INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Model Evaluation in Macroeconometrics: from early empirical macroeconomic models to DSGE models

Carlo A. Favero

Working Paper n. 327

September 2007

IGIER – Università Bocconi, Via Salasco 5, 20136 Milano –Italy
http://www.igier.uni-bocconi.it

Model Evaluation in Macroeconometrics: from early empirical macroeconomic models to DSGE models

Carlo A. Favero*
IGIER Università Bocconi and CEPR

September 2007

Abstract

This paper reconsiders the developments of model evaluation in macroeconometrics over the last forty years. Our analysis starts from the failure of early empirical macroeconomic models caused by stagflation in the seventies. The different diagnosis of this failure are then analyzed to classify them in two groups: explanations related to problems in the theoretical models that lead to problems in the identification of the relevant econometric model and explanations related to problems in the underlying statistical model that lead to misspecification of the relevant econometric model. Developments in macroeconometric model evaluation after the failure of the Cowles foundation models are then discussed to illustrate how the different critiques have initiated different approaches in macroeconometrics. The evolution of what has been considered the consensus approach to macroeconometric model evaluation over the last thirty years is then followed. The criticism moved to Cowles foundation models in the early seventies might apply almost exactly to DSGE-VAR model evaluation in the first decade of the new millenium. However, the combination of general statistical model, such as a Factor Augmented VAR, with a DSGE model seems to produce forecasts that perform better than those based exclusively on the theoretical and on the statistical model.

JEL Classification Numbers: C52, E17

Keywords: Macroeconometrics, Model Evaluation

*Paper prepared for the "Conference in honor of David Hendry", Oxford 23-25 August 2007. I am grateful to David Hendry, Rob Engle, Chris Gilbert,Kevin Hoover, Duo Qin, Bruno Sitzia, Conference participants and a referee for useful comments. All errors are mine.

1 Introduction

This paper reconsiders the developments of model evaluation in macroeconometrics over the last forty years. Our analysis starts from the failure of early empirical macroeconomic models caused by stagflation in the seventies. The different diagnoses of this failure are then analyzed to classify them in two groups: explanations related to problems in the theoretical models that lead to problems in the identification of the relevant econometric model and explanations related to problems in the underlying statistical model that lead to misspecification of the relevant econometric model.

Developments in macroeconometric model evaluation are then discussed to illustrate how the different critiques have initiated different approaches in macroeconometrics. The evolution of what has been considered the consensus approach to macroeconometric model evaluation over the last thirty years is then followed. The statistical based evaluation of the failure of empirical macroeconomic models has generated a dynamic modelling strategy that has posed great emphasis on the statistical model and very little emphasis on the explicit modelling of the economy based on intertemporal optimization. This approach has failed to reach the consensus of the profession. The evolution of what has been instead considered the consensus approach to macroeconometric model evaluation over the last thirty years is then followed. VAR based model evaluation of Real Business Cycle (RBC) models and model evaluation in Bayesian Dynamic Structural General Equilibrium (DSGE) models are discussed to show how the profession has become progressively more dominated by the theory related approach to model evaluation. The logic of the criticism moved to Cowles foundation models in the early seventies might apply almost exactly to DSGE-VAR model evaluation in the first decade of the new millennium. However, the combination of general statistical model, such as a Factor Augmented VAR , with a DSGE model seems to produce forecasts that perform better than those based exclusively on the theoretical and on the statistical model.

2 Model evaluation in empirical large econometric models

Ealry empirical large econometric models were derived in the sixties in tradition of the Cowles Commission. Theory was used in this model simply to suggest the list of regressors to be included in the estimated equation, and it was largely driven by the IS/LM framework , in which the supply side was left virtually unmodelled and relative price movements were not considered (see Fukac and Pagan(2006)). Large-scale models were obtained by specifying equations which described the determinants of variables in the national accounting identity for GDP e.g. investment, consumption. This approach aimed at the quantitative evaluation of the effects of modification in the variables controlled by the policy-maker (the instruments of economic policy) on the macroeconomic variables which represent the final goals of the policy-maker. Policy variables were considered exogenous. Simultaneity was one of the main concern of Cowles Commission. However, estimation was of much greater concern than model evaluation and this traditional approach to econometrics is well charactirezed by a proliferation of estimators rather than by its attention to model evaluation and diagnostic and misspecification testing (see Hendry (1976), Qin(1993)). Klein-Golberger(1955) is a typical example of this modelling strategy. A general representation of these early empirical models could be given as follows:

$$\begin{aligned}
\mathbf{A}\begin{pmatrix}\mathbf{Y}_t\\ \mathbf{P}_t\end{pmatrix} &= \mathbf{C}_1(L)\begin{pmatrix}\mathbf{Y}_{t-1}\\ \mathbf{P}_{t-1}\end{pmatrix}+\mathbf{C}_2(L)\left(\mathbf{X}_t\right)+\mathbf{B}\begin{pmatrix}\boldsymbol{\nu}_t^Y\\ \boldsymbol{\nu}_t^P\end{pmatrix}, \qquad (1)\\
\begin{pmatrix}\boldsymbol{\nu}_t^Y\\ \boldsymbol{\nu}_t^M\end{pmatrix} \mid \mathbf{X}_t, I_{t-1}\Bigg) &\sim (\mathbf{0},\mathbf{I})\,.
\end{aligned}$$

The vector of n variables of interest is partitioned into three subsets: $\mathbf{Y}$, which represents the vector of modelled macroeconomic variables of interest and $\mathbf{P}$, the vector of modelled policy variables, and $\mathbf{X}$ a vector of variables which are considered exogenous and left unmodelled. In fact, this type of models consider specify fully only $D\left(\mathbf{Y}_t,\mathbf{P}_t \mid \mathbf{X}_t, I_{t-1},\Theta\right)$,i.e. the joint conditional distribution of $\mathbf{Y}_t,\mathbf{P}_t$ given $\mathbf{X}_t$,the information set available, made of all lagged variables, and the relevant parameters Θ.

The probabilistic structure for the variables of interest is determined by the implied reduced form. This statistical model has the following representation:

$$\begin{aligned}
\begin{pmatrix}\mathbf{Y}_t\\ \mathbf{P}_t\end{pmatrix} &= D_1(L)\begin{pmatrix}\mathbf{Y}_{t-1}\\ \mathbf{P}_{t-1}\end{pmatrix}+\mathbf{D}_2(L)\left(\mathbf{X}_t\right)+\mathbf{u}_t, \qquad (2)\\
\mathbf{u}_t &= \begin{pmatrix}\mathbf{u}_t^Y\\ \mathbf{u}_t^P\end{pmatrix},\\
u_t &\mid I_{t-1} \sim n.i.d.\left(\mathbf{0},\sum\right),\\
\left(\begin{matrix}\mathbf{Y}_t\\ \mathbf{P}_t\end{matrix} \mid I_{t-1},\mathbf{X}_t\right) &\sim \left(D_1(L)\begin{pmatrix}\mathbf{Y}_{t-1}\\ \mathbf{P}_{t-1}\end{pmatrix}+\mathbf{D}_2(L)\left(\mathbf{X}_t\right),\sum\right).
\end{aligned}$$

This system specifies the statistical distribution for the vector of variables of interest conditional upon the information set available at time $t-1$. [1] In relating the structure of interest to the statistical model a crucial identification problem has to be solved, since there is more than one structure of economic interest which can give rise to the same statistical model for our vector of variables.

For any given structure,

$$\begin{aligned}
\mathbf{A}\begin{pmatrix} \mathbf{Y}_t \\ \mathbf{P}_t \end{pmatrix} &= \mathbf{C}_1(L)\begin{pmatrix} \mathbf{Y}_{t-1} \\ \mathbf{P}_{t-1} \end{pmatrix} + \mathbf{C}_2(L)\left(\mathbf{X}_t\right) + \mathbf{B}\begin{pmatrix} \boldsymbol{\nu}_t^Y \\ \boldsymbol{\nu}_t^P \end{pmatrix}, \qquad (3) \\
\begin{pmatrix} \boldsymbol{\nu}_t^Y \\ \boldsymbol{\nu}_t^P \end{pmatrix} \mid I_{t-1} &\sim \left(\mathbf{0}, \mathbf{I}\right),
\end{aligned}$$

which give rise to the observed reduced form (2) when the following restrictions are satisfied:

$$\begin{aligned}
\mathbf{A}^{-1}\mathbf{C}_1(L) &= D_1(L),\ \mathbf{A}^{-1}\mathbf{C}_2(L) = D_2(L) \\
\mathbf{A}\begin{pmatrix} \mathbf{u}_t^Y \\ \mathbf{u}_t^P \end{pmatrix} &= \mathbf{B}\begin{pmatrix} \boldsymbol{\nu}_t^Y \\ \boldsymbol{\nu}_t^P \end{pmatrix}.
\end{aligned}$$

There exists a whole class of structures which produce the same statistical model (2) under the same class of restrictions:

$$F\mathbf{A}\begin{pmatrix} \mathbf{Y}_t \\ \mathbf{P}_t \end{pmatrix} = F\mathbf{C}_1(L)\begin{pmatrix} \mathbf{Y}_{t-1} \\ \mathbf{P}_{t-1} \end{pmatrix} + F\mathbf{B}\begin{pmatrix} \boldsymbol{\nu}_t^Y \\ \boldsymbol{\nu}_t^P \end{pmatrix}, \qquad (4)$$

where F is an admissible matrix, i.e. it is conformable by product with $\mathbf{A}$, $\mathbf{C}_1(L)$, and $\mathbf{B}$, and $F\mathbf{A}$, $F\mathbf{C}_1(L)$, $F\mathbf{B}$ feature the same restrictions with $\mathbf{A}$, $\mathbf{C}_1(L)$, $\mathbf{B}$.

The identification problem is solved in the Cowles commission approach by imposing restrictions on the $\mathbf{A}$, $\mathbf{C}_1(L)$, and $\mathbf{B}$ matrices so that the only admissible F matrix is the identity matrix. This is typically achieved by attributing an exogeneity status to the policy variables, i.e. either by including them in the set of unmodelled variables $\mathbf{X}_t$ or by making $\mathbf{P}_t$ independent form present and past $\mathbf{Y}_t$.

The construction of diagnostics for model evaluation is related to the solution of the identification problem.. In fact, in the (very common) case

[1] In this case the statistical model is a VAR with some exogenous variables. When variables included in the VAR are non-stationary, the model can be re-parameterised as a VECM. In this case, after the solution of the identification problems of cointegrating vectors, the information set available at $t-1$ contains n lagged endogenous variables and r cointegrating vectors.

of over-identified models, a test of the validity of the over-identifying restrictions can be constructed by comparing the restricted reduced form implied by the structural model with the reduced form implied by the just-identified model in which each endogenous variables depend on all exogenous variables with unrestricted coefficients. The statistics are derived in Anderson and Rubin(1949) and Basman(1960). The logic of the test attributes a central role to the structural model. The statistical model of reference for the evaluation of the structural model is derived by the structural model itself.

