

How much structure in empirical models?

Fabio Canova *
ICREA-UPF, CREI, AMeN and CEPR

First Draft: October 2007

Abstract

This chapter highlights the problems that structural methods and SVAR approaches have when estimating DSGE models and examining their ability to capture important features of the data. We show that structural methods are subject to severe identification problems due, in large part, to the nature of DSGE models. The problems can be patched up in a number of ways but solved only if DSGEs are completely reparametrized or respecified. The potential misspecification of the structural relationships give Bayesian methods an hedge over classical ones in structural estimation. SVAR approaches may face invertibility problems but simple diagnostics can help to detect and remedy these problems. A pragmatic empirical approach ought to use the flexibility of SVARs against potential misspecification of the structural relationships but must firmly tie SVARs to the class of DSGE models which could have generated the data.

JEL classification numbers: C10, C52, E32, E50

Key words: DSGE models, SVAR models, Identification, Invertibility, Misspecification, Small Samples.

*Conversations with L. Sala and C. Michelacci are gratefully acknowledged. Financial support from the CREI and the Spanish Ministry of Education through the grant SEJ-2004-21682-E. is gratefully acknowledged. This paper was prepared for the Palgrave Handbook of Economics, volume 2, Applied Econometrics, edited by T. Mills and K. Patterson.

1 Introduction

The 1990's have witnessed a remarkable development in the specification of stochastic general equilibrium models. The literature has added considerable realism to the popular workhorses of the 1980's; a number of shocks and frictions have been introduced into first generation RBC models driven by a single technological disturbance; and our understanding of the propagation mechanism of structural shocks has been considerably enhanced. Steps forward have also been made in comparing the quality of the models' approximation to the data. While a few years ago it was standard to calibrate the parameters of a model and informally evaluate the quality of its fit to the data, now full information likelihood-based estimation of the structural parameters has become common practice (see e.g. Smets and Wouters (2003), Ireland (2004), Canova (2004), Rubio and Rabanal (2005), Gali and Rabanal (2005)) and new techniques have been introduced for model evaluation purposes (see Del Negro et. al. (2006)). Given the complexities involved in estimating stochastic general equilibrium models and the difficulties in designing criteria which are informative about their discrepancy with the data, a portion of the literature has also considered less demanding limited information methods and focused on whether a model matches the data only along certain dimensions. For example, following Rotemberg and Woodford (1997), Christiano, et. al. (2005), it is now common to estimate structural parameters by quantitatively matching the conditional dynamics in response to certain structural shocks. Regardless of the approach a researcher selects, the stochastic general equilibrium model one uses to restrict the data is taken very seriously: in both estimation and testing, it is in fact implicitly assumed that the model is the DGP of the actual data, up to a set of serially uncorrelated measurement errors. Despite the above mentioned progress, such an assumption is probably still heroic to be credibly entertained. As a consequence, estimates of the parameters may reflect this primitive misspecification and, as the sample size grows, parameter estimates need not converge to those of the true DGP.

The 1990s have also witnessed an extraordinary development of VAR techniques: from simple reduced form models, VARs have evolved into tools to analyze questions of interest to academics and policymakers. Structural VARs have enjoyed an increasing success in the profession for two reasons: they are easy to estimate and the computational complexities are of infinitesimal order relative to those of structural techniques; structural inference can be performed without conditioning on a single, and possibly

misspecified, model. Clearly, there is no free lunch and robustness against misspecification comes at the cost of limiting the type of policy exercises one can entertain. One additional advantage of structural VARs needs to be mentioned. While techniques to deal with parameter variations are sufficiently well developed in this literature (see e.g. Cogley and Sargent (2005), Primiceri (2005) or Canova and Gambetti (2007)), they are still at an infant stage when it comes to structurally estimating time variations in the parameters of a stochastic general equilibrium model (see Justiniano and Primiceri (2006) or Fernandez Villaverde and Rubio Ramirez (2007)).

When addressing an empirical problem with a finite amount of data, one has therefore to take a stand on how much theory she wants to use to structure the available data prior to estimation. If the former approach is taken (which we will call "structural" for simplicity), model based estimation can be performed but inference is valid only to the extent that the model correctly represents the DGP of the data. If the latter approach is taken (which we call "SVAR" for simplicity), one can work with a class of structural models, use implications which are common to the members of this class to identify shocks and trace out their effects on the endogenous variables of the system but can not say much about, e.g. preference or production function parameters, nor conduct certain policy exercises which involve changes in expectation formation. The choice between the two alternatives is easy in two extreme and unlikely situations: the stochastic models one writes down is in fact the DGP of the actual data; the mapping from structural models to reduced form ones is univocal. Under these two conditions, direct (structural) or indirect (SVAR) estimation will give similar answers to a set of core questions investigators like to study (transmission of certain disturbances, effects of shocks to certain policy rules, etc.) and for these questions, accuracy and computational time become the most important factors that determine the choice of technique.

Unfortunately, the reality is far from the ideal and both approaches have important shortcomings. Current dynamic stochastic general equilibrium (DSGE) models, even in the large scale versions that are now used in central banks and international institutions, are still too simple to capture the complexities of the macrodata. In addition, because they are highly non-linear in the structural parameters and the mapping between structural parameters and the coefficients of the aggregate decision rules is analytically unknown - the exact mapping is known only in a few but uninteresting cases - the identification of the structural parameters from the data is far from clear. Structural VAR estimation also faces identification problems. Identification restrictions

researchers use are often conventional, have little economic content, and are not derived from any class of models that macroeconomists use to interpret the results. Furthermore, there are DSGE models which do not admit a finite order VAR representation and others which can not be recovered when the Wold decomposition is used to setup a VAR. Omitted variables may play an important role in SVAR results and the use of small scale systems may distort the conclusions one draws from the exercise. In both cases, small samples, or samples which contain different regimes, may further complicate the inferential problem. All in all, the issues of misspecification, of identification, of low signal to noise ratio, of invertibility, of omitted variables and reduced number of shocks, and last but not least, small samples, should always be in the back of the mind of an investigator who is interested in studying an applied problem and/or suggest policy recommendations from her analysis.

The scope of this chapter is to highlight the problems one faces when using either of the two methodologies to conduct policy analyses, to address questions concerning the validity of models and their ability to capture features of the data, and, in general, empirical issues of interest to academics and to policymakers. In particular, we discuss identification problems and problems connected with the potential non-representability of the aggregate decision rules with VARs. The problems we describe do not have a solution yet and standard approaches to deal with them may make the problems worse. We provide a list of "do and don't" which applied investigators may want to keep in mind in their work and outline a methodology, combining ideas from both types of approaches, which can potentially avoid some of the problems we discuss and allow useful inference on interesting economic questions. Nevertheless, it should be clearly that asking too much from a model is equivalent to asking for trouble. One should use theory as a flexible mechanism to organize the data and avoid questions that the data, the nature of the model, or the estimation approach employed can not answer.

2 DSGE models

DSGE models are consistent theoretical laboratories where the preferences and the objective functions of the agents are fully specified, the general equilibrium interactions are taken good care, the stochastic structure of the driving forces exactly defined, the expectations of the agents consistently treated and the equilibrium of the economy clearly spelled out. The economic decisions of the agents are derived under the assumption

that they maximize their objectives in a rational, forward looking manner. Individual optimality conditions are highly non-linear functions of the parameters of agents' objective functions and constraints and of the variables which are predetermined and exogenous to their actions. Given the complicated nature of these conditions, explicit decision rules, expressing the choice variables as a function of the predetermined and exogenous variables and the parameters, are not generally available in a closed form. Hence, it is typical to use numerical procedures to approximate these functions, either locally or globally. The solutions to the individual problem are then aggregated into total demand and supply curves, the equilibrium for the economy is computed and perturbations produced by selected disturbances are analyzed to understand both the mechanics and the timing of the adjustments back to the original equilibrium.

Under regularity conditions, we know that a solution to agents' optimization problems exists and is unique. Hence, one typically guesses the form of the solution, uses a particular functional form to approximate the guess and calculates the coefficients of the approximating function which, given the stationarity of the problem, must be the same for every t . For most situations of interest, (log-)linear or second order approximations, computed around a carefully selected pivotal point, suffice. The optimality conditions of agent's problems in (log)-linearized deviation from the steady state are

$$0 = E_t[A(\theta)x_{t+1} + B(\theta)x_t + C(\theta)x_{t-1} + D(\theta)z_{t+1} + F(\theta)z_t] \quad (1)$$

$$0 = z_{t+1} - G(\theta)z_t - e_t \quad (2)$$

where θ is a vector which includes the parameters of preferences, technologies, and policies; $A(\theta), B(\theta), C(\theta), D(\theta), F(\theta)$ are continuous and differentiable functions of θ , x_t are the endogenous variables of the model, and z_t the uncontrollable driving forces, which are typically assumed to follow an AR(1) with possibly contemporaneously correlated errors. These approximate individual optimality conditions are numerically solved to produce individual decisions rules which can be equivalently written in a restricted state space format,

$$\begin{aligned} x_{1t} &= J(\theta)x_{1t-1} + K(\theta)e_t \\ x_{2t} &= G(\theta)x_{1t} \end{aligned} \quad (3)$$

where x_{1t} are the predetermined and exogenous variables and x_{2t} are the choice variables of the agents, or in a restricted VAR format

$$A_0(\theta)x_t = H_1(\theta)x_{t-1} + H_2(\theta)E_t \quad (4)$$

where $A_0(\theta) = \begin{bmatrix} I & 0 \\ I & -G(\theta) \end{bmatrix}$, $H_1(\theta) = \begin{bmatrix} J(\theta) & 0 \\ 0 & 0 \end{bmatrix}$, $H_2(\theta) = \begin{bmatrix} K(\theta) & 0 \\ 0 & 0 \end{bmatrix}$, $E_t = \begin{bmatrix} e_t \\ 0 \end{bmatrix}$.

