Back to square one: Identification issues in DSGE models ☆

Fabio Canova [a,*], Luca Sala [b]

[a] Department of Economics, ICREA-UPF, CREI, AMeN and CEPR, Ramon Trias Fargas 25-27, 08008 Barcelona, Spain
[b] Dipartimento di Economia, IGIER and Centro Baffi - Universitá Bocconi, Italy

ARTICLE INFO

ABSTRACT

We investigate identification issues in DSGE models and their consequences for parameter estimation and model evaluation when the objective function measures the distance between estimated and model-based impulse responses. Observational equivalence, partial and weak identification problems are widespread and typically produced by an ill-behaved mapping between the structural parameters and the coefficients of the solution. Different objective functions affect identification and small samples interact with parameters identification. Diagnostics to detect identification deficiencies are provided and applied to a widely used model.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The 1990s have seen a remarkable development in the specification of DSGE models. The literature has added considerable realism to the constructions popular in the 1980s and a number of shocks and frictions have been introduced into first generation RBC models driven by technological disturbances. Steps forward have also been made in comparing models and the data: while 10 years ago it was standard to calibrate the parameters and informally evaluate the quality of their fit, now likelihood-based estimation of the structural parameters is common (see e.g. Smets and Wouters, 2003; Ireland, 2004; Gertler et al., 2008; Canova, forthcoming) and new techniques have been introduced for evaluation purposes (see Del Negro et al., 2006).

Given the complexities involved in estimating state-of-the-art DSGE models and the difficulties in designing criteria that are informative about the reasons for their discrepancy with the data, a strand of literature has considered less demanding limited information methods and focused on whether the model matches the data only along certain dimensions. Following Rotemberg and Woodford (1997) and others, it is now common to estimate structural parameters by quantitatively matching conditional dynamics in response to certain structural shocks (Canova and Paustian, 2007 propose

a limited information strategy where only a qualitative matching of impact responses is sought). One crucial but often neglected condition needed for any methodology to deliver sensible results and meaningful inference is the one of identifiability: the objective function must have a unique extremum at the true parameter vector and display "enough" curvature in all relevant dimensions. Since the mapping between DSGE parameters and the objective function is highly nonlinear and, generally, not available in an analytical form, it is far from straightforward to check in practice whether identification conditions hold or not.

This paper investigates identifiability issues in DSGE models; explores their consequences for parameter estimation and model evaluation and provides diagnostics to detect problems in practice. While the approach we consider falls in the class of minimum distance estimators and some results concerning the interaction between identification and estimation in this class exist in the literature (see, among others, Choi and Phillips, 1992; Stock and Wright, 2000; Rosen, 2006; Kleibergen and Mavroeidis, 2008), to the best of our knowledge we are the first to address these issues in the context of DSGE models and to study the reasons for why they emerge. Our special interest in impulse response matching is motivated by the popularity of the technique among macroeconomists and the fact that a few peculiarities of the procedure make existing theoretical conclusions inapplicable.

Section 2 discusses the generics of identification, provides informal definitions for several practically relevant situations and studies, in the context of a simple example, four commonly encountered problems—observational equivalence; under-identification; partial and weak identification. These problems turn out to be inherent to DSGE models—they are typically produced because the mapping from the structural parameters to the coefficients of the solution is ill-behaved—and standard choices of data summaries and of objective functions may make identification deficiencies worse. Classical maximum likelihood and Bayesian approaches face similar difficulties in dealing with information failures intrinsic to the theory and arbitrarily chosen priors may hide severe identification problems. The common practice of fixing some troublesome parameters to arbitrary (calibrated) values may create distortions in the distribution of the estimates of parameters, making the results difficult to interpret. The section concludes examining the reasons for why identification deficiencies emerge and offering practical remedies to avoid them.

Section 3 investigates the population identification features of the parameters of a version of the Christiano et al. (2005)—Smets and Wouters (2003) model. Researchers may face important difficulties in identifying the parameters of this model and may be led astray in their inferential exercises. The mapping between the structural parameters and the population responses has, in fact, a unique local minimum at the true parameter vector, but it is extremely flat in a large portion of the parameter space and displays ridges in the parameters controlling price and wage stickiness and indexation. Thus, data generating processes (DGP) featuring different economic frictions may be close in terms of conditional dynamics. We demonstrate the extent of this problem showing that the range of structural parameters generating impulse responses close to the true ones (and therefore consistent with small values of objective functions) is large and include specifications with different features and potentially different welfare properties.

Section 4 studies how identification is affected when the sample version of the objective function is available. Curvature deficiencies combined with small samples produce a very ill-behaved mapping from the structural parameters to the estimated impulse responses and the features of this mapping hardly change in larger samples. Hence, when population problems are present, small samples tend to make identification problems especially severe and standard asymptotic approximations become grossly deficient.

Section 5 presents a number of diagnostics to detect identification failures and discusses why standard errors of the estimates do not necessarily convey relevant information about identification deficiencies. We shows that the pathologies discovered in Section 3 exist because the mapping between the structural parameters and the coefficients of the stationary rational expectations solution is ill-behaved. In turn, this occurs because the transition law of the states of the model is relatively insensitive to changes in many parameters. Since the mappings from the coefficients of the solution to the impulse responses and from the impulse responses to the objective function contribute little to information deficiencies, identification problems in this model can be resolved only through extensive reparameterization. Section 6 concludes providing some suggestions for empirical practice.