2.1 An illustrative example

To see a simple illustration of this modelling tradition consider the following quarterly econometric model of the United States proposed by Gallaway-Smith(GS,1961):

$$
\begin{aligned}
Y_t &= C_t + I_t + G_t \\
C_t &= c_{1,0} + +c_{1,9} M_t + c_{1,11} Y_{t-1}^d + u_{1,t} \\
I_t &= c_{2,0} + c_{2,1} \left(Y_{t-1} - Y_{t-2}\right) + c_{2,13} R_{t-1} + u_{2,t} \\
G_t &= c_{3,0} + c_{3,3} G_{t-1} + u_{3,t}
\end{aligned}
$$

where all variables are seasonally adjusted in current prices, Y_t is GDP, Y_t^d is disposable income, C_t is consumption expenditure, M_t is the quantity of money, R_t is property income. The adopted model is estimated by least squares and in first difference of all variables are taken "...to reduce the extent of autocorrelation and heteroscedasticity..."[2] Very loose theoretical foundation are given for the adopted specification, for example an accelerator model is cited to justify the adopted specification for investment and no justification is made for the inclusion of the quantity of money in the consumption function.

To cast GS in our proposed general framework set $\mathbf{Y}_t' = \begin{pmatrix} C_t & I_t \end{pmatrix}$,$\mathbf{P}_t = G_t, \mathbf{X}_t' = \begin{pmatrix} Y_t^d & M_t & R_t \end{pmatrix}$. We have then:

[2] Gallaway and Smith(1961), p.380

$$
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} C_t \\ I_t \\ G_t \end{bmatrix} = \begin{bmatrix} c_{1,0} \\ c_{2,0} \\ c_{3,0} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ c_{2,1} & c_{2,1} & c_{2,1} \\ 0 & 0 & c_{3,3} \end{bmatrix} \begin{bmatrix} C_{t-1} \\ I_{t-1} \\ G_{t-1} \end{bmatrix} +
$$

$$
\begin{bmatrix} c & 0 & 0 \\ -c_{2,1} & -c_{2,1} & -c_{2,1} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} C_{t-2} \\ I_{t-2} \\ G_{t-2} \end{bmatrix} + \begin{bmatrix} 0 & c_{1,9} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} Y_t^d \\ M_t \\ R_t \end{bmatrix} +
$$

$$
\begin{bmatrix} c_{1,11} & 0 & 0 \\ 0 & 0 & c_{2,13} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Y_{t-1}^d \\ M_{t-1} \\ R_{t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{bmatrix}
$$

which is a clearly overidentified structure, in the sense that a very special form is assumed for the $\mathbf{A}$ matrix (no simultaneous feedback among the variables of interests) and many restrictions are introduced $\mathbf{C}_1(L)$ and $\mathbf{C}_2(L)$ polynomial matrices. The validity of these over-identifying restrictions is tested by assuming as benchmark the statistical model in which the vector of the variables of interest is projected on the whole information set I_{t-1} and the exogenous variable without imposing any restrictions:

$$
\begin{bmatrix} C_t \\ I_t \\ G_t \end{bmatrix} = \begin{bmatrix} d_{1,0} \\ d_{2,0} \\ d_{3,0} \end{bmatrix} + \begin{bmatrix} d_{1,1} & d_{1,2} & d_{1,3} \\ d_{2,1} & d_{2,2} & d_{2,3} \\ d_{3,1} & d_{3,2} & d_{3,3} \end{bmatrix} \begin{bmatrix} C_{t-1} \\ I_{t-1} \\ G_{t-1} \end{bmatrix} +
$$

$$
\begin{bmatrix} d_{1,4} & d_{1,5} & d_{1,6} \\ d_{2,4} & d_{2,5} & d_{2,6} \\ d_{3,4} & d_{3,5} & d_{3,6} \end{bmatrix} \begin{bmatrix} C_{t-2} \\ I_{t-2} \\ G_{t-2} \end{bmatrix} + \begin{bmatrix} d_{1,7} & d_{1,8} & d_{1,9} \\ d_{2,7} & d_{2,8} & d_{2,9} \\ d_{3,7} & d_{3,8} & d_{3,9} \end{bmatrix} \begin{bmatrix} Y_t^d \\ M_t \\ R_t \end{bmatrix} +
$$

$$
\begin{bmatrix} d_{1,10} & d_{1,11} & d_{1,12} \\ d_{2,10} & d_{2,11} & d_{2,12} \\ d_{3,10} & d_{3,11} & d_{3,12} \end{bmatrix} \begin{bmatrix} Y_{t-1}^d \\ M_{t-1} \\ R_{t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{bmatrix}
$$

Importantly the statistical model is dependent on the strucutral model adopted in that the exogeneity assumption for the variables included in the vector $\mathbf{X}_t$ is maintaned and the length of all distributed lags in the statistical model is determined by the length of the distributed lag in the strucutral model.

3 The different diagnoses of the failure of early models

Stagflation in the late seventies condemned the early empirical models, that

> "...did not represent the data, ... did not represent the theory... were ineffective for practical purposes of forecasting and policy evaluation..." (Pesaran and Smith 1995).

Different explanations of the failure of this approach were proposed. We can classify them into diagnoses related to the solution of the structural identification problem and diagnoses related to the (lack of) solution of the statistical identification problem.

The distinction between structural and statistical identification has been introduced by Spanos(1990). As illustrated in the previous section structural models can be viewed statistically as a reparameterization, possibly (in case of over-identified models) with restrictions, of the reduced form. Structural identification refers to the uniqueness of the structural parameters, as defined by the reparameterization and restriction mapping from the statistical parameters in the reduced form, while statistical identification refers to the choice of a well-defined statistical model as reduced form.

In Spanos' terminology a well-defined statistical model is one whose underlying assumptions are valid for the data-chosen. A well-defined statistical model should be based valid conditioning and it should satisfy the underlying statistical assumptions of normality, homoscedasticity, and temporal independence of residuals. We have seen in the previous section that unrestricted reduced form models are determined by the underlying structural model and not by the data. It is then well possible that the just-identified model chosen as a benchmark to evaluate the valicity of over-identifying restrictions does not provide an adequate statistical description of the data. Reconsider the GS model and think of a situation where there is a feedback between the variables in $\mathbf{Y}_t$ and the variables in $\mathbf{X}_t$, this would clearly invalidate the unrestricted reduced form in which $\mathbf{X}_t$ is left unmodelled. Similar problems will occur if the length in the polynomial lags in the Data Generating Process are different and longer than the ones included in the structural model.

The Lucas(1976) critique and the Sims(1980) critique are the diagnoses related to the solution of the identification problem.

3.1 Structural Identification

Lucas questions the superexogeneity status of the policy variables. Lucas attacks the identification scheme proposed by the Cowles Commission by pointing out that these models do not take expectations into account explicitly and, therefore, the identified parameters within the Cowles Commission approach are a mixture of 'deep parameters' describing preference and technology into the economy, and expectational parameters which, by their nature, are not stable across different policy regimes. The main consequence of such instability is that traditional structural macro-models are useless for the purpose of policy simulation. To illustrate the point, assume the following DGP, in which expected policy matters for the determination of macroeconomic variables in the economy:

$$\begin{pmatrix} \mathbf{Y}_t \\ \mathbf{P}_t \end{pmatrix} = \begin{pmatrix} c_{01} \\ c_{02} \end{pmatrix} + \begin{pmatrix} c_{11} & c_{12} \\ 0 & c_{22} \end{pmatrix} \begin{pmatrix} \mathbf{Y}_{t-1} \\ \mathbf{P}_{t-1} \end{pmatrix} + \begin{pmatrix} \gamma \\ 0 \end{pmatrix} \left(\bar{\mathbf{P}}_{t+1}^{e} \right) + \begin{pmatrix} \mathbf{u}_t^Y \\ \mathbf{u}_t^P \end{pmatrix}. \tag{5}$$

A Cowles Commission model is estimated without explicitly including expectations and it will have the following specification:

$$\begin{pmatrix} 1 & a_{12} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{Y}_t \\ \mathbf{P}_t \end{pmatrix} = \begin{pmatrix} d_{01} \\ c_{02} \end{pmatrix} + \begin{pmatrix} d_{11} & d_{12} \\ 0 & c_{22} \end{pmatrix} \begin{pmatrix} \mathbf{Y}_{t-1} \\ \mathbf{P}_{t-1} \end{pmatrix} + \begin{pmatrix} \mathbf{u}_t^Y \\ \mathbf{u}_t^P \end{pmatrix}. \tag{6}$$

Under the assumed DGP the restrictions $a_{12} = \gamma c_{22}$ and $d_{01} = \gamma c_{02}$ apply and simulation of alternative policy regimes, i.e. alternative values of c_{02} and c_{22},cannot be implemented by keeping the estimated parameters constant.

Sims reinforced Lucas point by labelling the Cowles Commission restrictions as "incredible", in fact no variable can be deemed as exogenous in a world of forward-looking agents whose behaviour depends on the solution of an intertemporal optimization model. Optimality of policy cannot be consistent with the restrictions that $\mathbf{A}$, $\mathbf{C}_1(L)$, and $\mathbf{B}$ lower triangular. Note also that by invalidly imposing such restrictions the model might induce a spurious statistic efficacy of policy in the determination of macroeconomic variables. Endogeneity of policy does generate correlations between macroeconomic and policy variables, which, by invalidly assuming policy as exogenous, can be interpreted as a causal relation running from policy to the macroeconomic variables.

3.2 Statistical Identification

The diagnosis related to the specification of the statistical model explains the ineffectiveness of large models for the practical purposes of forecasting and policy as due to their incapability of representing the data. The root of the failure of the traditional approach lies in the little attention paid to the statistical model implicit in the estimated structure. Any identified structure that is estimated without checking that the implicit statistical model is an accurate description of the data is bound to fail if the statistical model is not valid. Spanos(1990) considers the case of a simple demand and supply model to show how the reduced form is ignored in the traditional approach. The example is based on the market for commercial loans discussed in Maddala(1988). Most of the widely used estimators allow the derivation of numerical values for the structural parameters without even seeing the statistical models represented by the reduced form. Following this tradition the estimated (by 2SLS) structural model is:

$$
\begin{aligned}
q_t^d &= \underset{(74.31)}{-210.43} - \underset{(1.60)}{20.2}\, r_t + \underset{(2.84)}{40.77} br_t + \underset{(0.45)}{2.34} x_t + \hat{u}_t^d \\
q_t^s &= \underset{(13.96)}{-87.94} + \underset{(1.89)}{6.09} r_t - \underset{(2.27)}{7.08}\, i_t + \underset{(0.008)}{0.334} d_t + \hat{u}_t^s \\
\xi_1(1) &= 28.106, \ \xi_2(1) = 4.5
\end{aligned}
$$

where r_t is the average prime rate, br_t the Aaa corporate bond rate, x_t is the industrial production index, i_t the three-month bill rate, d_t total bank deposits and q_t commercial loans. q_t and r_t are the endogenous variables, br_t, x_t, i_t and d_t are taken as, at least, weakly exogenous variables and no equation for these variables is explicitly estimated. Given that there are two omitted instruments in each equation one over-identifying restrictions is imposed both in the demand and in the supply equation, the validity of such restrictions is tested via the Anderson-Rubin tests($\xi_1(1)$ *and* $\xi_2(1)$), that leads to rejection of the restrictions at the 5 per cent level in both cases, while in the second equation the restrictions cannot be rejected at the 1 per cent level. Estimation of the statistical model (i.e. the implicit unrestricted reduced form) yields:

$$
\begin{aligned}
q_t &= \underset{(21.05)}{-128.20} - \underset{(0.810)}{3.007} i_t + \underset{(1.236)}{7.078} br_t + \underset{(0.156)}{0.497} x_t + \underset{(0.011)}{0.281} d_t + \hat{u}_{1t} \\
r_t &= \underset{(3.02)}{1.864} + \underset{(0.116)}{0.771} i_t + \underset{(0.178)}{0.763} br_t + \underset{(0.022)}{0.008} x_t - \underset{(0.001)}{0.005} d_t + \hat{u}_{2t}
\end{aligned}
$$

where the underlying statistical assumptions of linearity, homoscedasticity, absence of autocorrelation and normality of residuals are all strongly rejected. On the basis of this evidence the adopted statistical model is not considered as appropriate. An alternative model is then considered allowing for a richer dynamic structure (two lags) in the reduced form, such dynamic specification is shown to provide a much better statistical mode for the data than the static reduced from. Of course, the adopted structural model implies many more over-identifying restrictions than the initial one. When tested the validity of these restrictions is overwhelmingly rejected for both the demand and the supply equations. This evidence leads Spanos to conclude that statistical identification should be distinguished from structural identification. Statistical identification refers to the choice of a well-defined statistical model, structural identification refers to the uniqueness of the structural parameters as defined by the reparameterization and restriction mapping from the statistical parameters. Lucas and Sims concentrate on model failure related to structural identification problems but models can fail independently from structural identification problems as a consequence of statistical identification problems.

