}(p, L)Y_t + B_{12}(p, L)X_t = v_{1t} \\
B_{21}(p, L)Y_t + B_{22}(p, L)X_t = v_{2t}.
\]

(25)

This result has important consequences on several aggregation problems. Here we shall discuss two of them, namely Granger-causality and structural VAR models.

The fact that causality relations are destroyed by aggregation has firstly been pointed out in Lippi (1988). Results on VAR models are presented in Blanchard and Quah (1989). Here we follow Lippi and Forni (1997, Chapters 11 and 12), in which further references can be found.

Since representation (23) has been obtained by inverting (24), we have that \( B_{11}(p, 0) = B_{22}(p, 0) = 1, B_{12}(p, 0) = B_{21}(p, 0) = 0 \), i.e. the equations are the projections of \( Y_t \) and \( X_t \) respectively on past values of \( Y_t \) and \( X_t \). By definition, \( Y_t \) does not Granger-cause \( X_t \) if \( B_{21}(p, L) = 0 \), this meaning that past values of \( Y_t \) do not help in predicting \( X_t \) once the information contained in the past of \( Y_t \) has been fully exploited. The Alternative Principle entails that if there exists a \( \beta \in \Gamma_n \) such that \( Y_t \) Granger-causes \( X_t \), then the subset of \( \Gamma_n \) where \( Y_t \) does not Granger-cause \( X_t \) is negligible.

Now the question is: assuming that \( y_t \) does not Granger-cause \( x_{1t} \), can we conclude that \( y_t \) does not Granger-cause \( x_{2t} \)? The answer is negative and an illustration of the result can be given by using the example of Section 2:

\[
y_t = a_1 x_{1t}, \quad x_{1t} = (1 + a_2) u_{1t},
\]

where \( a_1 \neq a_2 \) and the micromodels \( y_{it} \) and \( x_{1t} \) do not help in predicting \( x_{2t} \) once the information contained in the past of \( Y_t \) has been fully exploited. The Alternative Principle entails that if there exists a \( \beta \in \Gamma_n \) such that \( Y_t \) Granger-causes \( X_t \), then the subset of \( \Gamma_n \) where \( Y_t \) does not Granger-cause \( X_t \) is negligible.

Now the question is: assuming that \( y_t \) does not Granger-cause \( x_{1t} \), can we conclude that \( y_t \) does not Granger-cause \( x_{2t} \)? The answer is negative and an illustration of the result can be given by using the example of Section 2:

\[
y_t = a_1 x_{1t}, \quad x_{1t} = (1 + a_2) u_{1t},
\]

Like in Section 2 we assume that the \( u_{it} \) are mutually orthogonal at any lead and lag and that there are two agents. The aggregate equations are

\[
Y_t = a_1(1 + a_1)u_{1t} + a_2(1 + a_2)u_{2t}, \\
X_t = (1 + a_1)u_{1t} + (1 + a_2)u_{2t}.
\]

The corresponding Wold representation is obtained by normalizing:

\[
Y_t = \begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = Q \begin{pmatrix} 1 + a_1 \\ 1 + a_2 \end{pmatrix} \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix},
\]

(26)

where \( Q = (a_1 - a_2)^{-1} \). The (2,1) entry of the corresponding autoregressive representation is

\[
(a_1 - a_2) L \begin{pmatrix} 1 \\ 1 + a_2 \end{pmatrix}.
\]

Thus if \( a_1 \neq a_2 \) and \( \beta \neq a_2 \) the macrovariable \( Y_t \) Granger-causes the macrovariable \( X_t \), even though no such Granger-causation occurs at the micro level. Here again we see how aggregation effects occur when both the independent variables and the responses of the agents are heterogeneous. On the other hand, given a micromodel, if there exists a point in \( \Gamma_n \) for which such aggregation effect occurs, that same effect occurs almost everywhere in \( \Gamma_n \).

Let us now turn to VAR models. Suppose that a VAR is estimated for the macrovariables \( Y_t \) and \( X_t \):

\[
B(L) \begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix}.
\]

Let \( A(L) = B(L)^{-1} \) and write

\[
\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = A(L) \begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix}.
\]