The problems this paper highlights emerge in standard settings where there is a unique solution to the problem of the representative agent under rational expectations. Therefore, the emphasis here differs from the one of Beyer and Farmer (2004), who consider cases where there may be multiple solutions; of Rosen (2006), where observational equivalence is a feature of the theory (it implies inequality constraints on some observable variables), and of Moon and Schorfheide (2007), where the empirical strategy used to recover structural parameters from the reduced form ones implies observational equivalence.

Christiano et al. (2006), Fernandez-Villaverde et al. (2007) and Chari et al. (2008) have studied invertibility problems in DSGE models and the ability of structural VARs to recover the dynamics induced by shocks. One interpretation of their evidence is that, when these problems are present, the empirical strategy of matching impulse responses is potentially flawed. Our work suggests that, even when these problems are absent, identification deficiencies may make impulse response matching exercises problematic and inference erratic.[1]

---

[1] Replication programs, the results of sensitivity analysis and additional material present in the original working paper (Canova and Sala, 2005) but omitted from the printed version of the paper are available at the JME Science Direct webpage.

2. A few definitions and an example

Identification problems have been extensively studied in theory; the literature on this issue goes back to Koopmans and Reiersol (1950), and more recent contributions include Rothenberg (1971), Pesaran (1981), Hsiao (1983) and Dufour (2003). While theoretical concepts are relatively straightforward, it is uncommon to see these issues explicitly considered in empirical analyses.

To set ideas, identification has to do with the ability to draw inference about the parameters of a theoretical model from an observed sample. Many reasons may prevent researchers from performing such an exercise. First, the mapping between structural parameters and the objective function may not display a unique minimum. In this case, structural models with potentially different economic interpretations may be indistinguishable. This problem is called here *observational equivalence*. Second, the objective function may be independent of certain structural parameters. This occurs, for example, when a parameter disappears from the rational expectations solution or enters only in certain functions of the data. We call this problem *under-identification*. A special case of this phenomenon emerges when two structural parameters enter the objective function only proportionally, making them separately unrecoverable. This problem, well known in traditional systems of linear simultaneous equations, is called here *partial identification*. Third, even though all parameters enter separately the population objective function and the population objective function has a unique minimum, its curvature may be small in certain regions of the parameter space. This phenomenon is named here *weak identification* problem, in analogy with the GMM literature on *weak instruments* (see, among others, Stock et al., 2002).

Clearly, identification problems depend on the location of the true parameters, i.e. they may be localized in a portion of the parameter space or may emerge globally; and on the choice of the objective function. Moreover, they may involve the relationship between the structural parameters and the *population* objective function or the relationship between the structural parameters and the *sample* objective function. The rest of this section concentrates on situations where identification pathologies occur in population and focuses on local problems. We emphasize population issues to highlight the fact that identification deficiencies are rooted in the theory and may appear even with samples of infinite size.

2.1. The role of the objective function

To understand why the choice of the objective function may affect identification, consider the optimality conditions of a DSGE model, which can be written as

$$A_\theta E_t x_{t+1} = H_\theta + B_\theta x_t + C_\theta x_{t-1} + D_\theta E_t z_{t+1} + F_\theta z_t \tag{1}$$
$$z_t = G_\theta z_{t-1} + e_t \tag{2}$$

where z_t is a vector of exogenous variables, x_t includes all the endogenous variables, E_t denotes expectations, conditional on the information set at t, and the matrices $A_\theta, B_\theta, C_\theta, D_\theta, F_\theta, G_\theta$ and H_θ are functions of the structural parameters θ. The unique stable rational expectation solution of (1) and (2) is

$$x_t = K_\theta + W_\theta z_{t-1} + J_\theta x_{t-1} + M_\theta e_t \tag{3}$$
$$z_t = G_\theta z_{t-1} + e_t \tag{4}$$

where $K_\theta, W_\theta, J_\theta$ and M_θ are nonlinear transformations of the matrices in (1)–(2). Letting

$$y_t = S[x_t' \ z_t']' \tag{5}$$

where S is a selection matrix picking the observables from the variables of the model, (3)–(5) represent a state space system. Let $\mu_\theta = S[K_\theta' \ 0']'$ represent the steady state of the observables, where 0 is a conformable zero vector, and $P_\theta = S[M_\theta' \ I]'$. Under normality of e_t, the likelihood function of (5) can be computed with the Kalman filter as

$$L(\theta|y) \propto |R_{t|t-1}|^{-0.5} \exp\left\{ -0.5 \sum_t [y_t - y_{t|t-1}]' R_{t|t-1}^{-1} [y_t - y_{t|t-1}] \right\} \tag{6}$$

where $y_{t|t-1}$ is the best predictor of y_t, given the information up to time $t-1$, and $R_{t|t-1}$ is the covariance matrix of the one-step ahead prediction error. Note that $y_{t|t-1}$ contains information about μ_θ, and $R_{t|t-1}$ information about the variance of the endogenous variables, $P_\theta Var(e_t) P_\theta'$. Thus, the likelihood function provides a natural upper bound to the identification information present in the data, since all features of the model are explicitly accounted for.