4 Model specification and model diagnostic when statistical identification matters.

The diagnosis related to the specification of the statistical model gave rise to the LSE approach to macroeconometric modelling[3] and to the "structural cointegrating VAR" approach..

There are several possible causes for the inadequacy of statistical models implicit in structural econometric models: omission of relevant variables, or of the relevant dynamics for the included variables, or invalid assumptions of exogeneity. The LSE solution to the specification problem is the theory of reduction. Any econometric model is interpreted as a simplified representation of the unobservable data generating process (DGP). For the representation to be valid or 'congruent', to use Hendry's own terminology, the information lost in reducing the DGP to its adopted representation, given by the adopted specification, must be irrelevant to the problem at hand. Any statistical model is reinterpreted as a simplified representation of the DGP, for a model to valid the information lost in the simplification should be irrelevant for the problem at hand. Large econometric models concentrate on $D\left(\mathbf{Y}_t, \mathbf{P}_t \mid I_{t-1}, \mathbf{X}_t, \boldsymbol{\theta}\right)$, this can be interpreted as a simplified representation of $D\left(\mathbf{z}_t \mid \mathbf{Z}_{t-1}, \boldsymbol{\theta}\right)$, where $\mathbf{z}$ contains all economic variables. For a model to be valid the several marginalization steps implictly taken to obtain $D\left(\mathbf{Y}_t, \mathbf{P}_t \mid I_{t-1}, \mathbf{X}_t, \boldsymbol{\theta}\right)$ from $D\left(\mathbf{z}_t \mid \mathbf{Z}_{t-1}, \boldsymbol{\theta}\right)$ must not induce a loss of relevant information. Adequacy of the statistical model is evaluated by ana-

[3] The LSE approach was initiated by Denis Sargan but owes its diffusion to a number of Sargan's students and is extremely well described in the book by David Hendry(1995).

lyzing the reduced form, i.e. by checking statistical identification via test on residuals and on the exogeneity assumptions[4]. Therefore, the prominence of the structural model, with respect to the reduced form representation in the Cowles Commission approach to identification and specification is reversed. The LSE approach starts its specification and identification procedure with a general dynamic reduced form model. The congruency of such a model cannot be directly assessed against the true DGP, which is unobservable. However, model evaluation is made possible by applying the general principle that congruent models should feature true random residuals; hence, any departure of the vector of residuals from a random normal multivariate distribution should signal a mis-specification. Stationarity of the statistical model is a crucial feature when the model has to be simulated. Non-stationarity in macroeconomic time-series is treated in the LSE methodology by treating the reduced form VAR as a cointegrated VAR, by imposing rank imposing rank reduction restrictions in the matrix determining the long-run equilibria of the system and by solving the identification problem of cointegrating vectors (see Johansen, 1995). Once the baseline model has been validated, the reduction process begins by simplifying the dynamics and reducing the dimensionality of the model by omitting the equations for those variables for which the null hypothesis of exogeneity is not rejected. Different tests are proposed for the different concepts of exogeneity proposed by Engle, Hendry and Richard (1983) and even the validity of the Lucas-critique becomes a testable concept (Engle and Hendry(1993), Hendry(1988)). If models are considered as the product of a reduction process, a very natural framework emerges to test for exogeneity. We have seen that early econometric model concentrate $D\left(\mathbf{Y}_t, \mathbf{P}_t \mid I_{t-1}, \mathbf{X}_t, \boldsymbol{\theta}\right)$. The validity of this assumption can be evaluated by specifying $D\left(\mathbf{Y}_t, \mathbf{P}_t, \mathbf{X}_t \mid I_{t-1}, \boldsymbol{\Psi}\right)$, where $\boldsymbol{\theta}$ is a subset of $\boldsymbol{\Psi}$, by evaluating if the inference based on $D\left(\mathbf{Y}_t, \mathbf{P}_t \mid I_{t-1}, \mathbf{X}_t, \boldsymbol{\theta}\right)$ differs from the one obtained starting from the joint distribution to derive the relevant conditional distribution by marginalization and integration. Different use of the model (estimation, forecasting and simulation) generate different condition for the validity of the reduction and therefore different concepts of exogeneity are introduced (see Hendry and Richard(1993)). The product of the process of reduction is a statistical model for the data, possibly discriminating between short-run dynamics and the long-run equilibria. Only after this validation procedure the structural model can be identified and estimated. A just-identified specification does not require any further test-

[4] For an extensive description of the reduction process, the concepts of exogeneity and the tests for the validity of reduction see Hendry(1995).

ing, as its implicit reduced form does not impose any further restrictions on the baseline statistical model. The validity of over-identified specification is instead tested by evaluating the validity of the restrictions implicitly imposed on the general reduced form. The most popular applications of the general-to-specific specification strategy are in the area of money demand (Baba et al.(1992)) and aggregate consumption expenditure (see, for example, Hendry et al.(1990)).

In practice, the LSE approach has almost exclusively concentrated on the statistical diagnosis of the failure of Cowles foundation models and has brought more attention to the dynamic specification and the long-run properties of models built in the Cowles foundation tradition and used by policy makers, but it has paid much less attention to the possibility of specifying a forward-looking microeconomically founded model consistent with the theory based diagnosis for the failure of traditional Cowles foundation models (an interesting example of this approach can be found in Juselius and Johansen(1999)). In a recent paper Juselius and Franchi(2007) propose to formulate as a set of hypotheses on the long-run structure of a cointegrated VAR all the basic assumptions underlying a theoretical model, they also argue in favour of using an identified cointegrated VAR as a way of structuring the data that offers a number a "sophisticated" stylized facts to be matched by empirically relevant theoretical models.

The idea of constructing empirical models based on the belief that economic theory is most informative about the long-run relationships between the relevant variables has been further developed by Hashem Pesaran and a number of co-authors (see, for example Pesaran and Shin (2002), Garratt et al.(2007)) in the so-called "structural cointegrating VAR approach". This approach is based on testing theory based over-identifying restrictions on the long-run relations to provide a statistically coherent framework for the analysis of the short-run. In practice, the implementation is based on log-linear VARX model were the baseline VAR model to analyze macroeconomic variables is augmented with weakly exogenous variables such as oil prices or country specific foreign variables. Theory based cointegrated relationships are tested and, whenever not rejected, imposed on the specification. No restrictions are imposed on the short-run dynamics on the model except for the, inevitable, choice of the lag length for the VARX. Models are then used to evaluate the effect of policies via generalized impulse response functions(see Pesaran and Shin(1998)) and for forecasting.

5 Model Specification and Model Diagnostic when structural identification matters

The great critiques made clear that the quantitative analysis of effect economic policy should be based on theoretical models in which parameters describing tastes and technology are clearly identified. To achieve identification theretically funded rather than ad hoc models should be used. As a consequence Dynamic Stochastic General Equilibrium (DSGE) models were introduced. Following Sims(2002) we represent a general linear (or linearized around equilibrium) rational expectations model as follows:

$$\mathbf{\Gamma}_0 \mathbf{Z}_t = \mathbf{\Gamma}_1 \mathbf{Z}_{t-1} + C + \Psi \epsilon_t + \Pi \eta_t \tag{7}$$

$$\mathbf{Z}_t = \begin{bmatrix} \mathbf{Y}_t \\ \mathbf{P}_t \end{bmatrix} \tag{8}$$

Where C is a vector of constants, ϵ_t is an exogenously evolving random disturbance, η_t is a vector of expectations errors, $\left(E_t\left(\eta_{t+1}\right) = \mathbf{0}\right)$, not given exogenously but to be treated as part of the model solution. The forcing processes here are the elements of the vector ϵ_t, this typically contains processes like Total Factor Productivity or policy variables that are not determined by an optimization process. Policy variables set by optimization, typically included $\mathbf{Z}_t$, are naturally endogenous as optimal policy requires some response to current and expected developments of the economy[5]. Expectations at time t for some of the variables of the systems at time t+1 are also included in the vector $\mathbf{Z}_t$,whenever the model is forward looking. Model like (7)can be solved using standard numerical techniques (see, for example, Sims, 2002), and the solution can be expressed as:

$$\mathbf{Z}_t = \mathbf{A}_0 + \mathbf{A}_1 \mathbf{Z}_{\mathbf{t-1}} + \mathbf{R}\epsilon_t$$

where the matrices $\mathbf{A}_0, \mathbf{A}_1$,and $\mathbf{R}$ contain convolutions of the underlying model structural parameters. Note that, the solution is naturally represented as a VAR, of course it is a VAR potentially with stochastic singularity, as the dimension of the vector of shocks is typically smaller than that of the vector of variables included in the VAR. However, this problem is promptly solved by adding the appropriate number of measurement errors[6].

[5] See Appendix 1 for an example of this representation applied to a simple macroeconomic model.

[6] Expressing the solution of a DSGE as a VAR might also involve solving some non

5.1 An illustrative example

To illustrate DSGE we consider a simple model, that is easily comparable to the GS model described in one of the previous sections. The model is obtained considering an economy which features a representative household optimizing over consumption, real money holdings and leisure, a continuum of monopolistically competitive firms with price adjustment costs and a monetary policy authority which sets the interest rate. The model is driven by three exogenous processes which determine government spending, g_t, the stationary component of technology, z_t, and the policy shock, $\epsilon_{R,t}$. A full description of the model can be found in Woodford (2003). For the purpose at hand we focus on its log-linear representation which takes each variable as deviations from its trend. The specification follows Del Negro and Schorfheide (2004)(DS) and it reads:

$$
\begin{aligned}
\tilde{x}_t &= E_t\tilde{x}_{t+1} - \frac{1}{\tau}(\tilde{R}_t - E_t\tilde{\pi}_{t+1}) + (1-\rho_G)\tilde{g}_t + \rho_z\frac{1}{\tau}\tilde{z}_t & (9)\\
\tilde{\pi}_t &= \beta E_t\tilde{\pi}_{t+1} + \kappa\,(\tilde{x}_t - \tilde{g}_t) & (10)\\
\tilde{R}_t &= \rho_R\tilde{R}_{t-1} + (1-\rho_R)(\psi_1\tilde{\pi}_t + \psi_2\tilde{x}_t) + \epsilon_{R,t} & (11)\\
\tilde{g}_t &= \rho_g\tilde{g}_{t-1} + \epsilon_{g,t} & (12)\\
\tilde{z}_t &= \rho_z\tilde{z}_{t-1} + \epsilon_{z,t} & (13)
\end{aligned}
$$

where $\tilde{x}_t$ is the output gap, $\tilde{\pi}_t$ is the inflation rate, $\tilde{R}_t$ is the short-term interest rate and $\tilde{g}_t$ and $\tilde{z}_t$ are two AR(1) stationary processes for government and technology, respectively.