The solution of a log-linearized DSGE model has therefore the same format as well known time series models and this makes it particularly attractive to applied macroeconomists with some time series background. However, several unique features of the individual decision rules produced by DSGE models need to be noted. First, (3)-(4) are non-linear in the structural parameters θ , and it is θ and not J , K or G , that a researcher is typically interested in. Second, the decision rules feature cross equation restrictions, in the sense the $\theta_i, i = 1, 2, \dots$ may appear in several of the elements of the matrices J , K and G . Third, the number of structural shocks is typically smaller than the number of endogenous variables the model generates. This implies singularities in the covariance of the x_t 's, which are unlikely to hold in the data. Finally, H_1 and H_2 are of reduced rank. Note that if A_0 is invertible (4) can be transformed into

$$x_t = M_1(\theta)x_{t-1} + v_t \quad (5)$$

where $M_1(\theta) = A_0(\theta)^{-1}H_1(\theta)$, $v_t = A_0(\theta)^{-1}H_2(\theta)E_t$ and a (reduced form) VAR representation for the theoretical model could be derived. As we will see, the non-linearity in the mapping between the θ and the J, K, G makes identification and estimation difficult, even when cross equation restrictions are present. System singularity, on the other hand, is typically avoided by adding measurement errors to the decision rules or by considering only the implications of the model for a restricted number of variables - in this case the number of variables is equal to the number of exogenous variables. Finally, rank failures are generally avoided integrating variables out of (4) and obtaining a new representation featuring invertible matrices. As we will see, such an integration exercise is not harmless. In fact, this reduction process will in general produce a VARMA representation for the individual decision rules of the DSGE model. Hence, aggregate decision rules may not be always representable with a finite order VAR.

Given the linearity of (3) or (4) in the predeterminate and exogenous variables, aggregate decision rules will also be linear in predetermined and exogenous variables. Therefore, given values for the θ vector, time series can be easily simulated, responses to exogenous impulses calculated and sources of business cycle fluctuations examined.

How does one select the θ vector used in simulation exercises? Until a few years ago, it was common to calibrate θ so that selected statistics of the actual and simulated data were close to each other. This informal selection process was justified by the fact that DSGE models were too simple and stylized to be faced with rigorous statistical

estimation. In recent years the complexity of models has increased; a number of frictions have been introduced on the real, the monetary and, at times, the financial side of the economy; a larger number of disturbances has been considered and a number of more realistic features added. Therefore, it has become more common to attempt structural estimation of the θ using limited information approaches, such as impulse response matching exercises, or full information ones, such as likelihood based methods.

A clear precondition for any structural estimation approach to be successful is that the parameters of interest are identifiable from the chosen objective function. In the next subsection we discuss why parameter identifiability may be hard to obtain in the context of DSGE models and why, perhaps, calibration was originally preferred by DSGE modelers.

2.1 Identification

Identification problems can emerge in three distinct situations. First, a model may face identification problems in population, that is, the mapping between the structural parameters and the parameters of the aggregate equilibrium decision rule is ill-conditioned. We call this phenomena "solution identification" problem. Since the objective functions is typically a deterministic transformation of either (3) or (4), failure to identify the θ from the entries of the aggregated versions of $J(\theta)$, $K(\theta)$, $G(\theta)$ matrices (or from the aggregate versions of $A_0(\theta)$, $H_1(\theta)$, $H_2(\theta)$ matrices) is sufficient for having population identification problems for all possible choices of objective functions.

Second, it could be that identification pathologies emerge because the selected objective function neglects important model information - for example, the steady states or the variance-covariance matrix of the shocks. In other words, one can conceive situations where all the structural parameters are identifiable if the whole model is considered, but some of them can not be recovered from, say, a subset of the equations of the model or a subset of the responses to shocks. We call this phenomenon "limited information identification" problem. A trivial example of why this may happen is the following. Suppose you have two variables, say output and inflation, and two shocks, say technology and monetary shocks. Obviously, the responses to technology shocks carry little information for the autoregressive parameter of the monetary shock. Hence, this parameter is unlikely to be identified from the dynamics induced by technology shocks. It should also be clear that limited information and solution identification problems are independent of each other and therefore may appear in isolation or jointly.

Finally, difficulties in identifying parameters may be the result of small samples. That is to say, even if the mapping between the structural parameters and the parameters of the aggregate decision rules is well behaved and even if the objective function considers all the implications of the model, it may be difficult to recover structural parameters because the sample does not contain enough information to invert the mapping from the $J(\theta)$, $K(\theta)$, and $G(\theta)$ or from the objective function to the θ . To understand why this problem may emerge, consider the likelihood function of one parameter, for a given data set. It is well known that, as the sample size increases, the shape of the likelihood function changes becoming more sharply peaked around the mode. Therefore, when the sample is small, the likelihood function may feature large flat areas in a relevant portion of the parameter space and this may make it difficult to infer the parameter vector which may have been generating the data.

Econometricians have been long concerned with identification problems (see e.g. Liu (1960) or Sims (1980) among others). When models are linear in the parameters, and no expectations are involved, it is relatively straightforward to check whether the first two types of problems are present: it is sufficient to use rank and order conditions and stare at the mapping between structural parameters and the aggregate decision rules. It is also easy to measure the extent of small sample issues - the size of the estimated standard errors or an ill-conditioned matrix of second order derivatives of the objective function evaluated at parameter estimates give us an indication of the importance of this problem. For DSGE models none of these diagnostic can really be used. Since the mapping between the θ and the parameters of (3) or (4) is non-linear, traditional rank and order conditions do not apply. Furthermore, the size of estimated standard errors is insufficient to inform us about identification problems.

If identification problems are detected what can one do? While for the first type of problems there is very little to be done, except going back to the drawing board and respecifying or reparametrizing the model, the latter two problems could in principle be alleviated by specifying a full information objective function and by adding external information. If one insists in using a limited information criteria, one then needs to experiment with the subset of the model's implications to be used in estimation. Such experimentation is far from straightforward because economic theory offers little guidance in the search, and because certain variables produced by the model are non-observable (e.g. effort) or non-measurable (e.g. capital) by the applied researcher. Information from external sources may not be always available; it may be plagued by

measurement errors or be not very informative about the parameters of interest (see Boivin and Giannoni (2005)).

DSGE models face a large number of population identification problems. Canova and Sala (2005) provide an exhaustive list of potentially interesting pathologies. To summarize their taxonomy: a number of DSGE models, with potentially different economic implications, may be observationally equivalent in the sense that the aggregate decision rules they produce will be indistinguishable; they may be subject to under or partial identification of their parameters, i.e. a set of parameters may disappear from the aggregate decision rules or enter only in a particular functional form; and they may face weak identification problems - the mapping between structural parameters and the coefficients of the aggregate decision rules may display little curvature or be asymmetric in some direction. All these problems could occur locally or globally in the parameter space. However, given the common practice of obtaining estimates using optimization routines which constrain the search of the maximum to an interval, we will consider only local problems in what follows. Also, while the econometric literature has often considered the latter as a small sample problem, weak identification problems easily occur in population. In other words, while it is generally true that when the sample size is small, the curvature of the mapping may not be sufficient to recover the underlying vector of structural parameters from the coefficients of the aggregate decision rules, there is nothing that ensures that such a mapping in DSGE models will be better behaved with an infinitely large sample.

Next, we present two examples which show the pervasiveness of population identification problems in standard DSGE models used. While the models are of small scale, it should be remembered that most of the larger scale DSGE models used in the literature feature the equations of these models as building blocks. Therefore, the problem we highlight are likely to emerge also in more complex setups.

2.1.1 Example 1: Observational equivalence

Consider the following three equations:

$$y_t = \frac{1}{\lambda_2 + \lambda_1} E_t y_{t+1} + \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} y_{t-1} + v_t \quad (6)$$

$$y_t = \lambda_1 y_{t-1} + w_t \quad (7)$$

$$y_t = \frac{1}{\lambda_1} E_t y_{t+1} \quad \text{where} \quad y_{t+1} = E_t y_{t+1} + e_t \quad (8)$$

where $\lambda_2 \geq 1 \geq \lambda_1 \geq 0$ and v_t , w_t and e_t are iid processes with zero mean and variance σ_v^2 , σ_w^2 , σ_e^2 respectively. It is well known that the unique stable rational expectations solution of (6) is $y_t = \lambda_1 y_{t-1} + \frac{\lambda_2 + \lambda_1}{\lambda_2} v_t$ and that the stable solution of (8) is $y_t = \lambda_1 y_{t-1} + e_t$. Therefore, if $\sigma_w = \sigma_e = \frac{\lambda_2 + \lambda_1}{\lambda_2} \sigma_v$, a unitary impulse in the three innovations will produce the same responses for y_{t+j} , $j = 0, 1, \dots$, in the three equations and these are given by $[\frac{\lambda_2 + \lambda_1}{\lambda_2}, \lambda_1 \frac{\lambda_2 + \lambda_1}{\lambda_2}, \lambda_1^2 \frac{\lambda_2 + \lambda_1}{\lambda_2}, \dots]$.

What makes the three processes equivalent in terms of impulse responses? Clearly, the unstable root λ_2 in (6) enters the solution only contemporaneously. Since the variance of the shocks is not estimable from normalized impulse responses (any value simply implies a proportional increase in all the elements of the impulse response function), it becomes a free parameter which we can arbitrarily select to capture the effects of the unstable root. Turning things around, the dynamics produced by normalized impulses to these three processes will be observationally equivalent because λ_2 is left underidentified in the exercise.

While equations (6)-(8) are stylized, it should be kept in mind that many refinements of currently used DSGE models add some backward looking component to a standard forward looking one, and that the current Great Moderation debate in the US hinges on the existence of determinate vs. sunspot solutions (see e.g. Lubik and Schorfheide (2004)). What this example suggests is that these features may be indistinguishable when one just looks at normalized impulse responses.