Suppose that the objective function measures the distance between a set of nonlinear functions of the data $\hat{f}(y_{1t}, \theta_0)$, where y_{1t} is a subset of y_t, and their population counterpart $f(Y_{1t}, \theta)$ (here, Y_{1t} denotes the random variable and y_{1t} its realization) and that θ is chosen to make this distance close to zero. Then

$$\hat{\theta}^{MD} = \arg\min_{\tilde{\theta}} S(\tilde{\theta}, y) = \|\hat{f}(y_1, \theta_0) - f(Y_1, \tilde{\theta})\|_\Omega \tag{7}$$

is a minimum distance estimator, where Ω is a weighting matrix. A maximum likelihood estimator can also be defined as a minimum distance estimator. In fact

$$\hat{\theta}^{ML} = \arg\min_{\tilde{\theta}} -L(\tilde{\theta}|y) \tag{8}$$

Hence, an objective function which measures the distance between some of the data and the model-based impulse responses may have additional informational deficiencies relative to the likelihood function for three reasons: it disregards information contained in the covariance matrix of the prediction errors, $R_{t|t-1}$ and may use only the information contained in some columns of P_θ; it neglects the steady state relationships, μ_θ; and may consider responses of only a subset of the variables in y_t.

2.2. The mappings from the structural parameters to the sample objective function

The mapping from the data to $\hat{\theta}^{MD}$ involves the interaction of four operators:

 (i) The "solution" mapping, linking the parameters θ to the coefficients of the solution, i.e. the matrices K_θ, W_θ, J_θ and M_θ. This mapping could be poorly behaved because structural parameters disappear from solution, do not have independent variation, for example, because of stability requirements, or induce small changes in the coefficients of the solution.
 (ii) The "moment" mapping, linking the coefficients of the solution to the functions of interest (impulse responses in our case). The problem here could be that the selection of particular impulse responses smears or poorly packages the information contained in the solution.
(iii) The "objective function" mapping, linking the functions of interest to the population objective function. It could happen that this mapping is deficient because it does not have a unique minimum or because the objective function is "insensitive" to changes in the functions of interest.
(iv) The "data" mapping, linking the population and the sample objective functions. This mapping may be ill-behaved when estimated responses are inconsistent estimators of the population ones—and this may occur, for example, because of errors in the identification of shocks—or when a subset of the endogenous variables appearing in the solution is omitted from the estimated VAR.

To illustrate how the first three mappings can induce identification problems, consider the following model:

$$y_t = k_1 + a_1 E_t y_{t+1} + a_2(i_t - E_t \pi_{t+1}) + e_{1t} \tag{9}$$
$$\pi_t = k_2 + a_3 E_t \pi_{t+1} + a_4 y_t + e_{2t} \tag{10}$$
$$i_t = k_3 + a_5 E_t \pi_{t+1} + e_{3t} \tag{11}$$

where y_t is the output gap, π_t the inflation rate, i_t the nominal interest rate and e_{1t}, e_{2t}, e_{3t} are i.i.d. contemporaneously uncorrelated shocks. The first equation is a IS curve, the second a Phillips curve, the third characterizes monetary policy, and k_1, k_2, k_3 are constants. The model is admittedly simple but this feature enables us to compute analytically the solution mapping and explain what additional problems the moment and the objective function mappings may induce. It is easy to verify that the solution for i_t, π_t, y_t is linear in $e_{jt}, j = 1, 2, 3$ and given by

$$\begin{bmatrix} y_t \\ \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & a_2 \\ a_4 & 1 & a_2 a_4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \end{bmatrix} \equiv \mu_\theta + P_\theta e_t \tag{12}$$

where

$$\mu_1 = \frac{k_2 a_2 + -k_2 a_2 a_5 + a_3 k_1 + a_3 a_2 k_3 - k_1 - a_2 k_3}{a_1 - a_3 a_1 + a_2 a_5 a_4 - a_2 a_4 + a_3 - 1}$$
$$\mu_2 = \frac{a_1 k_2 - k_1 a_4 - a_2 k_3 a_4 - k_2}{a_1 - a_3 a_1 + a_2 a_5 a_4 - a_2 a_4 + a_3 - 1}$$
$$\mu_3 = \frac{k_3 a_1 - k_3 a_3 a_1 - a_2 k_3 a_4 + a_3 k_3 - k_3 - a_5 k_1 a_4 + a_5 a_1 k_2 - k_2 a_5}{a_1 - a_3 a_1 + a_2 a_5 a_4 - a_2 a_4 + a_3 - 1}$$

Note that the example is rigged so that (12) has a clear empirical counterpart and the shocks e_{1t}, e_{2t}, e_{3t} can be recovered from the data using a simple triangularization of the reduced form shocks. We have done this to keep the issue of shock identification, which is tangential to the problems discussed here, separate from the issue of parameter identification.

2.3. The problems

A number of useful points can be made from Eq. (12). First, the parameters (a_1, a_3, a_5) disappear from the solution and therefore from the impulse responses. Interestingly, they are those characterizing the forward looking components of the

model. Hence, if the μ's are disregarded, observational equivalence could be widespread. For example, a model where information is sticky, that is, a version of the model where expectations are functions of lagged information, has the same dynamics as the basic one and models with determinate ($a_5 > 1$) or indeterminate ($a_5 < 1$) features may not distinguished from the dynamics of the model (see Canova and Gambetti, 2007, for a general case).

Second, different shocks carry different information for the remaining parameters: for example, responses to e_{1t} allow us to recover only a_4; while responses to e_{2t} have no information for either a_4 or a_2—they do not independently move any endogenous variable of the system. Similarly, responses of the inflation rate to the three shocks carry different information about the structural parameters than, say, responses of the nominal rate. Clearly, even considering responses to all shocks, a set of parameters remains under-identified.