The first equation is an intertemporal Euler equation obtained from the households' optimal choice of consumption and bond holdings. There is no investment in the model and so output is proportional to consumption up to an exogenous process that describes fiscal policy. The net effects of these exogenous shifts on the Euler equation are captured in the process $\tilde{g}_t$. The parameter $0 < \beta < 1$ is the households' discount factor and $\tau > 0$ is the inverse of the elasticity of intertemporal substitution. This first equation describes the demand side of the model. The second equation is the forward-looking Phillips curve which describes the dynamics of inflation and κ determines the degree of the short-run trade-off between output and inflation. This second equation describes the supply side of the model and represent a feature which was totally absent from the fixed price GS model. The third equation

invertibility problems of the matrix governing the simulataneous relation among variables originally considered in the theoretical model. This problem is carefully discussed in Canova(2007).

is the monetary policy reaction function, that endogeneizes monetary policy to the state of the economy. The central bank follows a nominal interest rate rule by adjusting its instrument to deviations of inflation and output from their respective target levels. The shock $\epsilon_{R,t}$ can be interpreted as unanticipated deviation from the policy rule or as policy implementation error. Fiscal policy is simply described by an autoregressive process. The set of structural shocks is thus $\epsilon_t = (\epsilon_{R,t}, \epsilon_{g,t}, \epsilon_{z,t})'$ which collects technology, government and monetary shocks.

To cast the model in the form of :

$$\Gamma_0 \widetilde{\mathbf{Z}}_t = \Gamma_1 \widetilde{\mathbf{Z}}_{t-1} + C + \Psi \epsilon_t + \Pi \eta_t \tag{14}$$

Specify the relevant matrices as follows:

$$
\widetilde{\mathbf{Z}}_t = \begin{bmatrix} \widetilde{x}_t \\ \widetilde{\pi}_t \\ \widetilde{R}_t \\ \widetilde{R}_t^* \\ \widetilde{g}_t \\ \widetilde{z}_t \\ E_t\widetilde{x_{t+1}} \\ E_t\widetilde{\pi_{t+1}} \end{bmatrix} \quad \epsilon_t = \begin{bmatrix} \epsilon_t^R \\ \epsilon_t^G \\ \epsilon_t^Z \end{bmatrix} \quad \eta_t = \begin{bmatrix} \eta_t^x = x_t - E_{t-1}(x_t) \\ \eta_t^\pi = \pi_t - E_{t-1}(\pi_t) \end{bmatrix}
$$

$$
\Gamma_0 = \begin{bmatrix} 1 & 0 & \frac{1}{\tau} & 0 & -(1-\rho_g) & -\frac{\rho_z}{\tau} & -1 & -\frac{1}{\tau} \\ -\kappa & 1 & 0 & 0 & \kappa & 0 & 0 & -\beta \\ 0 & 0 & 1 & -(1-\rho_R) & 0 & 0 & 0 & 0 \\ -\psi_2 & -\psi_1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

$$
\Gamma_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_R & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \rho_G & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_Z & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}
$$

$$
\Psi = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \Pi = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}
$$

As a solution to (**??**), we obtain the following policy function

$$
\tilde{Z}_t = T(\theta)\,\tilde{Z}_{t-1} + R(\theta)\,\epsilon_t \tag{15}
$$

To provide the mapping between the observable data and those computed as deviations from the steady state of the model we set the following measurement equations, as in DS:

$$\begin{aligned}
\Delta \ln x_t &= \ln \gamma + \Delta \tilde{x}_t + \tilde{z}_t && (16)\\
\Delta \ln P_t &= \ln \pi^* + \tilde{\pi}_t && (17)\\
\ln R_t &= 4[(\ln R^* + \ln \pi^*) + \tilde{R}_t] && (18)
\end{aligned}$$

which can be also cast into matrices as:

$$Y_t = \Lambda_0(\theta) + \Lambda_1(\theta)\tilde{Z}_t + v_t \tag{19}$$

where $Y_t = (\Delta \ln x_t, \Delta \ln P_t, \ln R_t)'$, $v_t = 0$ and Λ_0 and Λ_1 are defined accordingly. For completeness, we write the matrices T, R, Λ_0 and Λ_1 as a function of the structural parameters in the model, $\theta = \left(\ln \gamma, \ln \pi^*, \ln r^*, \kappa, \tau, \psi_1, \psi_2, \rho_R, \rho_g, \rho_Z, \sigma_R, \sigma_g, \sigma_Z\right)'$: such a formulation derives from the rational expectations solution.

The evolution of the variables of interest, Y_t, is therefore determined by (14) and (19) which impose a set of restrictions across the parameters on the moving average (MA) representation. Finally, the MA representation is approximated by a finite order VAR representation.

6 Estimation and Evaluation of DSGE models

We have seen that a DSGE model identifies clearly tastes and technology parameters, that a solved DSGE model is a VAR, but parameters in the VAR are very complicated convolutions of the structural parameters and it is therefore very hard to pin down values for the parameters of interest, given the estimation of a reduced form VAR.

This difficulty explains the historical evolution of estimation and evaluation of DSGE models

Early DSGE models were called RBC models as they feature non-friction and therefore no role for policy(see, for example, Kydland and Prescott, 1982 and King, Plosser and Rebelo, 1988). This type of model contained a limited number of parameters, which have been often calibrated rather than estimated.

Calibration is extensively described in Cooley (1997), among others; he states (p. 56) that

> *calibration is a strategy for finding numerical values for the parameters of artificial economic worlds...[it] uses economic theory extensively as the basis for restricting a general framework and mapping that framework into the measured data.*

The aim of calibration is not to provide a congruent representation of the data, but simply to find values for the structural parameters of the model that are jointly compatible with the theory and the data in particular well-specified dimensions. Calibration proceeds in several stages. First, a preliminary, non-theoretical, inspection of the data identifies some general stylized facts that any economic model should internalize. The theoretical framework at hand, then, integrated by these observed stylized facts, provides the relevant parametric class of models. Once a particular model has been developed, it precisely defines the quantities of interest to be measured, and suggests how available measurements have to be reorganized if they are inconsistent with the theory.

Then, measurements are used to give empirical content to the theory, and in particular to provide empirically based values for the unknown parameters. Value for the parameters are then chosen by specifying some features of the data for the model to reproduce, by finding some one-to-one relationships between these features and the deep parameters of the model and by inverting these relationships.

In fact, calibration can be interpreted as a method of moments estimation procedure that focuses on a limited parameters' subset, setting only the discrepancy between some simulated and observed moments to zero. Christiano and Eichenbaum (1992) generalize this idea and propose a variant of Hansen's (1982) GMM procedure to estimate and assess stochastic general equilibrium models using specific moments of the actual data. These procedures are formal developments of the basic methodological approach, and share with standard calibration the focus on a limited set of previously selected moments, while standard econometric methods use, in principle, the whole available information set, weighting different moments exclusively according to how much information on them is contained in the actual data, as for example in the maximum likelihood methods.

Generally, not all parameters can be calibrated, simply because there are more unknown parameters than invertible relationships. A subset of them has to be left to more standard econometric techniques.

Once a parameterization is available, the model is simulated and different kinds of numerical exercises are performed. At this stage model evaluation can be also implemented. Model evaluation was initially conducted by assessing the ability of the model to reproduce some particular features (of course, the ones that are different from those used to calibrate it) of the data. The comparison between the properties of actual and model simulated series was initially based on an informal measures of their distance.Moreover as RBC models are usually solved by linearizing them around equilibrium raw data cannot be used to generate the set of statistics relevant for model evaluation. Raw data contain trends, so they are usually de-trended using filtering techniques before using them to generate the relevant statistics[7].

Model Evaluation in RBC models became much more sophisticated when the practice started to exploit the fact that a solved DSGE model is a VAR.

A solved DSGE could be represented a Structural VAR:

$$\mathbf{A}\begin{pmatrix} \mathbf{Y}_t \\ \mathbf{P}_t \end{pmatrix} = \mathbf{C}(L)\begin{pmatrix} \mathbf{Y}_{t-1} \\ \mathbf{P}_{t-1} \end{pmatrix} + \mathbf{B}\begin{pmatrix} \boldsymbol{\nu}_t^Y \\ \boldsymbol{\nu}_t^P \end{pmatrix} \tag{20}$$

within this framework a new role for empirical analysis emerges, that is to provide evidence on the stylized facts to include in the theoretical model adopted for policy analysis and to decide between competing DSGE models. The operationalization of this research program is very well described in a paper by Christiano, Eichenbaum and Evans (1998) considering the case of monetary policy. There are three relevant steps:

1. monetary policy shocks are identified in actual economies, i.e. in VAR without theoretical restrictions;
2. the response of relevant economic variables to monetary shocks is then described;
3. finally, the same experiment is performed in the model economies to compare actual and model-based responses as an evaluation tool and a selection criterion for theoretical models.

The identification of the shocks to interest is the first and most relevant step in VAR-based model evaluation. VAR modelling recognizes that identification and estimation of structural parameters is impossible without explicit modelling expectations, therefore a structure like (**??**) can only be used to run special experiments that do not involve simulating different scenarios for the parameters of interests. A natural way to achieve these results is to experiment with shocks $\boldsymbol{\nu}_t^P$. *Facts* are then provided by looking at impulse response analysis, variance decomposition and historical decompositions. Impulse response analysis describes the effect in time of a policy shocks on the variables of interest, variance decomposition illustrates how much of the variance of the forecasting errors for macroeconomic variables

[7] Some abuses of this practice are present in the literature, the most common one is to compare the properties of filtered raw data with those of filtered model generated data. Filtering model generated data is clearly hard to justify given that model generated data are stationary by their nature.

at different horizons can be attributed to policy shocks, historical decomposition allows to evaluate the effect of zeroing policy shocks on the variables to interest. All these experiments are run by keeping estimated parameters unaltered. Importantly, running these experiments makes sense only if shocks to the different variables included in the VAR are orthogonal to each other, otherwise it would not be possible to simulate a policy shock by maintaining all the other shocks to zero. As a consequence, VAR models need a structure because orthogonal shocks are normally not a feature of the statistical model. This fact generates an identification problem. In the reduced form we have:

$$\begin{pmatrix} \mathbf{Y}_t \\ \mathbf{P}_t \end{pmatrix} = \mathbf{A}^{-1}\mathbf{C}(L)\begin{pmatrix} \mathbf{Y}_{t-1} \\ \mathbf{P}_{t-1} \end{pmatrix} + \begin{pmatrix} \mathbf{u}_t^Y \\ \mathbf{u}_t^P \end{pmatrix},$$

where $\mathbf{u}$ denotes the VAR residual vector, normally independently distributed with full variance-covariance matrix $\mathbf{\Sigma}$. The relation between the residuals in $\mathbf{u}$ and the structural disturbances in $\boldsymbol{\nu}$ is therefore:

$$\mathbf{A}\begin{pmatrix} \mathbf{u}_t^Y \\ \mathbf{u}_t^P \end{pmatrix} = \mathbf{B}\begin{pmatrix} \boldsymbol{\nu}_t^Y \\ \boldsymbol{\nu}_t^P \end{pmatrix}. \tag{21}$$