How can one avoid observationally equivalence? Clearly, part of the problem emerges because normalized impulse responses carry no information for the unstable root λ_2 . However, the variance of the shocks does have this information and e.g. the likelihood function of the first process will be different from those of the other two. Hence, adding information could help limit the extent of observationally equivalence problems. In the case one is not willing or can not use this information and only employs the dynamics in response to normalized shocks to recover structural parameters, information external to the models needs to be brought in to disentangle various structural representations (as it is done e.g. in Boivin and Giannoni (2006)).

2.2 Example 2: Identification problems in a NK model

Throughout this subsection we assume that the investigator knows the correct model and the restrictions needed to identify the shocks. Initially, we assume that she chooses as objective function to be minimized the distance between the responses in the model

and in the data. Later on, we examine how identification is affected when additional information is brought into estimation process.

We consider a well known small scale New-Keynesian model, which has become the workhorse in academic and policy discussions, and constitutes the building block of larger scale models currently estimated in the literature. Several authors, including Ma (2002), Beyer and Farmer (2004), Nason and Smith (2005) and Canova and Sala (2005) have pointed out that such a structure is liable to identification problems. Here we discuss where and how these problems emerge.

The log-linearized version of the model consists of the following three equations:

$$y_t = \frac{h}{1+h}y_{t-1} + \frac{1}{1+h}E_t y_{t+1} + \frac{1}{\phi}(i_t - E_t \pi_{t+1}) + v_{1t} \quad (9)$$

$$\pi_t = \frac{\omega}{1+\omega\beta}\pi_{t-1} + \frac{\beta}{1+\omega\beta}E_t \pi_{t+1} + \frac{(\phi+\nu)(1-\zeta\beta)(1-\zeta)}{(1+\omega\beta)\zeta}y_t + v_{2t} \quad (10)$$

$$i_t = \lambda_r i_{t-1} + (1-\lambda_r)(\lambda_\pi \pi_{t-1} + \lambda_y y_{t-1}) + v_{3t} \quad (11)$$

where h is the degree of habit persistence, ϕ the relative risk aversion coefficient, β is the discount factor, ω the degree of price indexation, ζ the degree of price stickiness, ν the inverse elasticity of labor supply while $\lambda_r, \lambda_\pi, \lambda_y$ are monetary policy parameters. The first two shocks follow an AR(1) process with parameters ρ_1, ρ_2 , while v_{3t} is iid. The variances of the shocks are denoted by $\sigma_i^2, i = 1, 2, 3$. For the sake of presentation, we assume that the shocks are contemporaneously uncorrelated even though, in theory, some correlation must be allowed for.

Since the model features three shocks and three endogenous variables, we can construct several limited information objective functions, obtained considering the distances of all the responses to only one type of shock, the distance of the responses of a subset of the endogenous variables to all shocks, and the distance of the responses of all variables to all shocks.

The model has 14 parameters: $\theta_1 = (\sigma_1^2, \sigma_2^2, \sigma_3^2)$ are under-identified from scaled impulse responses, just as in the previous example, $\theta_2 = (\nu, \zeta)$ can not be separately identified - they enter only in the slope of the Phillips curve and in a multiplicative fashion, while $\theta_3 = (\beta, \phi, h, \omega, \lambda_r, \lambda_\pi, \lambda_y, \rho_1, \rho_2)$ are the parameters of interest.

To construct aggregate decisions rules numerically, we set $\beta = 0.985, \phi = 2.0, \nu = 1.0, \zeta = 0.68, \omega = 0.75, h = 0.85, \lambda_r = 0.2, \lambda_\pi = 1.55, \lambda_y = 1.1, \rho_1 = 0.65, \rho_2 = 0.65$. With the aggregate decision rules we compute population responses and use twenty equally weighted responses to construct the distance function. We explore the shape

of the distance function in the neighborhood of this parameter vector, by tracing out how it changes when we change either one or two parameters belonging to θ_3 at the time, keeping the others fixed at the chosen value. As we have mentioned, identification problems could be due to solution or objective function pathologies. Here we convolute the two mappings, and directly examine how the shape of the objective function varies with θ , because the graphical presentation of these separate mappings is cumbersome.

Figure 1 plots the shape of the distance function when we vary β, ϕ, ω, h . Column 1 presents the distance function obtained using the responses of all three variables to monetary shocks, column 2 the distance function obtained using the responses of inflation to all shocks, column 3 the distance function obtained using the responses of all variables to all the shocks. The range for the parameters considered is on the x-axis while the height of the distance function for each parameter value is on the y-axis.

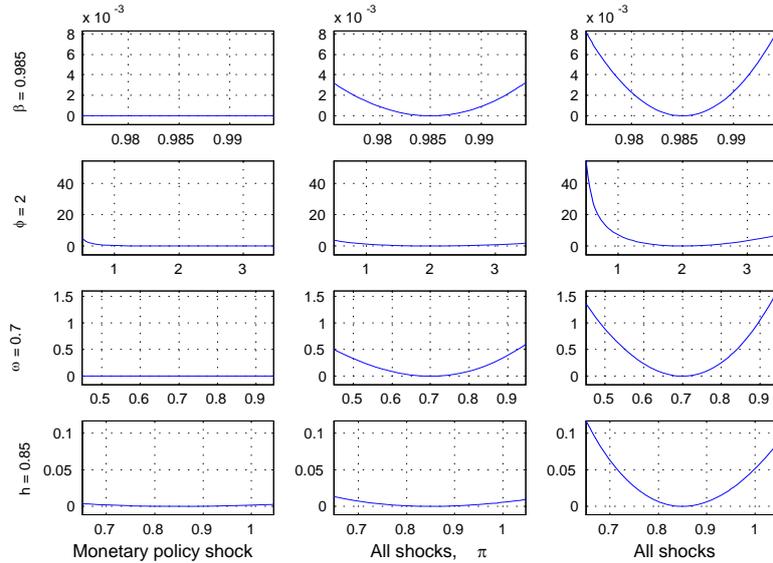


Figure 1: Shape of the distance function

It is easy to see that monetary shocks have hard time to identify the four structural parameters over the chosen intervals (the distance function is extremely flat in each of the parameters), that considering the responses of inflation to all shocks still leaves the coefficient of relative risk aversion pretty much underidentified, and that considering all the responses to all the shocks makes the distance function much better behaved. Still, asymmetries in the mapping between the risk aversion coefficient and the distance

function remain even in this latter specification. Hence, taking a limited information approach, either in the sense of considering the responses of all variables to one shock or of one variable to all shock, is problematic from an identification point of view.

One may wonder if this behavior is due to the choice of the parameters around which we map the distance function. The answer is negative. Canova and Sala (2005) construct a concentration statistics, defined as $C_{\theta_0}(i) = \int_{j \neq i} \frac{g(\theta) - g(\theta_0) d\theta}{\int (\theta - \theta_0) d\theta}$, $i = 1, \dots, 9$, where g represents the distance function and θ_0 the pivot point, and let θ_0 vary over a reasonable range. Such a statistic synthetically measures how the multidimensional slope of the distance function change around the selected parameter vector (see Stock, Wight and Yogo (2002)). Canova and Sala show that the minimum and maximum of this statistics in the range of θ_0 they consider varies very little, suggesting that the problems present in figure 1 are not specific to the selected parameter vectors.

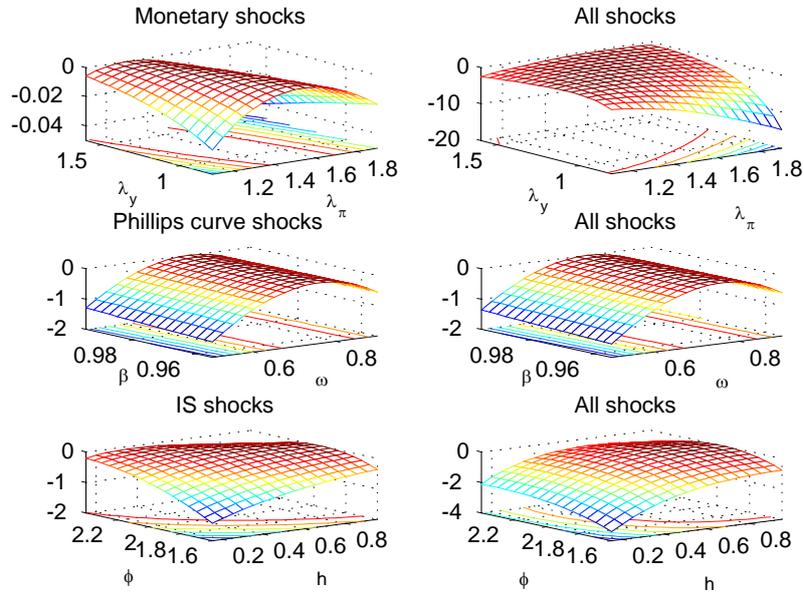


Figure 2: Distance Function and Contour Plots

Since Figure 1 considers one dimension at the time, partial identification problems, where only combinations of parameters are identifiable, can not be detected. Figure 2 shows that indeed ridges exist: for example, responses to monetary shocks carry little information about the correct combination of λ_y and λ_π ; IS shocks can not separately identify the risk aversion coefficient ϕ and the habit persistence parameter h , while

Phillips curve shocks have little information about the discount factor β . What is interesting is that when the responses to all shocks are considered, some problems are reduced. For example, there appear to be less difficulties in identifying the parameters of the policy rule when all the responses to all shocks are considered - the distance function is more bell shaped even though there is a significantly large flat area. However, even in this case, the true values of β , ϕ and h are difficult to pin down.

This model, in addition to partial, weak and under-identification problems faces generic observational equivalence problems. For example, it would be hard to detect whether the data is generated by an indeterminate version of the model (which would be the case if $\lambda_\pi < 1$) or an determinate one ($\lambda_\pi > 1$) so long as the other parameters are allowed to be adjusted. The figure below, which is reproduced from Canova and Gambetti (2007), shows that the shape and, in many cases, the size of the responses at almost all horizons to the three shocks are similar in the two regimes. Hence, if this were the only information available to the investigator, it would be difficult to detect which regime has generated the data.