Third, partial identification problems may emerge: simultaneously increasing a_2 and decreasing a_4 by the right amount leaves the objective function roughly unchanged. It is also easy to verify that for low values of a_2 the identification of a_4 will be weak. The contours of the distance function computed from the responses to e_{3t} are shown in Fig. 1, first row. In the left panel the true a_2 and a_4 are set to 0.2. Here, since the curvature of the objective function is small, any value of a_4 between 0 and 0.9 falls within the 0.01 distance contour. If a_2 were smaller (see right panel), even larger values of a_4 could fall into this contour.

Fourth, steady state information (the constants μ_1, μ_2, μ_3) helps in identifying some of the dynamic parameters. This can be seen in the second row, left panel of Fig. 1: the likelihood function computed assuming normality of the shocks is much more bell-shaped and better behaved in (a_2, a_4) than the distance function.

2.4. Two common but problematic solutions

A popular strategy employed when identification problems are present is to calibrate some of the parameters and estimate the others, conditional on the calibrated values. Calibrated values are generally obtained from micro-estimates, cross-country estimates, or educated guess-estimates researchers make. If the calibrated parameters are under-identified (as is the case, for example, for a_1 in the model), the distance function, and therefore estimates of the remaining parameters, will be unaffected. If, on the contrary, the calibrated parameters are partially identifiable, it is easy to build examples where improper settings will make the distance function shift, biasing estimates of the remaining parameters. We show this in the second row, right panel of Fig. 1, using the likelihood function. Here k_2 is set to 1.2, when the true value is 1.0. Such an apparently minor miscalibration induces significant changes in the shape and in the height of the likelihood function for a_2 and a_4 and the maximum for a_2 shifts toward a much lower value. Since an improper setting of partially identified parameters may lead to uninterpretable estimates of parameters originally free of identification problems, care should be exercised in interpreting estimates obtained with mixed calibration–estimation approaches and robustness checks extensively performed to verify the robustness of substantive economic conclusions to variations in the calibrated parameters.

Because the likelihood function of DSGE models is often ill-behaved, it has became common to employ Bayesian methods to estimate structural parameters. Given the recent emphasis, it is useful to discuss identification issues in Bayesian frameworks. The posterior distribution is proportional to the likelihood times the prior. If the parameter space is variation free, that is, there are no implicit constraints on combinations of parameters, the likelihood carries information for the parameters if the prior and the posterior have different features (see Poirier, 1998). When this is not the case, there is a simple diagnostic to detect identification failures. If prior information becomes progressively more diffuse, the posterior of the parameters with doubtful identification features will also become progressively more diffuse—there is no problem in eliciting diffuse prior distributions here, since the parameterization of a DSGE model is given.

When the parameter space is not variation free, e.g. because stability conditions or non-negativity constraints based on economic theory are imposed, the prior of non-identified parameters may be marginally updated because these restrictions shift the posterior away from the prior, even when the likelihood is flat. Thus, finding that prior and posterior differ, does not guarantee that parameters are identified. Nevertheless, also in this case, a sequence of prior distributions with increasing spreads may help detecting potential identification problems.

When identification failures are due to deficiencies in the shape of the population objective function, classical and Bayesian methods face similar difficulties. Prior restrictions may help to resolve sample identification problems, as long as the prior effectively summarizes information external to the data used, but an improper use of prior restrictions may cover up population pathologies. In fact, the combination of tightly specified priors and/or auxiliary restrictions on the parameter space can produce a well behaved posterior, even when the population objective function is poorly behaved. We show this fact in the last row of Fig. 1, which plots the likelihood function and the posterior distribution in k_3 and a_3, the latter obtained with a tightly specified prior, centered at the true values of the parameters. The ridge visible in the likelihood function has disappeared from the posterior.