Undoing the partitioning, we have

$$\mathbf{u}_t = \mathbf{A}^{-1}\mathbf{B}\boldsymbol{v}_t,$$

from which we can derive the relation between the variance-covariance matrices of $\mathbf{u}_t$ (observed) and $\boldsymbol{\nu}_t$ (unobserved) as follows:

$$E\left(\mathbf{u}_t\mathbf{u}_t'\right) = \mathbf{A}^{-1}\mathbf{B}E\left(\boldsymbol{v}_t\boldsymbol{v}_t'\right)\mathbf{B}'\mathbf{A}^{-1}.$$

Substituting population moments with sample moments we have:

$$\widehat{\sum} = \widehat{\mathbf{A}}^{-1}\widehat{\mathbf{B}}\mathbf{I}\widehat{\mathbf{B}}'\widehat{\mathbf{A}}^{-1}, \tag{22}$$

$\widehat{\sum}$ contains $n(n+1)/2$ different elements, that is the maximum number of identifiable parameters in matrices $\mathbf{A}$ and $\mathbf{B}$. Therefore, a necessary condition for identification is that the maximum number of parameters contained in the two matrices equals $n(n+1)/2$, such a condition makes the number of equations equal to the number of unknowns in system (22). As usual, for such a condition also to be sufficient for identification no equation in (22) should be a linear combination of the other equations in the system (see

Amisano and Giannini 1996, Hamilton 1994). As for traditional models, we have the three possible cases of under-identification, just-identification and over-identification. The validity of over-identifying restrictions can be tested via a statistic distributed as a χ^2 with a number of degrees of freedom equal to the number of over-identifying restrictions. Once identification has been achieved, the estimation problem is solved by applying generalized method of moments estimation.

Since VAR models are not estimated to yield advice on the best policy but rather to provide empirical evidence on the response of macroeconomic variables to policy impulses in order to discriminate between alternative theoretical models of the economy, it then becomes crucial to identify policy actions using restrictions independent from the theoretical models of the transmission mechanism under empirical investigation, taking into account the potential endogeneity of policy instruments. Restrictions based on the theoretical predictions of models are clearly inappropriate, so are Cowles commission type of restrictions as they do not acknowledge the endogeneity of systematic policy. The recent literature on the monetary transmission mechanism (see Strongin 1995, Bernanke and Mihov 1995, Christiano, Eichenbaum and Evans 1996, Leeper, Sims and Zha 1996), offers good examples on how these kind of restrictions can be derived. VARs of the monetary transmission mechanism are specified on six variables, with the vector of macroeconomic non-policy variables including gross domestic product (GDP), the consumer price index (P) and the commodity price level (Pcm), the vector of policy variables includes the federal funds rate (FF), the quantity of total bank reserves (TR) and the amount of non-borrowed reserves (NBR). Given the estimation of the reduced form VAR for the six macro and monetary variables, a structural model is identified by: (*i*) assuming orthogonality of the structural disturbances; (*ii*) imposing that macroeconomic variables do not simultaneously react to monetary variables, while the simultaneous feedback in the other direction is allowed, and (*iii*) imposing restrictions on the monetary block of the model reflecting the operational procedures implemented by the monetary policy-maker. All identifying restrictions satisfy the criterion of independence from specific theoretical models. In fact, within the class of models estimated on monthly data, restrictions (*ii*) are consistent with a wide spectrum of alternative theoretical structures and imply a minimal assumption on the lag of the impact of monetary policy actions on macroeconomic variables, whereas restrictions (*iii*) are based on institutional analysis. Restrictions (*ii*) are made operational by setting to zero an appropriate block of elements of the $\mathbf{A}$ matrix. Note that restrictions on the contemporaneous feedbacks among variables is not the only way of imposing restrictions consistent with a wide spectrum of theoretical models, in fact such aim could be achieved by imposing restrictions on the long-run effects of shocks (for example, there is a clear consensus among macroeconomist that demand shocks have zero effect on output in the long-run) or on the shape of some impulse response functions. These type of restrictions are easily imposed on SVAR (see, for example, Blanchard-Quah,1989 and Uhlig 1997), although one must be always aware of the effect of imposing invalid restrictions on parameter estimates (Faust and Leeper, 1997). Finally note that partial identification can be easily implemented in a VAR model. If the relevant dimension for model comparison is the response of the economy to monetary policy shocks there is no need to identify the non-monetary structural shocks in the model.

7 VAR Based Model Evaluation: an Assessment.

VAR based model evaluation can be assessed by discussing first the results achieved and their impact on model building, to then offer some consideration on the specification of the VAR and on the evaluation of the statistical model adopted.

The main results of the VAR based evaluation model is that, in order to match fluctuations in the data, any model must feature some attrition that causes temporary but rather persistent deviation from the long-run equilibrium defined by a frictionless neoclassical economy. In a series of recent papers, Christiano, Eichenbaum and Evans (1996*a*)–(1996*b*) apply the VAR approach to derive 'stylized facts' on the effect of a contractionary policy shock, and conclude that plausible models of the monetary transmission mechanism should be consistent at least with the following evidence on price, output and interest rates: *(i)* the aggregate price level initially responds very little; *(ii)* interest rates initially rise, and *(iii)* aggregate output initially falls, with a j-shaped response, with a zero long-run effect of the monetary impulse. Such evidence leads to the dismissal of traditional real business cycle models, which are not compatible with the liquidity effect of monetary policy on interest rates, and of the Lucas (1972) model of money, in which the effect of monetary policy on output depends on price misperceptions. The evidence seems to be more in line with alternative interpretations of the monetary transmission mechanism based on sticky prices models (Goodfriend and King 1997), limited participation models (Christiano and Eichenbaum 1992) or models with indeterminacy–sunspot equilibria (Farmer 1997). When model are extended to analyze the components of output some more frictions need to be added to explain the dynamics of consumption and investment, typically some habit persistence is needed to explain fluctuations in consumption and some adjustment costs are needed to match the dynamics of investment and the stock of capital in the data.

Specification of the VAR and its statistical adequacy is an issue that has not received much explicit attention in the literature. I think that, although it is well understood by now that identification of the VAR must be theory independent, much less reflection has been conducted on the specification of the VAR. It seems that the choice of the variables included in the VAR is driven by the theoretical model. This is natural: if the theoretical model is a restricted VAR, the natural benchmark is the same VAR without restrictions. But what about potential misspecification of the statistical model?

Statistical analysis of the unrestricted VAR is rather rare, although some implicit consideration has been clearly devoted to this issue. Think for example of the "liquidity puzzle" and the "price puzzle" for models of the money transmission mechanism.

VAR models of the monetary transmission mechanism were initially estimated on a rather limited set of variables, i.e. prices, money and output, and identified imposing a diagonal form on the matrix **B** and a lower triangular form on the matrix **A** with money coming last in the ordering of the variables included in the VAR (Choleski identification). The typical impulse responses obtained within this type of models show that prices slowly react to monetary policy, output responds in the short run, in the long run (from two years after the shock onwards) prices start adjusting and the significant effect on output vanishes. There is no strong evidence for the endogeneity of money. Macroeconomic variables play a very limited role in explaining the variance of the forecasting error of money, while money instead plays an important role in explaining fluctuations of both the macroeconomic variables.

Sims (1980) extended the VAR to include the interest rate on Federal funds ordered just before money as a penultimate variable in the Choleski identification. The idea is to see the robustness of the above results after identifying the part of money which is endogenous to the interest rate. Impulse response functions and FEVD raise a number of issues.

1. Though little of the variation in money is predictable from past output and prices, a considerable amount becomes predictable when past short-term interest rates are included in the information set.

2. It is difficult to interpret the behaviour of money as driven by money

supply shocks. The response to money innovations gives rise to the 'liquidity puzzle': the interest rate declines very slightly contemporaneously in response to a money shock to start increasing afterwards.

3. There are also difficulties with interpreting shocks to interest rates as monetary policy shocks. The response of prices to an innovation in interest rates gives rise to the 'price puzzle': prices increase significantly after an interest rate hike. An accepted interpretation of the liquidity puzzle relies on the argument that the money stock is dominated by demand rather than supply shocks. Moreover, the interpretation of money as demand shocks driven is consistent with the impulse response of money to interest rates. Note also that, even if the money stock were to be dominated by supply shocks, it would reflect both the behaviour of central banks and the banking system. For both these reasons the broad monetary aggregate has been substituted by narrower aggregates, bank reserves, on which it is easier to identify shocks mainly driven by the behaviour of the monetary policy maker. The 'price puzzle' has been attributed to mis-specification of the four-variables VAR used by Sims. Suppose that there exists a leading indicator for inflation to which the Fed reacts. If such a leading indicator is omitted from the VAR, then we have an omitted variable positively correlated with inflation and interest rates. Such omission makes the VAR mis-specified and explains the positive relation between prices and interest rates observed in the impulse response functions. It has been observed (see Christiano, Eichenbaum and Evans 1996, Sims 1996) that the inclusion of a Commodity Price Index in the VAR solves the 'price puzzle'.

As a result of these developments a consensus was reached on the specification of the VAR to provide facts on the MTM as a model including prices, output, a commodity price index, the policy rate and the narrow money indicators necessary to model the market for bank reserves.

Note that the final specification is very different from the initial one and the modification in the specification are driven by a number of puzzles found in the impulse responses of discarded VARs. One can of course interpret these puzzles as signals of misspecification of the VAR but it is not clear that puzzles are the best way to diagnose mis-specification of the statistical model. Think for example of the recent practice of identifying shocks by imposing constraint to shape of the impulse response functions. I think that it might be regarded as reasonable to impose that a monetary policy restriction has a non-positive effect on inflation. Obviously if VARs of the MTM would have been always identified by imposing this restriction the price puzzle would have never been observed and one is left to wonder if the consensus on the specification of the VAR to analyze the MTM would evolved differently from what it did.

Another issue of crucial importance is structural stability of the parameters estimated in the VAR. If the VAR is a reduced form of a forward-looking model it is of crucial importance to estimate its parameters on a single regime. Although this issue has been explicitly recognized in some papers, for example Bernanke-Mihov(1998), the consensus VAR is normally estimated on a sample including different monetary regimes. The main justification for this practice is that monetary policy shocks are robust to the different identification generated by the different monetary policy regime. Some authors have been left skeptical by such robustness and some criticisms has been moved to VAR based monetary policy shocks. Rudebusch, 1998, argues that VAR based monetary shocks do not make sense as they are very little correlated with monetary policy shocks directly derived from asset prices (the federal fund future). The mainstream reaction to this criticism is that even if the two type of shocks are very little correlated, the impulses responses of macroeconomic variables VAR based and financial market based monetary policy shocks are not significantly different from each other. Rudebusch criticism shared the same fortune with other criticism of the VAR approach. Lippi and Reichlin(1993) pointed out that a crucial assumption in structural VAR modelling is that structural shocks are linear combinations of the residuals in reduced form VAR models to argue that modern macroeconomic models which are linearized into dynamic systems tend to include non-invertible moving average components and structural shocks are therefore not identifiable. In fact, the linearized modern macroeconomic models of the monetary transmission mechanism deliver short VARs. In such models structural shocks are combinations of the residuals in the reduced form VARs (the Wold innovations) and the Lippi–Reichlin critique does not seem to be applicable (for a further discussion of this point see Amisano and Giannini 1996).