Suppose also that according to our theory the variables \( Y_t \) and \( X_t \) are driven by a supply shock \( W_{1t} \) and a demand shock \( W_{2t} \), and that (i) \( W_{1t} \) and \( W_{2t} \) have
unit variance and are orthogonal at any lead and lag, (ii) the shock $W_t$ has no contemporaneous effect on $X_t$. This is sufficient to identify $W_t$ and $W_{2t}$, and the matrix $C(L)$ such that
\[
\begin{pmatrix}
Y_t \\
X_t
\end{pmatrix} = C(L) \begin{pmatrix}
W_t \\
W_{2t}
\end{pmatrix},
\]
where $C_{00}(0) = 0$. Now assume that the microvariables are driven by $h$ common supply shocks and $k$ common demand shocks
\[
\begin{pmatrix}
y_{kt} \\
x_{kt}
\end{pmatrix} = \begin{pmatrix}
c_{11}(L) & c_{12}(L) \\
c_{21}(L) & c_{22}(L)
\end{pmatrix} \begin{pmatrix}
\epsilon_{kt} \\
\eta_{kt}
\end{pmatrix},
\]
where $c_{11}(L)$ and $c_{21}(L)$ are $1 \times h$ while $c_{12}(L)$ and $c_{22}(L)$ are $1 \times k$. Moreover, assume that $c_{22}(0) = 0$, so that our identification criterion is ‘correct’, i.e. not inconsistent with the underlying micromodel. Aggregating (27) and using (26) we obtain
\[
\begin{pmatrix}
W_{1t} \\
W_{2t}
\end{pmatrix} = C(L)^{-1} \begin{pmatrix}
\sum c_{11}(L) & \sum c_{12}(L) \\
\sum c_{21}(L) & \sum c_{22}(L)
\end{pmatrix} \begin{pmatrix}
\epsilon_{kt} \\
\eta_{kt}
\end{pmatrix},
\]
where $W_{1t}$ and $W_{2t}$ appear as linear combination of the $\epsilon_{kt}$’s and the $\eta_{kt}$’s. The question is: can we state that $W_{1t}$ is a combination of the micro supply shocks $\epsilon_{kt}$ only, so that it is reasonable to call $W_{1t}$ an aggregate supply shock, or a mixing occurs? Similarly, can we state that $W_{2t}$ is a demand shock? The answer is that, if the mixing occurs for one point of $\Gamma_m$, then it occurs almost everywhere. Moreover, a mild heterogeneity is sufficient to generate the mixing. Thus, under heterogeneity, the aggregate supply shock is a combination of both the supply and demand microshocks.

8. Conclusions

In this paper we have shown a number of unpleasant aggregation effects. If agents are heterogeneous, a macroequation can be dynamic even though the corresponding microequations are static. Cointegration is destroyed by aggregation, unless either the micro cointegration coefficients are equal, or there is only one permanent common shock driving the microvariables. When consumers have limited information, the cross-correlation and volatility properties of consumption and income changes implied by Hall’s model are lost at the macro level. In general, unidirectional Granger-causality is not robust with respect to aggregation. Lastly, the

shocks appearing in a structural VAR representation result from a mixing of both corresponding and non-corresponding micro shocks.

We can summarize all of these results by saying that when the representative-agent assumption is dropped and heterogeneity among individuals is introduced, we cannot expect that a macro model shares the same dynamic properties as the underlying micro model. As a consequence, given an estimated macroequation, we cannot interpret its dynamic shape or other features as revealing something about the behaviour of individual agents. In the same way, if the estimated macro parameters fail to fulfill the restrictions implied by the micro theory, this is not a good reason to reject the micro theory. Thus, on one hand, aggregation is not so bad after all, since theory could be reconciled with evidence, as we have seen for the permanent income model. But, on the other hand, it should be recognized that additional difficulties arise in macroeconomic modeling. In order to obtain testable implications at the macro level, information on the joint behaviour of individual independent variables is needed.

Here we have proposed a model for the independent variables: the dynamic factor analytic model, generalized to allow for cross-correlated idiosyncratic components. This model is flexible enough to accommodate a large amount of heterogeneity, while retaining a reasonably parsimonious parameterization. A remarkable feature of the model is that, when the number of individuals is large, the idiosyncratic components die out with aggregation. This enormously simplifies things, since microvariables moving in a huge-dimensional space are made consistent with macrovariables driven by a small-dimensional vector of shocks.

Unfortunately, long series of individual data are seldom available, so that information on the number of common shocks and the cross-sectional distribution of individual response functions may be very difficult to collect. Nevertheless, we do not think that such difficulties should convince us to stick to the representative-agent practice, which amounts to transforming a complete lack of information about a distribution in the assumption that such distribution is concentrated on a single point of the microparameter space.

References


Fabio Canova [1997] "Una Cifra Storica per la Valutazione di Investimenti" pp. 17

Elene Forni, Lucrezia Reichlin [1996] "Let's get Real on Regional Cohesion" pp. 22

Stefano Ermotti [1997] "Elaborazione Automatica dei Dati" pp. 40

Paolo Bosi e Massimo Matteuzzi [1997] "Nuovi Risultati sull'Equivalenza di Lungo Periodo e il ruolo dei licei e delle scuole di familiari" pp. 21


Antonio Ribba [1998] "Analytical and numerical solution of the cumulative assignment problem with a well-structured tableau" pp. 25

Paolo Bosi, Maria Cecilia Guarini e Paolo Silvestri [1998] "La spesa pubblica nel consenso Modena: Rapporto intermedio" pp. 37