2.5. Summary and remedies

What does one learn from this simple example? First, the dynamics of the model may fail to contain information about certain parameters. Second, matching responses to a limited number of shocks or choosing a subset of the responses to all
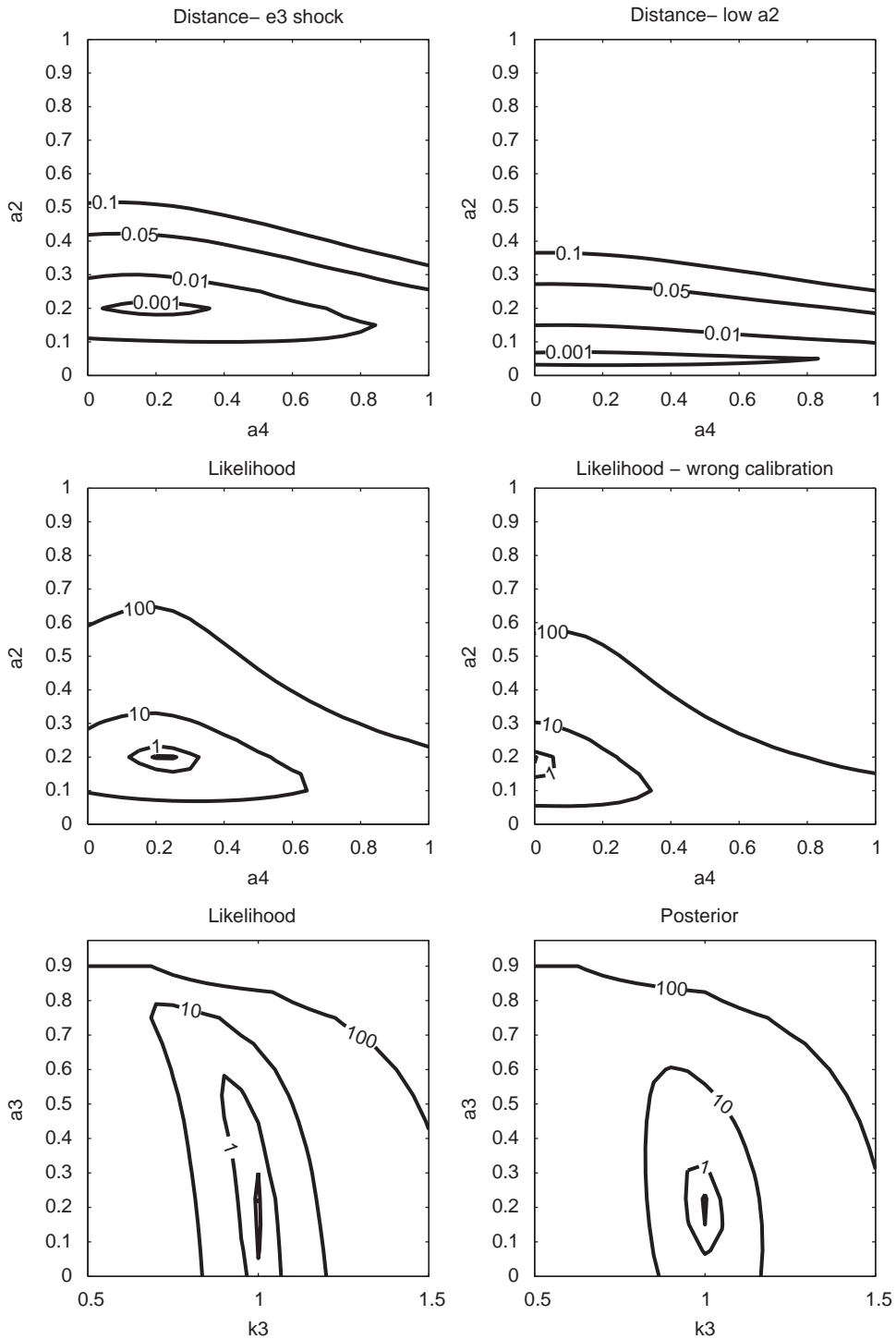
Fig. 1. Distance function, likelihood and posterior.

shocks may exacerbate identification problems. Since the choice of what shock to consider and of what variable response to look at is typically made disregarding the identification power of the different shock-variable combinations, identification pathologies are likely to be widespread. Third, since structural parameters enter nonlinearly in the coefficients of the solution, it is not hard to encounter situations where partial and weak identification problems emerge. Fourth, while appropriately choosing the objective function may reduce identification difficulties, there is no guarantee that it will solve them. Fifth, mixed calibration–estimation exercises may bias inference. Sixth, priors may help to reduce sample

identification problems but fail to resolve problems related to poorly conditioned population mappings. In fact, uncritical use of Bayesian methods, including employing prior distributions which do not truly reflect spread uncertainty, may hide identification pathologies.

How can one remedy these problems? The parameters on forward looking variables disappear from the solution because the shocks are i.i.d. and the model features no internal dynamics. Therefore, if interest centers in estimating these parameters, and one insists on using only the dynamics for that purpose, one should introduce some form of sluggishness in the model or allow the disturbances to be serially correlated. In the case both are used, care should be exercised because it is well known since at least Sargent (1978) that the parameters regulating the internal and the external propagation mechanisms are difficult to separately identify. In general, the restrictions implied by the steady states or the covariance of the shocks should always be used to identify the parameters of interest. Clearly, even if all model information is employed, partial and weak identification problems may remain. In this case, additional theoretical restrictions or changes in the parameterization of the model are necessary.

To conclude, we would like to stress that the analysis of identification issues was straightforward in this example because the mapping between DSGE parameters and the coefficients of the solution was available analytically. In more general situations, detection of identification problems is extremely difficult. Diagnostics that help in this enterprise are in Section 5.

3. Population identification issues in a prototype DSGE model

The problems highlighted in Section 2 indicate that structural inference may be difficult, even when the true model is known. When it is unknown, and identification problems present, researchers may be misled into believing that certain features characterize the data, when they are in fact absent. Since the literature has added frictions in standard DSGE models to enhance their fit without caring too much about their identifiability, we dedicate this section to investigate (i) whether models with different frictions could have similar population responses and (ii) whether it is possible to obtain meaningful estimates of parameters that are in fact absent from the DGP.

3.1. The model

The economy we consider is standard: it features real and nominal frictions, and has been shown to capture well important features both of the US (see Christiano et al., 2005; Dedola and Neri, 2007) and the EU economies (see Smets and Wouters, 2003). The log-linearized optimality conditions of the model are in Table 1 together with the definition of the variables. Eq. (T.1) describes capital accumulation; Eq. (T.2) determines capacity utilization; Eq. (T.3) is the resource constraint; Eq. (T.4) represents the monetary policy rule; Eq. (T.5) represents the production function; Eq. (T.6) is a labor demand equation, Eq. (T.7) is an Euler equation for consumption, Eq. (T.8) is an Euler equation for investment, Eq. (T.9) describes the dynamics of Tobin's q; finally, Eqs. (T.10) and (T.11) represent the wage setting and the price setting equations. Labels and definitions of the parameters are in Table 2.