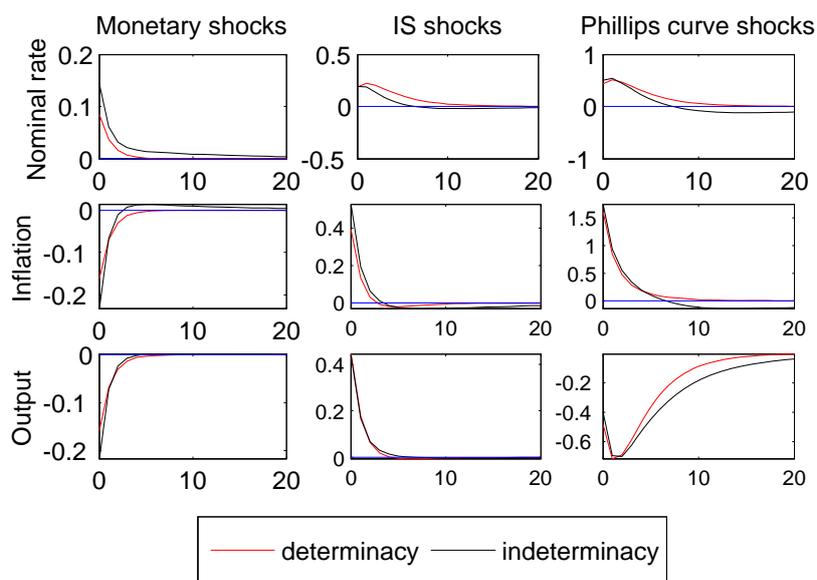


Figure 3: Impulse responses: determinate vs indeterminate equilibrium

This latter problem is a special case of a general pathology that applied investigators

often face when dealing with DSGE models: the objective functions that one constructs from the aggregate decision rules may display multiple peaks, which may be clearly separated (as is the case in the above example, see also Lubik and Schorfheide (2004)) or not (see the example discussed in section 5 of Canova and Sala (2005)). Observation equivalence, probably more than any other identification problem, prevents to attach any meaningful economic interpretation to the outcomes of the estimation process and, obviously, to conduct any meaningful policy analysis with the estimated model.

What generates the identification problems we have detected? All the nonlinear transformations, which are necessary to go from the structural parameters to the distance function, contribute. For example, consider the case of the price indexation parameter ω , which enters nonlinearly in the model and in several of the coefficients of the aggregate decision rules, but always in combination with other parameters. The coefficients of the restricted VAR solution are inverted to compute impulse responses and their distance from the "truth" is then squared and summed. One would guess that it is just by chance that such a complex set of operations will allow the mapping from ω to the objective function to be well behaved.

The standard answer to the problems shown in figures 1 and 2 is to fix parameters with difficult identification features (after all it does not matter what value we select) and estimate the remaining ones. While this approach is common, there is no insurance that it will give meaningful answers to the questions of interest. In fact, while such a mixed calibration-estimation approach will be successful, at least in population, if the parameters which are treated as fixed are set at their true value, setting them at values which are only slightly different from the true ones, may lead estimation astray. Intuitively this happens because, for example, setting β to the wrong value implies adjustments in parameters which enter jointly with β in the coefficients of the aggregate decision rules and this may move the minimum of the function in a somewhat unpredictable way. Canova and Sala (2005) show, in the context of a simple RBC example, that these shifts may be significant and drive inference the wrong way.

What can one then do to conduct structural estimation? The distance function we have employed can be obtained approximating the likelihood function of the model. Therefore, the resulting estimators can be thought as quasi-ML estimators of the structural parameters. However, there is no reason to use such an approximation. Once the decision rules are written in a state space format, the likelihood function can be easily and efficiently computed with the Kalman filter. Therefore, identification problems

could be reduced if information about the covariance matrix of the shocks or the steady states of the model - which are not used when normalized impulse response matching is performed - are brought into the estimation. Figure 4, which plots the distance function when all the shocks are considered and the likelihood function in β and ω , and ϕ and h , indeed suggests that these parameters could be better identified from the likelihood than from the distance function - the curvature of the latter is much larger than the curvature of the former. Nevertheless, the problem with ridges remains. Since the likelihood has all the information that the model delivers, one can conclude that it is the solution mapping, rather than the objective function mapping, that induces under and partial identification problems in this example.

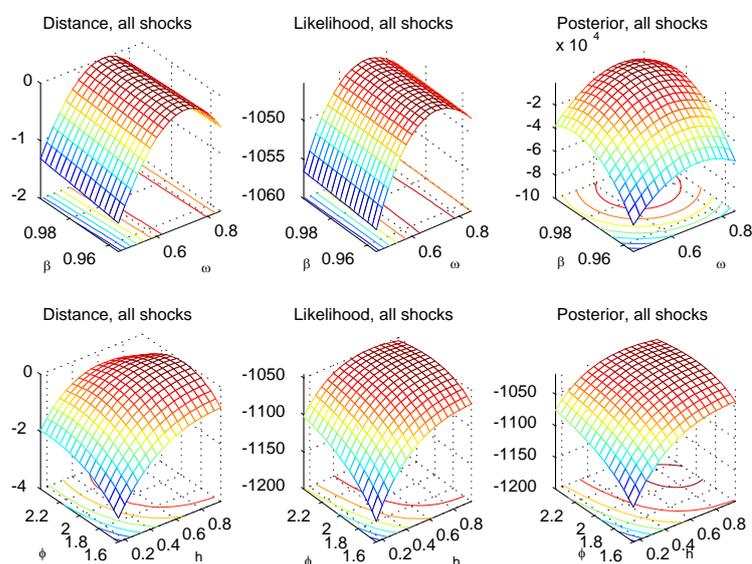


Figure 4: Distance function, Likelihood and Posterior plots

It has become quite common lately to estimate the parameters of a DSGE model by Bayesian methods. Bayesian methods attempt to trace out the shape of the posterior distribution of the structural parameters, which is proportional to the likelihood times the prior. The use of prior information could add curvature to the likelihood function therefore making identification problems apparently disappear. We show how this can happen in the last column of figure 4. A sufficiently tight prior has given the posterior a nice bell shape appearance with round contours in (β, ω) . Clearly, the use

of Bayesian methods is not the solution to identification problems we have highlighted in this subsection - it could however help when identification problems are caused by small samples. Achieving identification via prior restrictions, does not change the fact that the likelihood function constructed through the lenses of the aggregate decision rules of the model, has little information about the structural parameters. In this case the shape of the posterior distribution will, to a large extent, mimic the shape of the prior so that structural estimation is nothing more than sophisticated calibration - rather than calibrating to a point we calibrate to an interval and within the interval we assume that some parameters values are more likely than others. When population identification problems exist and a researcher is interested in estimating the structural parameters, it is necessary to reparametrize the model. If this is unfeasible or undesirable, then informal calibration is one simple and internally consistent device to make the model operative for inference and forecasting. The deep issue here is that DSGE models are not typically designed with an eye to the estimation of its parameters and this is clearly reflected in the identification problems they display.

Prior information on the parameters of macroeconomic models may come from different sources. It may be accumulated knowledge about a phenomena repeatedly studied in the literature (e.g. the properties of the transmission of monetary policy shocks), evidence obtained from micro studies or from the experience of other countries. All this information may be valuable to the applied investigator and should be formally introduced in the structural estimation of the model, if available. However, if the likelihood has little information about the structural parameters, and this additional information were the only one available to identify the parameters, structural estimation will not be particularly useful - it would resemble confirmatory analysis where prior expectations are verified a-posteriori. In this situation, policy exercises are difficult to interpret and the alternative of measuring the range of outcomes produced by the model using a selected range of parameters, as suggested in Canova (1995), is a feasible and more plausible approach.

What are the consequences of the identification problems we have described? For the sake of presentation, we will focus on estimates obtained matching responses to monetary policy shocks, which appear to produce the distance function with the worst identification properties, and are those on which the literature has paid most of its attention. In this exercise we still assume that shocks are correctly identified - in our model reduced form interest rate innovations are the true monetary policy shocks. If

this were not true, additional problems, as those discussed in e.g. Canova and Pina (2005), are compounded with those discussed here. We consider different sample sizes to highlight, on the one hand, some of the properties of the estimates of parameters with problematic identification features and, on the other, examine whether additional identification problems may emerge just because of small samples.

We simulate 200 time-series for interest rates, the output gap and inflation for $T = 120, 200, 1000$, fixing $\nu = 1$ and $\sigma_i^2 = 1.0$ in all cases, we estimate an unrestricted VAR(2), which is the correct empirical reduced form model to use in this case, compute impulse responses and bootstrap confidence bands which are then used to build a diagonal matrix of weights: the weights are inversely proportional to the uncertainty in the estimates. Table 1 presents a summary of the estimation results. It reports the true parameters, the mean estimate, the numerical standard errors computed across replications (in parenthesis) and the percentage bias (in brackets).

Table 1: NK model. Matching monetary policy shocks

	True	T = 120	T = 200	T=1000
β	0.985	0.984(0.007) [0.6]	0.985 (0.007) [0.7]	0.986 (0.008) [0.7]
ϕ	2.00	2.39 (2.81) [95.2]	2.26 (2.17) [70.6]	1.41 (1.19) [48.6]
ζ	0.68	0.76 (0.14) [19.3]	0.76 (0.12) [17.5]	0.83 (0.10) [23.5]
λ_r	0.2	0.47 (0.29) [172.0]	0.43 (0.27) [152.6]	0.41 (0.24) [132.7]
λ_π	1.55	2.60 (1.71) [98.7]	2.22 (1.51) [78.4]	2.18 (1.38) [74.5]
λ_y	1.1	2.82 (2.03) [201.6]	2.56 (2.01) [176.5]	2.16 (1.68) [126.5]
ρ_1	0.65	0.52 (0.20) [30.4]	0.49 (0.21) [34.3]	0.50 (0.19) [31.0]
ρ_2	0.65	0.49 (0.20) [32.9]	0.48 (0.21) [34.8]	0.48 (0.21) [34.7]
ω	0.25	0.76 (0.39) [238.9]	0.73 (0.40) [232.3]	0.65 (0.38) [198.1]
h	0.85	0.79 (0.35) [30.9]	0.76 (0.37) [32.4]	0.90 (0.21) [21.3]

A few features of the table are worth commenting upon. First, biases are evident in the estimates of the partially identified parameters (λ_π, λ_y), the weakly identified parameters (ζ, ω, h), and the under-identified parameters (ρ_1, ρ_2). Note that even with 250 years of quarterly data major biases remain. Second, numerical standard errors are large for the partially identified parameters and invariant to sample size for the under-identified ones. Third, parameter estimates do not converge to population values as T increases. Finally, and concentrating on $T = 200$, estimates suggest an economic behavior which is somewhat different from the one characterizing the DGP. For example, it appears that price stickiness is stronger and the Central Bank reaction to the output gap and inflation is equally strong.