To sum up, although the original idea of the Cowles Commission to use the implied unrestricted reduced form as a benchmark to evaluate the structural model is clearly reflected in the VAR based evaluation of DSGE models, the potential importance of the formal evaluation of the adequacy of the statistical model adopted has not certainly received the same attention. However, in the practice of VAR specification some attention to issue of potential misspecification has been clearly paid, such attention has been clearly more related to the economic interpretation of results than to the implementation of formal statistical criteria for model evaluation.

8 Model Evaluation in Bayesian Analysis of DSGE models

VAR based evaluation of early DSGE model made clear that a large number of nominal and real frictions should be added to the traditional new-classical RBC models to replicate relevant features in observed data (see, for example, Christiano, Eichenbaum and Evans, 2005). Adding frictions implies increasing the number of parameters, especially along the dimension of parameters little related to theory. As a consequence calibration became impractical to attribute numerical values to the DSGE parameters and estimation came back into fashion. However, estimating DSGE models by classical maximum likelihood methods proved to be very hard as the convergence of the estimates to values that ensure a unique stable solution[8] turned out to be practically impossible to achieve when implementing unconstrained maximum likelihood estimation. A paper by Ireland(2004) made an exception and obtained convergence of numerical estimates of parameters of a DSGE model to values that allow economic policy simulation. In practice the Ireland method consists of penalizing the likelihood function along some dimension so that the range of variation of many parameters is limited (for an interesting discussion of the estimation implemented in Ireland, see Johansen(2004)).

In practice, one can think of Ireland method as a naive Bayesian one in which some from of (very tight) prior is imposed on (at least a subset of) the parameters. A natural development of Ireland proposal was to extend the naive Bayesian framework to a proper Bayesian framework. This is what happened as soon as the use of MCMC methods to derive the relevant posterior distribution of parameters became widespread (see An and Schorfeide, 2005, for a survey and Del Negro and Smets and Wouters, 2003, and Ruge-Murcia(2003) for applications).

Once adopted, the Bayesian framework offered naturally some additional model evaluation tools. These tools were generated by pairing the tradition of model evaluation in the Bayesian approach to macroeconometrics with the

[8] Three types of solution are possible for a DSGE model, depending on its parameterization: no stable rational expectations solution exists, the stable solution is unique (determinacy), or there are multiple stable solutions(indeterminacy). Determinacy is a prerequisite in order to use a model to simulate the effects of economic policy.

VAR nature of a solved DSGE model. The Bayesian approach made its way into applied macroeconometrics to solve the problem of the lack of parsimony of VARs. In practice, data availability from a single regime poses a binding constraint on the number of variables and the number of lags that can be included in a VAR without overfitting the data. A solution of the problem of over-parameterization is constraining the parameters by shrinking them toward some specific point in the parameter space. The Minnesota prior, proposed by Doan, Litterman and Sims(1984) uses the Bayesian approach to shrink the estimates toward the univariate random walk representation for all variables included in the VAR. Within this framework, Bayesian methods are used to save degrees of freedom on the basis of the well established statistical evidence that no-change forecasts are known to be very hard to beat for many macroeconomic variables. De Jong, Ingram and Whiteman(1996, 2000) and Ingram and Whiteman(1994) proposed to evaluate RBC models by comparing the forecasting performance of a Bayesian VAR estimated via the Minnesota prior with that of a VAR in which the a-theoretical prior information in the Minnesota prior was supplanted by the information in a RBC model.

In a series of papers Del Negro and Schorfeide (2004, and 2006) and Del Negro, Schorfeide, Smets and Wouters(2004) use this approach to model evaluation with the fact that a solved DSGE model is a restricted VAR to develop a method that tilt coefficient estimates of an unrestricted VAR toward the restriction implied by a DSGE model. The weight placed on the DSGE model is controlled by an hyperparameter called λ. This parameter takes values ranging from 0 (no-weight on the DSGE model) to ∞ (no weight on the unrestricted VAR). Therefore, the posterior distribution of λ provides an overall assessment of the validity of the DSGE model restrictions. To see how the approach is implemented, consider that the solved DSGE model generates a restricted MA representation for the vector of n variables of interest $\mathbf{Z}_t = \begin{pmatrix} \mathbf{Y}_t & \mathbf{P}_t \end{pmatrix}$,that can be approximated by a VAR of order p:

$$
\begin{aligned}
\mathbf{Z}_t &= \mathbf{\Phi}_0^*(\theta) + \mathbf{\Phi}_1^*(\theta)\,\mathbf{Z}_{t-1} + ... + \mathbf{\Phi}_p^*(\theta)\,\mathbf{Z}_{t-p} + \mathbf{u}_t^* \\
\mathbf{u}_t^* &\sim N\left(\mathbf{0}, \Sigma_u^*(\theta)\right) \\
\mathbf{Z}_t' &= \mathbf{X}_t'\mathbf{\Phi}^*(\theta) + \mathbf{u}_t', \\
\mathbf{X}_t &= \left[1, \mathbf{Z}_{t-1}', ..., \mathbf{Z}_{t-p}'\right] \\
\mathbf{v} &= \left[\mathbf{\Phi}_0^*(\theta), \mathbf{\Phi}_1^*(\theta), ..., \mathbf{\Phi}_p^*(\theta)\right]'
\end{aligned}
$$

where all coefficients are convolutions of the structural parameters in the model included in the vector θ. The chosen benchmark to evaluate this model is the unrestricted VAR derived from the solved DSGE model

$$
\begin{aligned}
\mathbf{Z}_t' &= \mathbf{X}_t'\mathbf{\Phi} + \mathbf{u}_t', \\
\mathbf{X}_t &= \left[1, \mathbf{Z}_{t-1}', ..., \mathbf{Z}_{t-p}'\right] \\
\mathbf{\Phi} &= \left[\mathbf{\Phi}_0, \mathbf{\Phi}_1, ..., \mathbf{\Phi}_p\right]
\end{aligned}
$$

where:

$$
\begin{aligned}
\mathbf{\Phi} &= \mathbf{\Phi}^*(\theta) + \mathbf{\Phi}^\Delta \\
\Sigma_u &= \Sigma_u^*(\theta) + \Sigma_u^\Delta
\end{aligned}
$$

the DSGE restrictions are imposed on the VAR by defining:

$$
\begin{aligned}
\Gamma_{XX}(\theta) &= E_\theta^D\left[\mathbf{X}_t\mathbf{X}_t'\right] \\
\Gamma_{XZ}(\theta) &= E_\theta^D\left[\mathbf{X}_t\mathbf{Z}_t'\right]
\end{aligned}
$$

where E_θ^D defines the expectation with respect to the distribution generated by the DSGE model, that of course have to be well defined. We then have:

$$
\mathbf{\Phi}^*(\theta) = \Gamma_{XX}(\theta)^{-1}\Gamma_{XZ}(\theta)
$$

Beliefs about the DSGE model parameters θ and model misspecification matrices $\mathbf{\Phi}^\Delta$ and Σ_u^Δ are summarized in prior distributions, that, as shown in Del Negro et al.(2004) can be transformed into prior for the VAR parameters $\mathbf{\Phi}$ and Σ_u.In particular we have:

$$
\begin{aligned}
\Sigma_u\,|\theta &\sim IW\left(\lambda T\Sigma_u^*(\theta), \lambda T - k, n\right) \\
\mathbf{\Phi}\,|\Sigma_u, \theta &\sim N\left(\mathbf{\Phi}^*(\theta), \frac{1}{\lambda T}\left[\Sigma_u^{-1}\otimes\Gamma_{XX}(\theta)\right]^{-1}\right)
\end{aligned}
$$

where the parameter λ controls the degree of model misspecification with respect to the VAR: for small values of λ the discrepancy between the VAR and the DSGE-VAR is large and a sizeable distance is generated between unrestricted VAR and DSGE estimators, large values of λ correspond to small model misspecification and for $\lambda = \infty$ beliefs about DSGE mis-specification degenerate to a point mass at zero. Bayesian estimation could be interpreted as estimation based a sample in which data are augmented by an hypothetical sample in which observations are generated by the DSGE model, within this framework λ determines the length of the hypothetical sample.

Given the prior distribution, posterior are derived by the Bayes theorem:

$$
\begin{aligned}
\Sigma_u \left| \theta, Z \right. &\sim IW\left(\left(\lambda+1\right) T \hat{\Sigma}_{u,b}\left(\theta\right), \left(\lambda+1\right) T - k, n \right) \\
\mathbf{\Phi} \left| \Sigma_u, \theta, Z \right. &\sim N\left(\hat{\mathbf{\Phi}}_b\left(\theta\right), \Sigma_u \otimes \left[\lambda T \Gamma_{XX}\left(\theta\right) + \mathbf{X}'\mathbf{X}\right]^{-1} \right) \\
\hat{\mathbf{\Phi}}_b\left(\theta\right) &= \left(\lambda T \Gamma_{XX}\left(\theta\right) + \mathbf{X}'\mathbf{X}\right)^{-1} \left(\lambda T \Gamma_{XZ}\left(\theta\right) + \mathbf{X}'\mathbf{Z}\right) \\
\hat{\Sigma}_{u,b}\left(\theta\right) &= \frac{1}{\left(\lambda+1\right) T} \left[\left(\lambda T \Gamma_{ZZ}\left(\theta\right) + \mathbf{Z}'\mathbf{Z}\right) - \left(\lambda T \Gamma_{XZ}\left(\theta\right) + \mathbf{X}'\mathbf{Z}\right) \hat{\mathbf{\Phi}}_b\left(\theta\right) \right]
\end{aligned}
$$

which shows that the smaller λ, the closer the estimates are to the OLS estimates of an unrestricted VAR, the higher λ the closer the estimates are to the values implied by the DSGE model parameters θ.

In practice, a grid search is conducted on a range of values for λ to choose that value that maximize the marginal data density. The typical results obtained when using DSGE-VECM(λ) to evaluate models with frictions is that " ... the degree of misspecification in large-scale DSGE models is no longer so large as to prevent their use in day-to-day policy analysis, yet is not small enough that it cannot be ignored...".

8.1 DSGE-VAR based model evaluation: an assessment

DSGE-VAR model evaluation takes the Lucas and Sims critique very seriously but ignores the issue of specification of the statistical model. The VAR used as a benchmark is the solved DSGE model that is generalized only by relaxing restrictions on parameters. The validity of the statistical model underlying the empirical specification is never questioned. Although the models are different, the evaluation strategy in the DSGE-VAR approach is very similar to the approach of evaluating models by testing over-identifying restrictions without assessing the statistical model implemented in Cowles foundation models. In fact, the DSGE-VAR approach is looser than the Cowles foundation approach in that model based restrictions are not imposed and tested but a different question is asked: restrictions are made fuzzy by imposing a distribution on them and then the relevant question becomes what is the amount of uncertainty that we have to add to model based restrictions in order to make them compatible not with the data but with a model-derived unrestricted VAR representation of the data. The natural question here is how well does this procedure do in rejecting false models? Spanos(1991) has shown clearly that modification in the structure of the statistical model could lead to dramatic changes in the outcome of tests for over-identifying restrictions. Why is this worry so strongly de-emphasized in the DSGE-VAR literature?