This model is sufficiently rich that it is difficult to know *a priori* which parameters are identifiable from the conditional dynamics. Since the analytical mapping between structural parameters and the objective function is not available, we explore its features by constructing population responses using the posterior mean estimates of Dedola and Neri (2007) (see Table 2) and examining the shape of the distance function in a neighborhood of the true parameter vector. The distance function is constructed using 20 equally weighted responses of inflation, the nominal interest rate, investment, output, hours, capacity utilization to monetary policy shocks. Varying the number of horizons and of variables used in the objective function has virtually no influence on the results we present.

The vector of structural parameters θ is divided into $[\theta_1, \theta_2]$ where $\theta_1 = [\phi, \nu, h, \eta, \chi, \psi, \gamma_p, \gamma_w, \zeta_p, \zeta_w, \varepsilon_w, \lambda_r, \lambda_\pi, \lambda_y]$; and $\theta_2 = [\beta, \delta, \bar{\pi}, \rho_z, \rho_x, \rho_b, \sigma_z, \sigma_r, \sigma_p, \sigma_w, \sigma_x, \sigma_b]$. In the analysis, θ_2 is treated as fixed. The standard errors of the shocks $(\sigma_z, \sigma_r, \sigma_p, \sigma_w, \sigma_x, \sigma_b)$ and the persistence parameters (ρ_z, ρ_x, ρ_b), belong to θ_2 because they are non-identifiable from the responses to monetary shocks. On the other hand, β and δ belong to θ_2 since they are typically estimated from steady state relationships. If they were obtained from impulse responses, the problems we highlight below would be even more pronounced.

3.2. Some graphical evidence

Fig. 2 plots the elasticity of (one plus) the distance function with respect to each of the parameters in θ_1 (vertical axis) against the parameter values (horizontal axis). To calculate the elasticity we vary one parameter over the range on the horizontal axis, keeping all other parameters fixed at their true values. True values for each parameter are on the top of each box. Elasticities are preferable to any other summary of the slope of the objective function since they are unit free measures. Plots of the objective function against the parameters could be difficult to interpret since scaling matters (if the distance is multiplied by 1000, the curvature would look different without affecting the nature of the problem). The distance function has a unique minimum at the true parameter vector (indicated by the star in each box) but variations in the distance function in a neighborhood of the true vector are generally very small. For example, the elasticities with

Table 1
(a) The variables of the model; (b) the equations of the model

Label	Definition
(a)	
i_t	Investment
u_t	Capacity utilization
r_t	Real rate
c_t	Consumption
y_t	output
R_t	Nominal rate
K_t	Capital
N_t	Total hours
w_t	Real wage rate
π_t	Inflation rate
q_t	Tobin's q

(b)

$$k_{t+1} = (1 - \delta)k_t + \delta i_t \tag{T.1}$$

$$u_t = \psi r_t \tag{T.2}$$

$$0 = \frac{\eta\delta}{\bar{r}}i_t + \left(1 - \frac{\eta\delta}{\bar{r}}\right)c_t - \eta k_t - (1 - \eta)N_t - \eta u_t - e_t^z \tag{T.3}$$

$$R_t = \lambda_r R_{t-1} + (1 - \lambda_r)(\lambda_\pi \pi_t + \lambda_y y_t) + e_t^r \tag{T.4}$$

$$y_t = \eta k_t + (1 - \eta)N_t + \eta u_t + e_t^z \tag{T.5}$$

$$N_t = k_t - w_t + (1 + \psi)r_t \tag{T.6}$$

$$c_t = \frac{h}{1+h}E_t c_{t+1} + \frac{h}{1+h}c_{t-1} - \frac{1-h}{(1+h)\phi}(R_t - E_t\pi_{t+1}) + \frac{1-h}{\phi}[e_t^b - E_t e_{t+1}^b] \tag{T.7}$$

$$i_t = \frac{\beta}{1+\beta}E_t i_{t+1} + \frac{1}{1+\beta}x_{t-1} + \frac{\chi^{-1}}{1+\beta}q_t + \frac{\beta}{1+\beta}E_t x_{t+1} - \frac{1}{1+\beta}e_t^x \tag{T.8}$$

$$q_t = E_t\pi_{t+1} - R_t + \beta(1-\delta)E_t q_{t+1} + \beta\bar{r}E_t r_{t+1} \tag{T.9}$$

$$\pi_t = \frac{\beta}{1+\beta\gamma_p}E_t\pi_{t+1} + \frac{\gamma_p}{1+\beta\gamma_p}\pi_{t-1} + T_p(\eta r_t + (1-\eta)w_t - e_t^z + e_t^p) \tag{T.10}$$

$$(1 + T_w)w_t = \frac{1}{1+\beta}w_{t-1} + \frac{\beta}{1+\beta}(E_t\pi_{t+1} + E_t w_{t+1}) - \frac{1+\beta\gamma_w}{1+\beta}\pi_t + \frac{\gamma_w}{1+\beta\gamma_w}\pi_{t-1} + T_w\left[\nu N_t + \frac{\varphi}{1-h}(c_t - hc_{t-1}) + e_t^w\right] \tag{T.11}$$

The six disturbances are: technology shock (e_t^z); monetary policy shock (e_t^r); investment shock (e_t^x); price markup shock (e_t^p); wage markup shock (e_t^w); preference shock (e_t^b). The compound parameters in Eqs. (T.10) and (T.11) are defined as: $T_p \equiv (1 - \beta\zeta_p)(1 - \zeta_p)/(1 + \beta\gamma_p)\zeta_p$ and $T_w \equiv (1 - \beta\zeta_w)(1 - \zeta_w)/(1 + \beta)(1 + (1 + \varepsilon_w)\nu\varepsilon_w^{-1})\zeta_w$.

respect to η, h, χ, ϕ, ν, $\psi\gamma_w$, ε_w, λ_y, λ_π, are of the order of 10^{-5} or smaller and those with respect to γ_p, ζ_w of the order of 10^{-4}. ζ_p is the only parameter for which variations in the baseline value induce sufficiently large variations in the distance function, at least for extreme values of this parameter.