In sum, identification problems lead to biased estimates of certain structural parameters (see also Choi and Phillips (1992)), to inappropriate inference when conven-

tional asymptotic theory is used to judge the significance of estimated parameters and, possibly, to wrong economic interpretations. For unconditional forecasting, identification problems are unimportant: as long as the fit and the forecasting performance is the same, the true nature of the DGP does not matter. However, policy analyses and conditional forecasting exercises conducted with estimated parameters may lead to conclusions which are very different from those obtained with the true one. Hence, it is generally unwise to attach any economic interpretation to the estimates or draw conclusions about how the economy works from structural exercises which are plagued by identification problems.

What is left to the applied investigator to do? Apart from attempting to reparametrize the model, not much. One interesting issue still unexplored in the literature is to take population identification problems as being the norm rather than the exception and try to find estimation techniques or objective functions which, given a sample size, are able to minimize the distortions produced by identification pathologies. While some progress has been made in the context of moment estimation (see Stock and Wright (2000) or Rosen (2005)), these procedures are applicable only in restrictive situations (the weighting matrix must be chosen in a particular way) and are awkward to use in DSGE models, which are highly parametrized and non-linear.

How does one detect identification problems? The univariate and bivariate exploratory analysis we have presented e.g. in figures 1 and 2 can definitely help in spotting potential problems and this analysis could be easily complemented with local derivatives of the objective function in the dimensions of interest. Alternatively, numerically computing the Hessian of the objective function around particular parameter values and calculating the size of its eigenvalues can give more formal indications on how flat or how information deficient the objective function is locally. For example, if the rank of the Hessian is less than the number of structural parameters, one of its eigenvalues is zero and at least one parameter is underidentified. If the rank of the Hessian is close to be deficient, one or more of its eigenvalues are close to zero and either weak or partial identification problems or both are likely to be present. Experimentation with number of shocks used to construct the objective function and the number of variables can also give useful information about what statistic may identify a particular structural parameter, as is the experimentation with different objective functions and with different features of the data (e.g. steady state vs. dynamics).

Clearly, diagnostics of this type have to be run prior to estimation, but such an

exercise is not much more complicated or time consuming than the type of exercises one performs to measure the sensitivity of the results to the selection of calibrated parameters. In general, the following rules of thumb are useful to limit the extent of identification problems: given a model, always choose a likelihood based objective function, which has the highest informational content; given a model and the likelihood function, and if it is the sample which is problematic, add information in the form of additional data or prior restrictions, which synthetically reproduces it.

It is important to stress that looking at the minimized value of the objective function, at standard errors of the estimates or to the resulting impulse responses it is not generally useful as an ex-post device to detect identification problems. The distance function is within the tolerance level (10^{-7}) for all the parameter combinations generating table 1 and the practice of blowing up the objective function, by appropriately choosing the matrix of weights, will not change the fact that the gradient or the Hessian display problematic features. Furthermore, it can be shown that population responses fall within a 68 percent band centered around the estimates of the responses to monetary shocks computed with the parameter estimates, even when the sample size is $T=120$. Therefore, the practice of showing that model's responses computed using the estimated parameters lie within the confidence bands of the responses estimated from the data is not particularly informative as far as identification problems are concerned. Large standard errors do emerge when identification failures exist but also when other problems are present (e.g. very noisy data or regime switches). Hence, univocally associating large standard error with identification issues is, in general, incorrect.

It is also important to stress that the addition of measurement errors for estimation purposes can distort the identification properties of structural parameters. It is not particularly difficult to conceive situations where a parameter which was identified by certain features of the model becomes free to move and fit other properties of the data it was not designed for, once measurement error is added. Therefore, while there is some logic in adding measurement errors to link the model variables to the observables, one should be careful and investigate the consequences that such a process has on the identification properties of the parameters.

3 Structural VARs

Structural VAR (SVAR) inference is typically perceived to be at the antipodes of structural model-based inference. SVAR models take a minimalist approach to the estimation problem and consider only a very limited subset of the large number of restrictions that DSGE models impose on the data. For example, the fact that the matrices H_1 and H_2 depend on θ is typically neglected and only a part of the information present in $A_0(\theta)$ is used. Furthermore, the singularity that the model imposes on the data is completely disregarded. This minimalist approach has one obvious disadvantage: if less structure is imposed on the data, fewer interesting economic questions can be asked. However, such a limited information approach is advantageous when some of the model's restrictions are dubious, which would be the case if the model is misspecified in some dimensions, or fragile, which would be the case if the restrictions depend on the functional forms or the parameter values one specifies. In this case, neglecting these restrictions, can robustify estimation and inference.

As we have mentioned in the introduction, and despite recent attempts to make them more realistic, the current generation of DSGE models is still far from reproducing the DGP of the actual data in many respects: models fail to capture, e.g., the heterogeneities present in the actual world; important relationships are modeled with black-box frictions; timing restrictions are used to make them compatible with the dynamics observed in the data; and ad-hoc shocks are often employed to dynamically span the probabilistic space of the data. Since misspecification is likely to be pervasive, system wide and even limited information classical structural methods are problematic even when identification problems are absent.

Bayesian methods have an hedge in structural estimation when model misspecification is present. Inference in this context, in fact, does not require the asymptotic correctedness of the model under the null. Furthermore, these methods can potentially deal with model misspecification, for example, by imposing prior distributions over models and weighting the posterior information contained in each of them by their posterior probability. However, this potential advantage of Bayesian methods is often unexpressed: except for Schorfheide (2000), it is very unusual for researchers to consider an array of models, all of which can potentially be useful to answer the question of interest. In this situation, one is often left wondering what posterior estimates obtained from a misspecified model mean in practice and whether policymakers could and should

trust these estimates when taking important policy decisions.

The difficulties of the current generation of DSGE models to represent the DGP of the data have been highlighted by Del Negro et. al. (2006), who take a workhorse model, popular among academics and central bankers, and show that it is possible to improve its fit by considerably relaxing the cross equation restrictions that it imposes on the matrices $H_1(\theta)$ and $H_2(\theta)$. Their approach, which uses a DSGE model as a prior for a VAR, is useful to design a metric to assess the distance between the model and the VAR of the data and represents a promising way to evaluate model fit, to suggest ways to bring models in closer contact with the data and, in general, to conduct structural inference in misspecified models.

If one takes the inherent misspecification that the current generation of DSGE models display seriously and heavily weights inferential mistakes, one may then want to proceed in a more agnostic way. Rather than conditioning on one model and attempting to estimate its structural parameters, one could be much less demanding in the estimation process, and employ a subset of the model restrictions, which are either uncontroversial or likely to be shared by a class of economies with potentially different features, to identify structural shocks. One way of doing this is precisely to neglect the restrictions present in the matrices H_1 and H_2 , which are often not robust, and use some of those present in $A_0(\theta)$, for which a large a-priori consensus can be found in theory, and then trace out the dynamics of the variables of interest in response to disturbances or measure the relative importance of each shock for business cycle fluctuations. Therefore, with such an approach, most of the detailed cross equation restrictions imposed by a model will be eschewed from the estimation process and only constraints which are likely to hold in many models, are used to identify structural shocks. Unfortunately, it has become common in the literature to employ constraints which unrelated to any specific class of models or so generic that they lack economic content. While 20 years ago researchers spent considerable time and effort justifying their identification restrictions from a theoretical point of view (see e.g. Sims (1986) or Bernanke (1986)), now it is often the case that these restrictions are not even spelled out in details, and the only justification for them a reader can find is that they are used because someone else in the literature has used them before. In general, delay type of restrictions, which use the flow of information accrual to agents in the economy, and place zeros in the impact matrix of shocks, are the preferred identification devices.

Canova and Pina (2005) have shown that delay type of restrictions do not naturally

arise in general equilibrium models, are often inconsistent with their logic and one has to work hard to cook up general equilibrium environments with such features (see e.g. Rotemberg and Woodford (1997)). Long run restrictions have been hailed in the past as the answer to these problem since restrictions of this type are common to a variety of theories (e.g. money neutrality or the idea that technological progress explain the long run path of variables are features which are shared by macro models with different microfounded) and allow inference without tying one's hand to a particular specification for the short run dynamics around these long run paths. However, this alternative identification approach is non-operative: long run restrictions are scarce relative to the number of shocks researchers are interested in recovering. Therefore, when four or five shocks need to be identified, one is forced to use a mixture of long run and delay restrictions. Furthermore, as pointed out by Faust and Leeper (1997), long run restrictions are weak and prone to observational equivalence problems.

The sign and shape approach, suggested in Canova and De Nicolò (2002) and Uhlig (2005), we advocate in the next section, can bridge SVAR and DSGE models in a more solid way and provide a constructive answer to the quest for identification restrictions. Unfortunately, such an approach does not yet have a widespread use in the profession (exceptions are, among others, Dedola and Neri (2007) and Pappa (2005)) and the science of identification in SVARs is still very much the craft of finding restrictions that would not bother anyone in the profession.