What are the potential sources of model derived VAR specification? An obvious candidate are all those variables that are related to the mis-specification of the theoretical model, but there are also all those variables that are not theory related but are important to model the actual behaviour of policy makers. Think for example of the commodity price index and the modelling of the behaviour of monetary policy authority. We have discussed in one of the previous section how the inclusion of this variable in a VAR to identify monetary policy shocks has been deemed important to model correctly the information set of the monetary policy maker when forecasting inflation and then to fix the "price- puzzle". DSGE model do not typically include the commodity price index in their specification as a consequence the VAR derived by relaxing the theoretical restrictions in a DSGE model is misspecified. So the evaluation of the effects of conducting model misspecification with a "wrong" benchmark is a practically relevant one.

As a matter of fact DSGE model tend to produce a high number of very persistent shocks (see Smets and Wouters, 2003), this would have been certainly taken as a signal of model mis-specification by an LSE type methodology. Still the model do not do too badly when judged in the metric of the λ test. It would be important to have some evaluation of phenomena like this.

Another dimension potentially relevant for evaluating the statistical model underlying VAR-DSGE is structural stability of the VAR parameters. If the DSGE restrictions are valid, then parameters in the VAR are convolutions of structural parameters that, by their nature, should be constant over time. It is well known that tests for structural stability have problems of power, especially in presence of multiple breaks at unknown dates. Detecting structural breaks in parameters of interest becomes even harder when structural innovations in the DSGE are allowed to have volatilities that vary over time. Justiniano and Primiceri(2005) have extended the Bayesian framework to develop an algorithm for inferring DSGE model parameters and time varying volatilities of structural shocks. Allowing for time varying volatilites makes the DSGE model consistent with structural breaks while keeping the deep parameters constant. However, it is hard to distinguish empirically the case for genuine stochastic volatility against a situation in which allowing for stochastic volatility in the estimation picks up parameters instability in a VAR model with constant volatility of structural shocks.

There are alternatives to the use of a VAR as a benchmark. The limited information problem in VAR could be solved by combining traditional VAR analysis with recent developments in factor analysis for large data sets and in using a factor-augmented VAR (FAVAR) as the relevant statistical model to conduct model evaluation. A recent strand of the econometric literature[9] has shown that very large macroeconomic datasets can be properly modelled using dynamic factor models, where the factors can be considered as an exhaustive summary of the information in the data. This approach has been successfully employed to forecast macroeconomic time series and in particular inflation. As a natural extension of the forecasting literature, Bernanke and Boivin (2003), Bernanke, Boivin and Eliasz(2005) proposed to exploit these factors in the estimation of VAR. A FAVAR benchmark for the evaluation of a DSGE model will take the following specification:

$$\begin{pmatrix} \mathbf{Z}_t \\ \mathbf{F}_t \end{pmatrix} = \begin{bmatrix} \mathbf{\Phi}_{11}(L) & \mathbf{\Phi}_{12}(L) \\ \mathbf{\Phi}_{21}(L) & \mathbf{\Phi}_{22}(L) \end{bmatrix} \begin{pmatrix} \mathbf{Z}_{t-1} \\ \mathbf{F}_{t-1} \end{pmatrix} + \begin{pmatrix} \mathbf{u}_t^Z \\ \mathbf{u}_t^F \end{pmatrix},$$

where $\mathbf{Z}_t$ are the variables included in the DSGE model and $\mathbf{F}_t$ is a small vector of unobserved factors extracted from a large data-set of macroeconomic time series, that capture additional economic information relevant to model the dynamics of $\mathbf{Z}_t$. The system reduces to the standard VAR used to evaluate DSGE models if $\mathbf{\Phi}_{12}(L) = 0$, therefore, within this context, the relevant λ test would add to the usual DSGE model-related restrictions on $\mathbf{\Phi}_{11}(L)$ the restrictions $\mathbf{\Phi}_{12}(L) = 0$.

Consolo et al.(2007) apply this idea to find that FAVAR models dominate VAR specification generated by adopting unrestricted version of the solution of DSGE models. Such dominance is clearly established by analysis of residuals and evaluation of forecasting performance. However, when the Bayesian approach is applied to the DSGE-FAVAR instead of the DSGE-VAR some support for the DSGE model is still found in the data (the optimal λ in the DSGE-FAVAR is different from zero). Moreover, the optimal combination between the DSGE model and the statistical model based on a larger information set (the FAVAR) delivers a forecasting model (the DSGE-FAVAR) that dominates all alternatives. This evidence leads to a

[9] Stock and Watson (2002), Forni and Reichlin (1996, 1998) and Forni et al. (1999, 2000)

new interaction between theory and empirical analysis where the theoretical DSGE model should not be considered as a model for the data but as a generator of prior distribution for the empirical model. The use of the FAVAR as an empirical model allows to include in the analysis also the information that is not considered in the theoretical model.

Beside this application there has been no work using FAVAR to evaluate DSGE, interestingly what has instead happened is that FAVAR have been interpreted as the reduced form of a DSGE model. This result has been achieved by removing the assumption that economic variables included in a DSGE are properly measured by a single indicator and by treating theoretical concepts of the model as partially observed to use the information set in factors to map them (Boivin and Giannoni,2005). This approach makes a FAVAR the reduced form a DSGE model, although the restrictions implied by DSGE model on a general FAVAR are very difficult to trace and model evaluation becomes even more difficult to implement. In fact, a very tightly parameterized theory model can have a very highly parameterized reduced form if one is prepared to accept that the relevant theoretical concept in the model are combination of many macroeconomic and financial variables. Identification of the relevant structural parameters, that is very hard also in DSGE model with observed variables (see Canova and Sala,2006), becomes even harder. Natural advantages of this approach are increased efficiency in the estimation of the model and improved forecasting performance. However, model evaluation becomes almost impossible to pursue and a theoretical model can only by rejected by another theoretical model, while the implied statistical model is made so general that virtually no room is left to the data to reject a DSGE model.

9 Conclusions

In this paper we have analyzed developments in the evaluation of macroeconomic models designed for policy simulation analysis. We started from the failure of the large empirical macroeconomic models in the seventies. We analyzed two different diagnoses for this failure: the first one mainly related to the unsatisfactory theoretical background of the models and heralded by the Lucas and Sims critique, the second one mainly related to the unsatisfactory statistical background of the Cowles Commission models. We have then illustrated how the different explanation for the failure have generated different streams of literature. The diagnosis based on the statistical identification problem has generated empirical approaches concentrating almost exclusively on the dynamic specification of the statistical model by de-emphasizing the importance of explicit microeconomic foundations of the model used for simulating macroeconomic policy. The diagnosis based on the structural identification problem has generated the DSGE approach to macroeconomic modelling in which the main emphasis has been clearly posed on the identification of the structural parameters of interest via microeconomic foundations of the adopted model.

The realization that the solution of DSGE model can be approximated by a restricted VAR, which is also a statistical model, has generated a potential link between the two approaches. This link has been not fully exploited so far and little research has been devoted to the evaluation of the statistical adequacy of the VAR used as benchmark to evaluted the validity of the DSGE model.

In fact there are two interesting avenue for pursuing this type of research.

A first approach looks at theoretical DSGE models as the natural way to generate prior distribution for the empirical model, which should be an (optimal) combination of a tightly parameterized theoretical model and of a more general empirical model, possibly a FAVAR. This approach requires the application of Bayesian methods. A second approach looks at theory as informative only on the long-run relations between economic variables, so theory should be used to specify a cointegrated VAR in which the short-run dynamics is determined by the data but the long-run properties of the model depend on testable (and tested) theoretical assumption.

Importantly both these approaches recognize the importance of both the theoretical and the statistical model, although the relative weights can be very different.

10 References

Amisano, G., and Giannini, C. (1996). *Topics in Structural VAR Econometrics.* Springer–Verlag.

An, Sungbae, and Frank Schorfheide (2006): "Bayesian Analysis of DSGE Models", *Working Paper, 06-5, Federal Reserve Bank of Atlanta.*

Anderson, Gary (1997): "Continuous Time Application of the Anderson-Moore (AIM) Algorithm for Imposing the Saddle Point Property in Dynamic Models," *Unpublished manuscript, Board of Governors of the Federal Reserve System.*

Anderson, T.W. and H. Rubin (1949) "Estimation of the parameters of a single equation in a complete system of stochastic equations" Annals of

Mathematical Statistics 20, 46—63.

Arouba, S.Boragan, Jesus Fernandez-Villaverde and Juan F. Rubio-Ramirez (2003): "Comparing Solution Methods for Dynamic Equilibrium Economies," *PIER Working Paper 04-003*.

Baba, Y., D.F. Hendry, and R.M. Starr (1992). The demand for M1 in the U.S.A., 1960-1988. Review of Economic Studies, 59, 25-61.

Basmann, R.L. (1960). "On finite sample distributions of generalized classical linear identifiability test statistic." Journal of the American Statistical Association 55. 650—659.

Bernanke, Ben S., and Jean Boivin (2003): "Monetary Policy in a Data-Rich Environment," *Journal of Monetary Economics*, L, 52-5464.

Bernanke, Ben S., Jean Boivin and Piotr Eliasz (2005): " Measuring the Effects of Monetary Policy a Factor-Augmented Vector Autoregressive (FAVAR) Approach," *The Quarterly Journal of Economics*, MIT Press, vol. 120(1), pages 387-422, January.

Bernanke, B. S. and Mihov, I. (1998). 'Measuring monetary policy'. *Quarterly Journal of Economics,* 113, 3: 869-902.

Blanchard, O. J., and Quah, D. T. (1989). 'The dynamic effects of aggregate demand and supply disturbances'. *American Economic Review*, 79: 655-73.

Blanchard, Oliver, and Charles M. Kahn (1980): " The Solution of Linear Difference Models under Rational Expectation," *Econometrica*, 48, 1305-1313.

Boivin, Jean, and Marc P. Giannoni (2005):"DSGE Models in a Data-Rich Environment," *Working Paper*.

Canova, Fabio (1994): "Statistical Inference in Calibrated Models," *Journal of Applied Econometrics*, 9, S123-144.

Canova, Fabio, (2007): Methods for Applied Macroeconomic Research, *Princeton University Press.*

Canova Fabio,(2007) "How much structure in empirical models", Chapter in the Palgrave Handbook of Applied Econometrics, vol 2..

Canova, Fabio and Gianni De Nicolò (2002): "Monetary Disturbances Matter for Business Cycle Fluctuations in the G-7," *Journal of Monetary Economics.*

Canova, F. and L. Sala (2005) "Back to Square one: identification issues in DSGE models" , IGIER Working Paper 303, Università Bocconi

Clarida, Richard, Jordi Galì, and Mark Gertler (2000): "Monetary Policy Rules and Macroeconomics Stability: Evidence and some Theory," *Quarterly Journal of Economics*, 115, 147-180.