3.3. The size of the weak identification region

Armed with this preliminary evidence, we want to determine how large is the region of "weakly identified" parameterizations, that is, the region of the parameter space in which vectors with different economic implications deliver *small* deviations from the true model. The precise meaning of small will be made clear below.

We draw 100 000 parameter vectors uniformly from the intervals presented on the horizontal axis in Fig. 2, solve the model, compute impulse responses and measure their distance from the benchmark impulse responses obtained with the true parameter vector. We then construct the distribution of the distance function and select those draws whose distance falls in the 0.1 percentile of the distribution of objective functions (i.e. only parameter configurations that produce values for the objective functions in a small neighbor around zero are selected). Five experiments are considered: a benchmark specification, where all features of the model are present; and three alternative ones, where either stickiness or indexation in prices or wages is eliminated from the data generating process. In these four experiments, responses to monetary shocks are used to construct the distance function. In the fifth experiment all the features of the model are present, but responses to both monetary and technology shocks are used to construct the distance function.

Table 2
Description of the parameters of the model and benchmark values.

Label	Definition	Benchmark value
ψ	Sensitivity of capacity utilization to the real rate	0.564
$\bar{r}$	Steady state rental rate of capital	0.0271
λ_r	Interest rate smoothing	0.779
λ_π	Response of monetary policy to inflation	1.454
λ_y	Response of monetary policy to output	0.234
η	Capital share	0.209
h	Habit persistence	0.448
ϕ	Risk aversion	3.014
χ	Inverse of investment's elasticity to Tobin's q	6.300
γ_p	Price indexation	0.862
γ_w	Wage indexation	0.221
ζ_p	Price stickiness	0.887
ζ_w	Wage stickiness	0.620
v	Inverse elasticity of labor supply	2.145
ε_w	Steady state wage markup	1.2
δ	Depreciation rate (F)	0.0182
β	Discount factor (F)	0.991
$\bar{\pi}$	Steady-state gross inflation (F)	1.016
ρ_z	Technology shock persistence (F)	0.997
ρ_b	Preference shock persistence (F)	0.972
ρ_x	Investment shock persistence (F)	0.522
σ_p	Price mark-up shock st. dev. (F)	0.221
σ_w	Wage mark-up shock st. dev. (F)	0.253
σ_z	Technology shock st. dev. (F)	0.0064
σ_x	Investment shock st. dev. (F)	0.557
σ_r	Monetary policy shock st. dev. (F)	0.0026
σ_b	Preference shock st. dev. (F)	0.066

Parameters followed by (F) are treated as known during simulation and estimation exercises.

If the model is well identified, the intervals for the parameters should be small and concentrated around the true value—only values close to the true ones should be compatible with small distances. If, on the other hand, weak and partial identification problems are present, the intervals need not be centered at the true value and could be large. Finally, if observational equivalence problems are relevant, the intervals could include not only the true parameters of the data generating process, but also the true parameters of alternative specifications. Table 3 reports the results of the experiments: each box presents the true parameter vector and summarizes the content of the intervals for ζ_p, γ_p, ζ_w and γ_w reporting their minimum, their median and their maximum value. Fig. 3 concentrates on the benchmark specification (case 1) and shows the true impulse responses together with the impulse responses generated by parameter vectors whose distance falls within the 0.1 percentile of the simulated distribution.

Table 3 displays three important features. First, in case 1, the intervals are very wide, almost any admissible value of ζ_p, γ_p, ζ_w, γ_w is compatible with a small distance and the impulse responses that these vectors produce are close to each other (see Fig. 3). Thus, with this DGP, it will be very hard to pin down how important stickiness and indexation are. Second, when the DGP eliminates some of the frictions, responses to monetary shocks have a hard time to distinguish what friction matters. When price stickiness (case 2) or price indexation (case 3) are absent, the intervals are practically indistinguishable from those of case 1, even though the median value changes, and jointly setting wage stickiness and price indexation to zero (case 4) still does not help to identify the wage indexation parameter. Third, problems do not necessarily disappear when more shocks are used to construct responses. In fact, the intervals for case 5 are almost identical those of case 1.

Why do these results obtain? To answer this question it is useful to study how the distance function vary with respect to ζ_p, γ_p, ζ_w and γ_w, only. Therefore, we repeat the experiment conducted in case 1, keeping all parameters, but the four controlling price and wage stickiness and indexation, at their true value. Fig. 4 plots parameter pairs generating small distance functions: there is a strong negative relation between price and wage stickiness and between price and wage indexation and the wage indexation parameter is nearly non-identified, suggesting that there is a multi-dimensional ridge in the function. Thus, even if the objective function displays a unique local minimum at the true parameter value, there are severe weak and partial identification problems that may lead researchers to confuse models with different features.