Apart from identification issues, which have received attention in the VAR literature since, at least, Cooley and Leroy (1985), a number of authors have recently questioned the ability of structural VARs to recover the true DGP of the data, even when an appropriate identification approach is used. To see why this could be the case, consider the following alternative restricted state space representation for the log-linearized aggregate decision rules of a DSGE model

$$\begin{aligned}x_{1t} &= J(\theta)x_{1t-1} + K(\theta)e_t \\x_{2t} &= N(\theta)x_{1t-1} + M(\theta)e_t\end{aligned}\tag{12}$$

where $e_t \sim iid(0, \Sigma_e)$. The questions we ask are the following: (i) does there exist a VAR representation for a subset of the variables of the model, say x_{2t} ? (ii) Would the resulting VAR be of finite order? (iii) What would happen to inference if only a sample of limited size is available? We have already mentioned that if both x_1 and x_2 were observable (12) is simply a restricted, although reduced rank VAR(1).

However, this is not a very realistic setup: usually x_{1t} includes non-observable variables; furthermore, only a subset of the variables appearing in x_{2t} may be of interest, could be reasonably measured, or have relevant information for the exercises one may want to conduct. Therefore, it is legitimate to ask what would the process of integrating out non-observable, uninteresting or badly measured variables imply for the restricted time series representation of the aggregate decision rules of the model.

3.1 Invertibility

If $M(\theta)$ is a square matrix, and if $J(\theta) - K(\theta)M(\theta)^{-1}N(\theta)$ has all eigenvalues less than 1 in absolute value, it is easy to show that:

$$x_{2t} = N(\theta)\{[1 - (J(\theta) - K(\theta)M(\theta)^{-1}N(\theta))]^{-1}K(\theta)M(\theta)^{-1}\}x_{2t-1} + u_t \quad (13)$$

where $u_t \sim (0, M(\theta)' \Sigma_e M(\theta))$. Therefore, if only x_{2t} is observable, the aggregate decision rules have a restricted VAR(∞) representation. If instead $N(\theta)$ is a square matrix, then:

$$x_{2t} = N(\theta)J(\theta)N(\theta)^{-1}x_{2t-1} + (I + (N(\theta)K(\theta)M(\theta)^{-1} - N(\theta)J(\theta)N(\theta)^{-1})\ell)u_t \quad (14)$$

where ℓ is the lag operator. Under this alternative assumption, the aggregate decision rule for x_{2t} has therefore a VARMA(1,1) representation.

Hence, if a reduced number of variables is considered, the aggregate decision rules of the model have a much more complicated structure than a restricted VAR(1). The question of interest is whether we can still use a VAR with a finite number of lags to approximate the aggregate decision rules for x_{2t} . Straightforward algebra can be used to show that if the exogenous driving forces are AR(1) and if both the predetermined states and x_{2t} are observed, then the correct representation for the vector of predetermined states and choice variables is a restricted VAR(2) with singular covariance matrix. On the other hand, if only x_{2t} is observable and the dimension of x_{2t} is the same as the dimension of e_t , Ravenna (2006) has shown that the aggregate decision rules for x_{2t} can be approximated with a finite order VAR if and only if the determinant of $\{I - [J(\theta)K(\theta)M(\theta)^{-1}N(\theta)]\ell\}$ is of degree zero in ℓ .

What does this all mean? It means that the aggregate decision rules for a subset of the variables of the model can be represented with a finite order VAR only under a set of restrictive conditions. These conditions include invertibility of the mapping between structural shocks and the variables included in the VAR, a fundamentalness condition,

which implies that the information contained in the observables is the same as the information contained in disturbances of the model, and the condition that random perturbations produce fluctuations around the steady state that are not too persistent.

Note that the condition we have used to derive (13), is never satisfied in practice. One can think, at best, of four or five truly structural sources of disturbances and this typically is much less than the size of vector x_{2t} . Therefore, it is only after ad-hoc disturbances and/or measurement errors are ex-post included that $M(\theta)$ is a square matrix. Similarly, the restriction that $N(\theta)$ is a square matrix is difficult to be satisfied in practice - the number of states is typically smaller than the number of endogenous variables. The other conditions clearly depend on the structure of the model but, for example, specifications in which agents react to news that may materialize in the future fail to satisfy the first condition - the resulting MA representation of the model is nonfundamental. Finally, the convergence of the economy to its steady state when perturbed by shocks depends on the details of the specification. Therefore, it is difficult to assess how important in practice this assumption is. Given that many DSGE models have fairly weak internal propagation mechanism, and as long as the structural shocks are stationary, such a condition is likely to be satisfied in practice.

In sum, one should not be surprised to find DSGE models featuring aggregate decision rules for a subset of the variables which are not representable with a finite order VAR (see e.g. Fernandez-Villaverde, et. al. (2005) for examples). Nevertheless, a large class of models does have aggregate decision rules with these properties. To be sure that SVAR inference is valid, one must preliminary select a class of models which could have generated the data and check whether the required conditions are satisfied for alternative parameterizations. While this requires a SVAR investigator to take much more seriously a certain class of models before drawing any inference from her analysis, it also makes SVAR estimation less straightforward and more time consuming since the number of parameters, functional form and friction permutations that need to be checked before the analysis is conducted is large. Furthermore, since bizarre counterexamples can always be found, it may become difficult for an applied macroeconomist to assess in practice whether a finite order VAR is a good approximation to the class of DSGE models one is interested in or not.

For the final question, Chari, et. al. (2006) have recently shown that one may be led astray in evaluating the relevance of economic theories using SVAR simply because with small samples, the population properties of the aggregate decision rules may be

very poorly approximated with a VAR. That is to say, even when there exists a VAR representation for the variables in x_{2t} , when this representation is of finite order, and when identification of shocks is properly performed, small sample biases in the estimates of the reduced form parameters and the covariance matrix of the shocks may make inference whimsical. For example, they show that a short sample of data simulated from a RBC model driven by a neutral technology shock may lead a researcher to believe that it could have been generated by a model with different microfoundations - in population hours worked increases in response to a technology shock; in small samples hours may fall in response to the correctly identified technology shocks.

An applied investigator has to leave with small sample biases. Long samples, when they are available, are rarely used because causal relationships are often subject to important regime shifts and when regime shifts are absent, changes in the definition or in the way the data is sampled or computed, make empirical analysis difficult. Econometrics can help here: it is well known that in a variety of experimental designs and with samples of about 100 observations, estimates of the AR(1) coefficient are downward bias by up to 30 percent. While this type of analysis could be easily extended to more realistic and interesting economic models - for example, measuring the size of the bias in the largest autoregressive root of the aggregate decision rule (which roughly determines the dynamics of the system) and in the eigenvalues of the covariance matrix of reduced form shocks (which determines the size of the impact effects) - one needs to consider models where the impact effect is fairly weak to have important sign reversals in small samples. Therefore, while such an issue should be kept in mind, its practical relevance appears to be limited.

There is another way of seeing these representation problems from a different and probably more informative viewpoint - the one of omitted variables and shock mis-aggregation, which have a long tradition in the VAR literature (see e.g. Braun and Mittnik (1993), Faust and Leeper (1997)). Suppose the aggregate decision rules for the endogenous variables of a DSGE model can be written as a VAR(1):

$$\begin{bmatrix} I - A_{11}\ell & A_{12}\ell \\ A_{21}\ell & I - A_{22}\ell \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} e_t$$

where y_{1t} are the variables included and y_{2t} the variables excluded from the empirical model and these vectors do not necessarily coincide with those of the state variables x_{1t} and the choice variables x_{2t} . Then the representation for y_{2t} is

$$(I - A_{22}\ell - A_{21}A_{12}(1 - A_{11}\ell)^{-1}\ell^2)y_{2t} = [B_2 - (A_{21}(1 - A_{11}\ell)^{-1}B_1)]e_t \equiv v_t \quad (15)$$

When y_{1t} and y_{2t} are of the same dimensions, simplifies to

$$[I - (A_{11} + A_{22})\ell + (A_{11}A_{22} - A_{21}A_{12})\ell^2]y_{2t} = [B_2 + (A_{21}B_1 - A_{11}B_2)\ell]e_t \equiv v_t \quad (16)$$

What does this reduced system representation imply? First, it is easy to see that the model for y_{2t} is an ARMA(∞, ∞) and the lagged effect of the disturbances mixes up the contemporaneous effects of different structural shocks (B_1e_{t-1} has smaller dimension than e_{t-1}). Furthermore, it is clear that even if e_t 's are contemporaneously and serially uncorrelated, v_t 's are contemporaneously and serially correlated and that two small scale VARs featuring different y_{2t} 's will have different v_t 's. Finally, since v_t is a linear combination of current and past e_t , the timing of the innovations in y_{2t} is not preserved unless A_{11} and A_{21} are both identically equal to zero, which is true, e.g., if y_{2t} includes the states and y_{1t} the controls of the problem.

In other words, (16) implies that shocks extracted from SVAR featuring a reduced number of the model's variables are likely to confound not only structural shocks of different types, but also display time series properties which are different from those of the true shocks to these variables. Hence, even if the correct identifying restrictions are used, VAR models which are small relative to the universe of variables potentially produced by a DSGE model are unlikely to be able to capture either its primitive structural disturbances nor the dynamics they induce unless some strong, and not very practically relevant conditions, hold.

Contrary to the previous representation of the invertibility problem which provides little guideline on how to check for failures, this latter representation does give researchers a way to measure the importance of potentially omitted variables. In fact, if omitted variables are important, reduced form VAR residuals will be correlated with them. Therefore, for any given set of variables included in the VAR, it is sufficient to check whether variables potentially belonging to y_{1t} display significant correlation with the residuals. If so, they should be included in the VAR and estimation repeated; if not they can be omitted without further ado.

To conclude, we present two examples illustrating the issues we have discussed in this section. In the first example, non-invertibility emerges because the model has a non-fundamental representation. In the second the MA of the model is invertible, but the dynamics of the reduced system are different from those of the full one.