Christiano, L. J., and Eichenbaum, M. (1992). ‘Liquidity Effects and the Monetary Transmission Mechanism’. *American Economic Review*, 82, 2: 346-353.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1996a). ‘The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds’. *Review of Economics and Statistics*, 78: 16-34.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1996b). ‘Monetary Policy Shocks and their Consequences: Theory and Evidence’. Paper presented at ISOM.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1998). ‘Monetary Policy shocks: what have we learned and to what end?’. NBER working paper No. 6400.

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans (2005): "Nominal Rigidities and the Dynamic effects of a Shcok to Monetary Policy," *Journal of Political Economy*, 113, 1-45.

Cogley, Timothy, and James Nason (1994): "Testing the Implications of Long-Run Neutrality for Monetary Business Cycle Models," *Journal of Applied Econometrics*, 9, S37-70.

Consolo A., C.A. Favero and A.Paccagnini(2007) "On the Statistical Identification of DSGE models" IGIER Discussion Paper

Cooley, T. F. (1995). *Frontiers of Business Cycle Research.* Princeton: Princeton University Press.

Cooley, T. F. (1997). ‘Calibrated Models’. *Oxford Review of Economic Policy*, 13: 55–69.

Doan, Thomas, Robert Litterman and Christopher Sims (1984): "Forecasting and Conditional Projections Using Realistic Prior Distributions," *Econometric Reviews*, 3, 1-100.

DeJong, David, Beth Ingram, and Charles Whiteman (1996):"A Bayesian Approach to Calibration," *Journal of Business economics and Statistics*, 14, 1-9.

DeJong, David, Beth Ingram, and Charles Whiteman (2000): "A Bayesian Approach to Dynamic Macroeconomics," *Journal of Econometrics*, 98, 203-223.

Del Negro, Marco, and Frank Schorfheide (2004): " Priors from General equilibrium Models for VARs," *International Economic Review*, 45, 643-673.

Del Negro, Marco, and Frank Schorfheide (2006): "How Good is what You’ve Got? DSGE-VAR as a Toolkit for evaluating DSGE Models," *Federal Reserve Bank of Atlanta Economic Review*

Del Negro, Marco, Frank Schorfheide, Frank Smets and Raf Wouters (2006): "On the Fit of New-Keynesian Models," *Working Paper.*

Engle R., Hendry D.F. and J.F. Richard(1983) "Exogeneity", Econometrica, 51, 277-302

Engle R. and D.F. Hendry(1993) "Testing superexogeneity and invariance in regression models", Journal of Econometrics, 56:119-139

Faust, Jon (1998): "The Robustness of Identified VAR Conclusions about Money," *Carnegie Rochester Conference Series* 49 , 207-44.

Faust, J., and Leeper, E. M. (1997). 'When do long run identifying restrictions give reliable results?'. *Journal of Business and Economic Statistics*, 15, 3: 345-353.

Favero C.A.(2001) *Applied Macroeconometrics*, Oxford University Press, Oxford

Fernandez-Villaverde, Jesus, and Juan Francisco Rubio-Ramirez (2004): " Comparing Dynamic Equilibrium Models to Data: A Bayesian Approach," *Journal of Econometrics*, 123, no. 1:153-87.

Forni, M. and L. Reichlin (1996), "Dynamic common factors in large cross-sections", *Empirical Economics*, 21, 27-42.

Forni, M. and L. Reichlin (1998), "Let's get real: A dynamic factor analytical approach to disaggregated business cycle", *Review of Economic Studies*, 65, 453-474.

Forni, M., Hallin, M., Lippi, M. and L. Reichlin (2000), "The generalized factor model: identification and estimation", *The Review of Economic and Statistics*.

Gallaway L.E., P.Smith(1961) "A Quarterly Econometric Model of the United States", Journal of the American Statistical Association, 56, 294, 379-383.

Garratt, A., Lee K., M.H.Pesaran and Y.Shin (2006) "Global and National Macroeconometric Modelling. A Long-run structural approach.", Oxford University Press, Oxford.

Geweke, John (1999): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communications," *Econometrics Reviews*, 18, 1-127.

Goodfriend, M., and King, R. (1997). 'The Newclassical Synthesis and the Role of Monetary Policy'. Paper presented at the 12th NBER Annual Macroeconomic Conference.

Hendry, D.F. (1976). The structure of simultaneous equations estimators. Journal of Econometrics, 4, 51-88

Hendry, D.F.(1988) "The econcompassing implications of feedback versus feedforward mechanisms in econometrics", Oxford Economic Papers,4:133-147.

Hendry, D.F. (1995). *Dynamic Econometrics.* Oxford: Oxford University Press.

Hendry, D.F., J.N.J. Muellbauer, and T.A. Murphy (1990). The Econometrics of DHSY. In J.D. Hey and D. Winch (Eds.), A Century of Economics, pp. 298-334. Oxford: Basil Blackwell.

Ingram, Beth, and Charles Whiteman(1994): "Supplanting the Minnesota Prior - Forecasting Macroeconomics Time Series using Real Business Cycle Model Priors," *Journal of Monetary Economics,* 34, 497-510.

Ireland, Peter N. (2004): "A Method for Taking Models to the Data," *Journal of Economic Dynamics and Control,* 28, 1205-1226.

Johansen, S. (1995). *Likelihood Based Inference on Cointegration in the Vector Autoregressive Model.* Oxford: Oxford University Press.

Johansen, S. (2004) "How not to take a model to the data", mimeo

Judd , Kenneth L. (1992): "Projection Methods for Solving Aggregate Growth Models,", *Journal of Economic Theory*, 58, 410-452.

Judd , Kenneth L., and Sy-Ming Guu (1997): "Asymptotic methods for aggregate growth models," *Journal of Economic Dynamics and Control*, Elsevier, vol. 21(6), pages 1025-1042, June.

Juselius K, and S. Johansen(1999) "Macroeconomic transmission mechanism. Empirical applications and econometric methods." paper available at http://www.econ.ku.dk/okokj/

Juselius K., and M.Franchi(2007) "Taking a DSGE Model to the Data Meaningfully", Economics Discussion Papers,
www.economics-ejournal.org/economics/journalarticles/2007-4

King, Robert G., Charles I. Plosser and Sergio T. Rebelo (1988): "Production, Growth and Business Cycles," *Journal of Monetary Economics*, 21, 195-232.

Klein, L.R. and A.S.Golberger(1955) An econometric Model of the United States, 1929-1952, Amsterdam, North Holland Publishing Company

Kydland Finn, and Edward Prescott (1982): " Time to Build and Aggregate Fluctuations," *Econometrica*, 50, 1345-70.

Kydland, F., and Prescott, P. (1996). 'The Computational Experiment: an Econometric Tool'. *Journal of Economic Perspectives*, 10: 69-85.

Koop, Gary (2003): "Bayesian Econometrics," *John Wiley & Sons.*

Justiniano A. and G.E. Primiceri(2005) "The time-varying volatility of macroeconomic fluctuations", mimeo Princeton University

Leeper, E. M., Sims, C. A., and Zha, T. (1996). 'What does monetary policy do?'. Available at http://eco-072399b.princeton.edu/yftp/bpea/bpeaf.pdf

Lippi, M., and Reichlin, L. (1993). 'The dynamic effects of aggregate demand and supply disturbances: comment'. *American Economic Review*, 83, 3: 644-652.

Lubik, Thomas, and Frank Schorfheide (2003): "Computing Sunspot Equilibria in Linear Rational Expectations Models," *Journal of Economic Dynamics and Control*, 28(2), 2003, 273-285.

Lubik, Thomas, and Frank Schorfheide (2004):"Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 94(1), 2004, 190-217.

Lucas, R. E. Jr. (1972). 'Expectations and the Neutrality of Money'. *Journal of Economic Theory*, 4, April: 103-124.

Lucas, R. E. Jr. (1976). 'Econometric Policy Evaluation: A Critique'. In K. Brunner and A. Meltzer (eds.) *The Phillips curve and labor markets.* Amsterdam: North-Holland.

Maddala (1988) *Introduction to Econometrics*, Mac Millan, New York, NY

McGrattan, Ellen R. (1999): "Application of Weigthed Residuals Methods to Dynamic Economic Models," in R. Marimom and A. Scott (eds) *Computational Methods for the Study of Dynamic Economies. Oxford University Press.*

Pesaran M.H and Y.Shin (1998) "Generalized Impulse Response Analysis in Linear Multivariate Models" Economics Letters, 58, 17-29

Pesaran M.H and Y.Shin (2002) "Long-run structural modelling", Econometric Reviews, 21, 49-87

Qin D.(1993). The formation of econommetrics. An historical perspective. Oxford, Oxford University Press

Rotemberg, Julio, and Micheal Woodford (1997): "An Optimization-Based Econometric framework for the Evaluation of Monetary Policy," *NBER Macroeconomics annual*, 12, 297-246.

Rudebusch, G. D. (1998). 'Do Measures of Monetary Policy in a VAR Make Sense?'. *International Economic Review*, 39: 907-931.

Ruge-Murcia, Francisco J. (2003): "Methods to estimate Dynamic Stochastic General Equilibrium Models," *CIREQ, Cahier 17-2003*.

Schmitt-Grohé, Stephanie, and Martin Uribe (2002): "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function," *NBER Technical Working Paper* 282.

Schorfheide, Frank (2000): "Loss Function-Based evaluation of DSGE Models," *Journal of Applied Econometrics*, 15, S645-670.

Sims, C. A. (1980). 'Macroeconomics and Reality'. *Econometrica*, 48: 1-48.

Sims, Christopher (1992): "Interpreting the Macroeconomic Time Series facts: The Effects of Monetary Policy," *European Economic Review*, XXXVI, 975-1000.

Sims, Chistopher (2002): "Solving Linear rational expectations Models," *Computational Economics*, 20 (1-2), 1-20.

Sims, Christopher and Tao Zha (1998): "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review*, 39, 949-968.

Sims, C. A. (1980b). 'Comparison of interwar and postwar business cycles: monetarism reconsidered'. *American Economic Review*, May: 250-257.

Sims, C. A., Stock, J. H., and Watson, M. (1990). 'Inference in linear time-series models with some unit roots'. *Econometrica*, 58: 113-144.

Sims, C. A., and Zha, T. (1996). 'Does Monetary Policy Generate Recessions?'. Mimeo. Available at ftp://ftp.econ.yale.edu/pub/sims/mpolicy

Smets, Frank, and Raf Wouters (2003): "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1, 1123-75.

Smith, Anthony (1993): "Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions," *Journal of Applied Econometrics*, 8, S63-S84.

Stock, James, and Mark Watson (2002): "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business Economics and Statistics*, XX:II, 147-162.

Tauchen, George (1986): "Finite State Markov-chain approximations to Univariate and Vector Autoregressions," *Economics Letters* 20, 177-181.

Theil, Henry, and Arthur S. Goldberg (1961): "On Pure and Mixed Estimation in Economics," *International Economic Review*, 2, 65-78.

Uhlig, Harald (1999): "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models easily," in R. Marimom and A. Scott (eds) *Computational Methods for the Study of Dynamic Economies. Oxford University Press.*

Uhlig, Harald (2001): "What are the effects of Monetary Policy? Evidence from an Agnostic Identification Procedure," *Mimeo, Humboldt University.*