3.1.1 Example 3: A Blanchard and Quah economy

The example we present belongs to the class of partial equilibrium models which was popular in the late 1980s. While it is not difficult to build general equilibrium models with the required features, the stark nature of this model clearly highlights how invertibility problems could occur in practice. The model that Blanchard and Quah (1989) consider has implications for four variables (GDP, inflation, hours and real wages) but the solution is typically collapsed into two equations, one for GDP growth (ΔGDP), the other for the unemployment rate (UN_t) of the form

$$\Delta GDP_t = \epsilon_{3t} - \epsilon_{3t-1} + a(\epsilon_{1t} - \epsilon_{1t-1}) + \epsilon_{1t} \quad (17)$$

$$UN_t = -\epsilon_{3t} - a\epsilon_{1t} \quad (18)$$

where ϵ_{1t} is a supply shocks, ϵ_{3t} a money supply shock and a is a parameter measuring the impact of supply shocks on aggregate demand. Hence, the aggregate decision rule for these two variables is an VMA(1). It is easy to check that a finite order VAR may approximate the theoretical dynamics of this model only if $a > 1$.

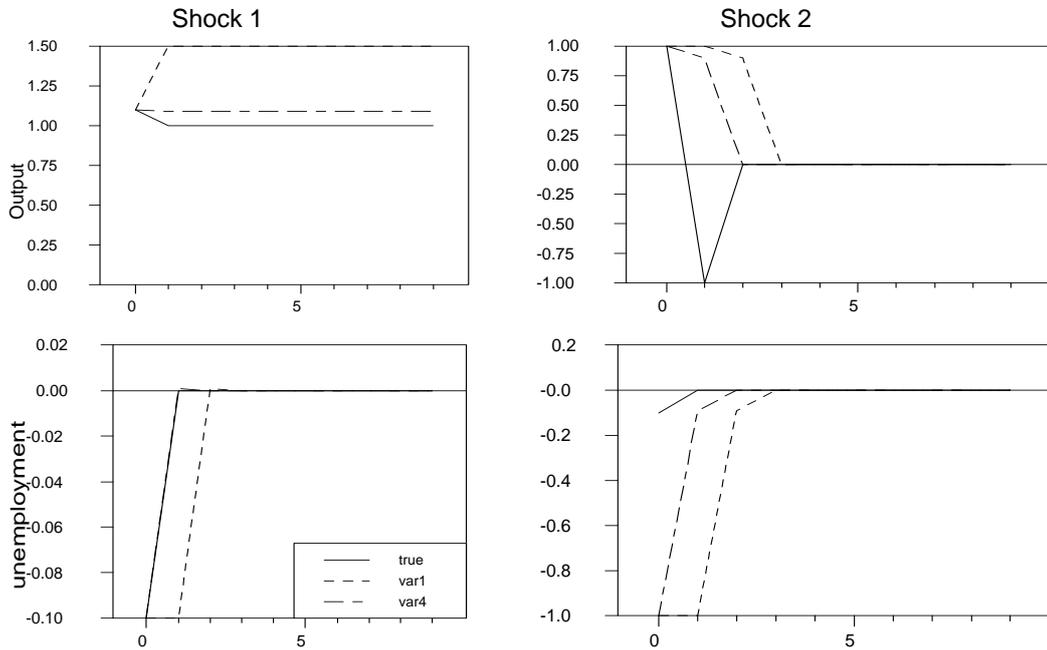


Figure 5: Responses in the Blanchard and Quah model

To see this, we set $a = 0.1$ and plot in figure 5 the theoretical responses of output and unemployment to the two shocks and the responses obtained using a VAR(1) and a VAR(4), where the econometrician uses the correct (but truncated) vector autoregressive representation of the model. Note that while the signs of the responses are correct the dynamics are pretty different. Also, while there is some improvement in moving from a VAR(1) to a VAR(4), some of the theoretical responses are very poorly approximated even with a VAR(4). Since a VAR(q), $q > 4$, has responses which are indistinguishable from those of a VAR(4) - matrices on longer VAR lags are all very close to zero - no finite order VAR can capture (17) and (18).

What generates this result? When $a < 1$ the aggregate decision rules of the model are non-fundamental, that is, innovations to output growth and unemployment do not have the same information as the variables themselves. Therefore, there is no convergent VAR representation for these two variables where the roots of the VAR are all less than one in absolute value, and this is true even when a infinite lag length is allowed for.

3.2 Example 4: A RBC model

We work with the simplest version of the model since more complicated structures do not bring additional insights into the problem. The social planner maximizes $E_0 \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\phi}}{1-\phi} - AN_t$ and the resource constraint is $c_t + k_t + g_t = k_{t-1}^\eta N_t^{1-\eta} z_t + (1 - \delta)k_{t-1}$, where c_t is consumption and ϕ is the risk aversion coefficient, A is a constant and N_t are hours worked; z_t is a first order autoregressive process of with persistence ρ_z , steady state value z^{ss} and variance σ_z^2 ; g_t is a first order autoregressive process of with persistence ρ_g , steady state value g^{ss} and variance σ_g^2 , k_{t-1} is the current capital stock, η is the share of capital in production and δ the depreciation rate of capital. Using the method of undetermined coefficients and letting output be $y_t \equiv k_{t-1}^\eta N_t^{1-\eta} z_t$, and investment be $i_t = k_t - (1 - \delta)k_{t-1}$ the aggregate decision rules for $(k_t, c_t, N_t, y_t, r_t, i_t)$, where r_t is the real rate imply standard dynamics in response to the two shocks. In particular, as z_t increases hours, consumption, output, the real rate and investment increase contemporaneously while the dynamics of the capital stock have a hump shaped pattern. On the other hand, as g_t increases, consumption falls, hours, output, the real rate and investment increase contemporaneously and the capital stock has an hump shaped pattern.

How would the dynamics induced by the two shocks in a system which includes only the interest rate and investment look like? That is, what would happen if we integrate

out the effect of the shocks on the other four variables? Figure 6 plots the responses of the two variables of interest to the two shocks in the full and the reduced systems. Clearly, while the impact effect is identical, lagged dynamics are pretty different.

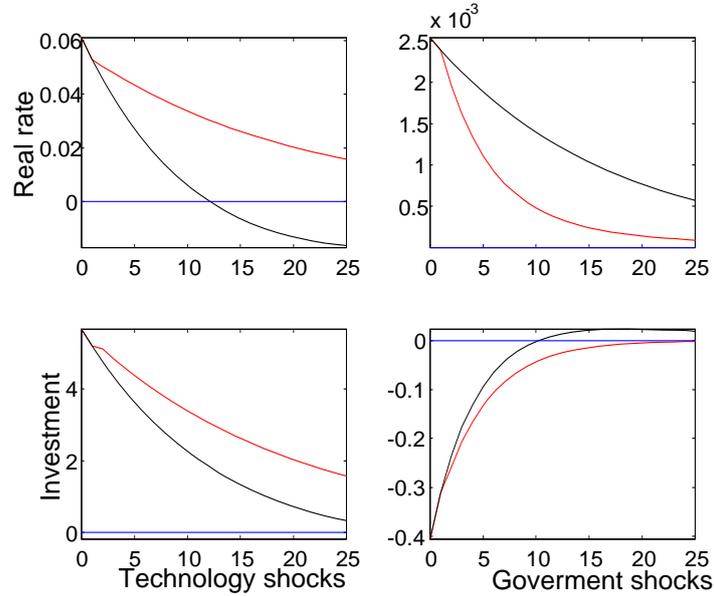


Figure 6: Dynamics in a RBC model

What is the reason for this result? Mechanically, since A_{11} and A_{21} are not small, shocks last more than one period and persist for a number of periods. Notice that the persistence in the reduced system is strong (see e.g. the effect of technology shocks on the real rate) suggesting that the process of marginalizing part of the system has serious consequences on the responses of the variables to shocks, at least in this example.

It goes without saying that it makes a lot of difference which of the two systems one uses as a benchmark to represent the DSGE model and in trying to see whether actual and simulated data are similar or not.

4 Some final thoughts

The previous two sections may have given the reader a rather pessimistic view about the possibility of conducting meaningful inference with DSGE models and the impression that not many alternatives are left to the applied investigator. If structural estimation

is pursued, misspecification of the structural relationships may make the interpretation of estimates difficult; identification problems are likely to be widespread and even in the unlikely case when they are not present, a number of additional statistical and specification assumptions need to be made, making it very difficult to judge what is causing what. The alternative of using SVARs seems to be equally problematic. While VAR are less prone to misspecification of the structural relationships, identification problems are still present and non-invertibility of the DSGE model aggregate decision rules may also make SVARs analyses uninterpretable.

Chari, et. al. (2007) have suggested to use the so-called business accounting method to evaluate DSGE models but the logic of the approach represents a step backward relative to what we discuss here - only reduced form relationships are used to judge what is missing from the model - and it is hard to avoid important observational equivalence problems when judging different structural models of the business cycle.

What should then one do? No matter which approach one takes, one should be very careful and learn how to interpret the information contained in the diagnostics obtained experimenting with the structure of the model and investigating the properties of the data. If structural estimation is performed, methods which allow for misspecification should be preferred and extra information, in the form of micro data or data from other countries, may help to break the deadlock of parameters identification when problems are due to small samples. We have suggested that to solve population identification problems it is necessary to reparametrize or respecify DSGE models, but obviously this is a more longer term goal, since such an approach brings us back to the very basic foundation of DSGE-based exercises. Nevertheless, if theorists would build models having in mind that they will be estimated, certain issues could be completely avoided.

If SVAR analysis is preferred, one should link the empirical model to DSGE theories much better than it has done so far, explicitly write down the class of models one will employ to interpret its results (as it is done e.g. Canova and De Nicrolo (2002)) and perform the preliminary analysis necessary to check whether the aggregate decision rules of such a class of models do have a finite order VAR format for the subset of relevant variables used in the VAR. Identification should also be clearly linked to the class of structural models of interest and artificial delay restrictions avoided. One way of doing this is in Canova (2002), where robust restrictions on the sign of responses to shocks derived from a class of models are used to identify shocks, and the results of the analysis are discussed through the lenses of such models. Canova and Paustian

(2007) show that such an approach has good size and power properties against local alternatives and give reasonable results in inappropriately marginalized systems.

Integrating structural and VAR analyses, as suggested by Del Negro and Schorfheide (2004), (2005) and (2006) also provides an interesting avenue for future research, where structural models and empirical analyses can cross fertilize each other.

From the point of view of policymakers DSGE models are useful if they can forecast well, since it is much easier to tell stories with estimates of their parameters than with SVAR estimates or estimates of pure time series models. However, to forecast at least as well as with more unrestricted models, the DSGE models popular in the academic literature must produce restrictions which are not rejected in the data and this is pretty hard to do when one considers, e.g., prices rather than quantities and financial or monetary variables rather than real ones. In addition, to test the quality of these restrictions one needs substantial "cosmetic surgery" in the form of additional shocks, frictions and other black-box jingles, which are difficult to justify from a theoretical point of view and make any hypothesis a joint test of the restrictions and the chosen adds on. Realizing these facts should probably lead academics and policymakers to be less demanding of the models they write down and use. Typically, small models forecast better than larger ones and different models can be used for different purposes. Having an array of models at disposal, which are built to answer different economic questions, and averaging their forecasting results may not only robustify the outcomes of the investigation but also give an entirely different perspective on the reasons driving certain economic phenomena.

While one can envision the disappearance of the "model" of the economy as conceived in the 1970s, constructed patching up pieces of theoretical structures and a lot of empirical wisdom, and used to answer all possible questions policymakers may have, it is very likely that smaller scale, more or less structurally oriented models will coexist in the portfolio of research departments of central banks and international institutions for a while, serving different purposes and different objectives.

To go back to the main question of this chapter, how much structure should there be in an empirical model? The solomonic and, probably, obvious answer, is that it depends on the scope of the analysis and the information available in the data. Different models can have different structural content if they serve different purposes. Nevertheless, it should be clear that certain policy exercises can be conducted only in models where expectations and general equilibrium features are fully taken into account and that

the predictive content of pure time series models is close to inexistent as the horizon of the forecast surpasses one year. Small scale structural models that allow a large number of policy exercises and at the same time offer some indications on the potential developments one-two years ahead are probably the ones that will survive the dust of time in the longer run.

References

- [1] An, S. and Schorfheide, F. (2007) Bayesian analysis of DSGE models, *Econometric Reviews*, 26, 113-172.
- [2] Beyer, A. and Farmer, R. (2004) On the indeterminacy of new Keynesian Economics, ECB working paper 323.
- [3] Blanchard, O. and Quah, D. (1989), The Dynamic Effect of Aggregate Demand and Supply Disturbances, *American Economic Review*, 79, 655-673.
- [4] Bernanke, B. (1986) Alternative Explanations of the Money Income Correlation, *Carnegie-Rochester Conference Series on Public Policy*, 25, 49-100.
- [5] Boivin, J. and Giannoni, M. (2005) DSGE Models in Data Rich environments, NBER working paper 12272.
- [6] Boivin, J. and M. Giannoni (2006), Has Monetary Policy Become More Effective Powerful? ,*Review of Economics and Statistics* 88(3), 445-462.
- [7] Braun, P. and Mittnik, S. (1993), Misspecification in VAR and their effects on impulse responses and variance decompositions, *Journal of Econometrics*, 59, 319-341.
- [8] Canova, F. (1995), Sensitivity Analysis and Model Evaluation in Simulated Dynamic General Equilibrium Economies, *International Economic Review*, 36, 477-501.
- [9] Canova, F. (2002) Validating monetary DSGE models through VARs, CEPR working paper 3442
- [10] Canova, F. and De Nicoló, G. (2002), Money Matters for Business Cycle Fluctuations in the G7, *Journal of Monetary Economics*, 49, 1131-1159.
- [11] Canova, F. (2005) Structural changes in the US economy, available at www.econ.upf.edu/crei/canova.
- [12] Canova, F. and Gambetti, L. (2007), Do inflation expectations matter? The Great Moderation revisited, manuscript.

- [13] Canova, F. and Pina, J. (2005), What VARs tell us about DSGE models?, in C. Diebolt, and Krystou, C. (eds.), *New Trends In Macroeconomics*, Springer Verlag, New York, NY.
- [14] Canova, F. and Paustian, M (2007) Measurement with theory: using sign restriction to evaluate business cycle models, manuscript.
- [15] Canova, F. and Sala, L. (2005) Back to square one: identification issues in DSGE models, ECB working paper 583.
- [16] Chari, V.V., Kehoe, P. and McGrattan, E. (2006) A critique of SVAR using Real Business Cycle theory, Fed of Minneapolis, working paper 631.
- [17] Chari, V.V., Kehoe, P. and McGrattan, E. (2007) Business cycle accounting, *Econometrica*, 75, 781-836.
- [18] Choi, I and P.C. Phillips (1992) Asymptotic and Finite Sample Distribution Theory for IV Estimators and Tests in Partially Identified Structural Equations, *Journal of Econometrics* 51, 113-150.
- [19] Christiano, L., Eichenbaum, M. and Evans, C. (2005) Nominal Rigidities and the Dynamic Effects of a shock to Monetary Policy, *Journal of Political Economy*, 113, 1-45.
- [20] Cogley, T. and T.J. Sargent (2005), Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. , *Review of Economic Dynamics*, 8, 262-302.
- [21] Cooley, T. and LeRoy, S. (1985), Atheoretical macroeconomics: A critique, *Journal of Monetary Economics*, 16, 283-308.
- [22] Dedola, L. and Neri, S. (2007) What does a technology shock do? A VAR analysis with model-based sign restrictions, *Journal of Monetary Economics*, 54, 512-549.
- [23] Del Negro, M. and Schorfheide, F. (2004), Priors from General Equilibrium Models for VARs, *International Economic Review*, 95, 643-673.
- [24] Del Negro, M. and Schorfheide, F (2005) Policy predictions if the model does not fit, *Journal of the European Economic Association*, 3, 434-443.
- [25] Del Negro, M. and Schorfheide, F (2006), Monetary Policy Analysis with Potentially Misspecified models, NBER working paper 13099.

- [26] Del Negro, M, Schorfheide, F., Smets, F. And R. Wouters (2006) On the fit of New Keynesian models, *Journal of Business and Economic Statistics*, 25, 143-162.
- [27] Faust, J. and Leeper, E. (1997), Do Long Run Restrictions Really Identify Anything?, *Journal of Business and Economic Statistics*, 15, 345-353.
- [28] Fernandez-Villaverde J. and Rubio-Ramirez J. F. (2007), How structural are structural parameters, forthcoming, *NBER Macroeconomic Annual* .
- [29] Fernandez-Villaverde, J., Rubio-Ramirez, J. , Sargent, T. and Watson, M. (2005) The ABC (and D) of structural VARs, NBER working paper.
- [30] Gali, J. and Rabanal, P. (2005) Technology shocks and aggregate fluctuations: How well does the RBC model fit Postwar US data?, *NBER Macroeconomic Annual*, 20.
- [31] Ireland, P. (2004) Technology Shocks in the New Keynesian model, *Review of Economics and Statistics*, 86, 923-936.
- [32] Justiniano, and Primiceri, G. (2006), Time Varying volatilities of Macroeconomics fluctuations, Northwestern University manuscript.
- [33] Liu, T. (1960) Underidentification, Structural estimation, and Forecasting, *Econometrica*, 28, 855-865.
- [34] Lubik, T. and Schorfheide, F. (2004) Testing for indeterminacy: An application to US Monetary Policy, *American Economic Review*, 94, 190-217.
- [35] Ma, A. (2002) GMM estimation of the new Phillips curve, *Economic Letters*, 76, 411-417.
- [36] Nason, J. and Smith, G. (2005) Identifying the New Keynesian Phillips Curve, Federal Reserve of Atlanta, working paper 2005-1.
- [37] Pappa, P. (2005) RBC or New-Keynesian transmission: the labor market effects of government expenditure shocks, CEPR working paper 5313.
- [38] Primiceri, G. (2005), Time Varying Structural VAR and Monetary Policy, *Review of Economic Studies*, 72, 453-472.

- [39] Ravenna, F. (2007) VAR and reduced form representation of DSGE models, *Journal of Monetary Economics*, 54, 2048-2064.
- [40] Rosen, A. (2006) "Confidence sets for partially identified parameters that satisfy a finite number of moment inequalities", UCL working paper.
- [41] Rotemberg, J. and Woodford, M. (1997) , An Optimization based Econometric Framework for the evaluation of Monetary Policy, *NBER Macroeconomic Annual*, 12, 297-346.
- [42] Schorfheide, F. (2000), Loss function based evaluation of DSGE models, *Journal of Applied Econometrics*, 15, 645-670.
- [43] Sims, C. (1980) Macroeconomic and Reality, *Econometrica*, 48, 1-48.
- [44] Sims, C. (1986), Are Forecasting Models Usable for Policy Analysis?, *Federal Reserve Bank of Minneapolis, Quarterly Review*, Winter, 1-16.
- [45] Smets, F. and Wouters, R. (2003) An estimated Dynamic Stochastic General Equilibrium models of the Euro area, *Journal of the European Economic Association*, I, 1123-1175.
- [46] Stock, J. and Wright, J. (2000) GMM with weak identification, *Econometrica*, 68, 1055-1096.
- [47] Stock, J., Wright, J. and Yogo, M. (2002) A survey of weak instruments and weak identification in Generalized methods of moments, *Journal of Business and Economics Statistics*, 20, 518-529.
- [48] Uhlig, H. (2005), What are the Effects of Monetary Policy? Results from an agnostic Identification procedure, *Journal of Monetary Economics*, 52, 381-419.