3.4. Summary

Since important debates in the academic and policy literature focus on the identification of the frictions that are most relevant to characterize cyclical fluctuations, and since several parameters of the model are so poorly identified from population responses that it is hard to distinguish in theory situations when some frictions are present from others when
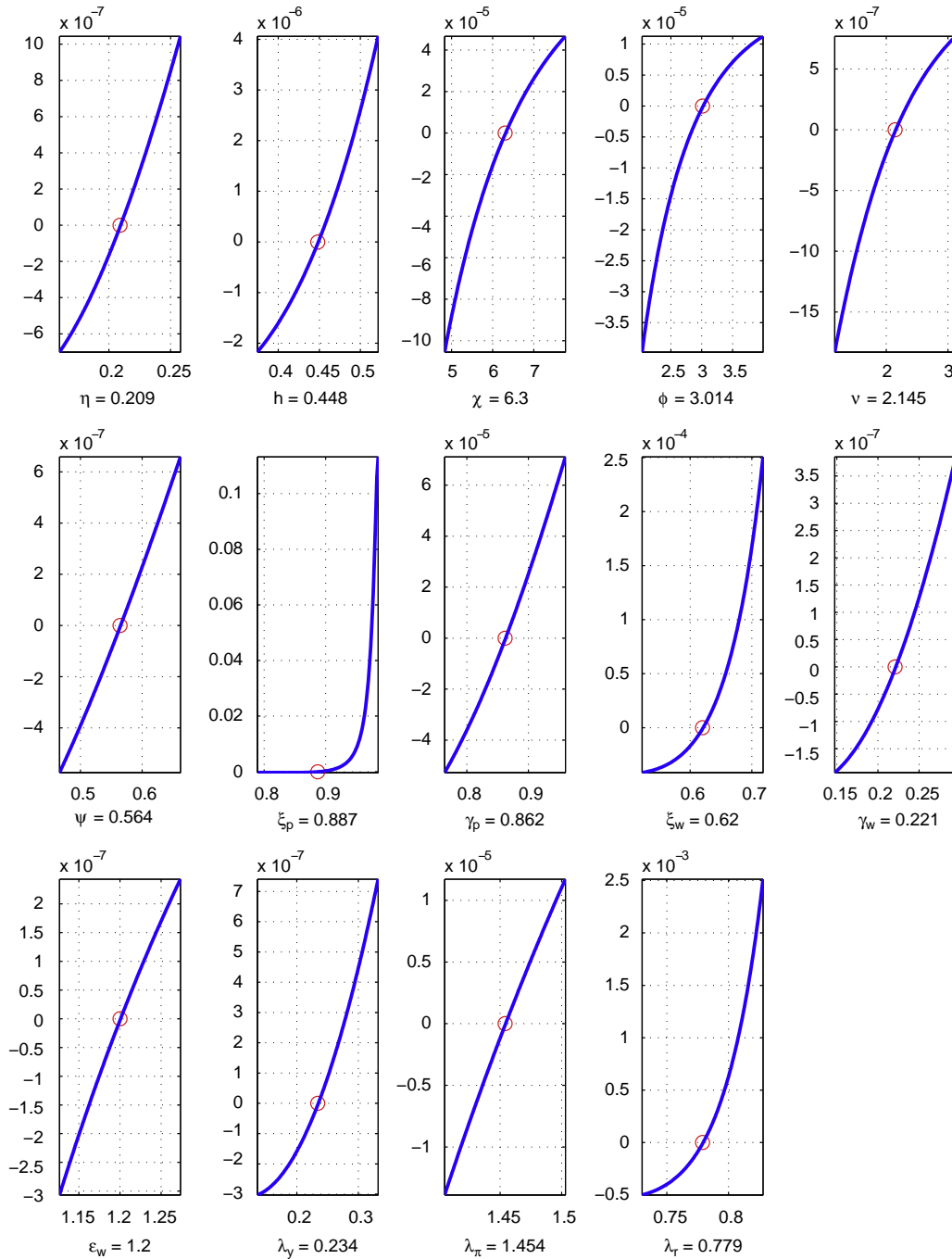
Fig. 2. Elasticities of objective function with respect to the parameters.

they are absent, it is clear that these deficiencies have a crucial impact on the conclusions one reaches. For forecasting, these informational failures may not be important: near-observationally equivalent specifications will forecast equally well. However, knowing, e.g., if price indexation is important or not, has important consequences for the choice of policies. Furthermore, welfare analysis can be easily distorted. For example, using the median estimates obtained in case 1, and equally weighting the variability of output and inflation in a quadratic loss function, welfare turns out to be about twice as worse with the estimated parameters than with the true ones (−0.0011 vs. −0.0005).

In general, in models where partial and weak identification deficiencies are present, one needs to bring information external to the dynamics, as it is done in Christiano et al. (2005), to be able to interpret estimates. It then becomes crucial where this information comes from and whether it is credible or not.

Table 3
Population intervals producing small values of the objective functions.

	ζ_p	γ_p	ζ_w	γ_w
Case 1	0.887	0.862	0.620	0.221
Max.	0.979	0.910	0.884	0.980
Median	0.845	0.515	0.641	0.657
Min.	0.034	0.015	0.028	0.009
Case 2	0.001	0.862	0.620	0.800
Max.	0.843	0.989	0.986	0.987
Median	0.264	0.406	0.906	0.563
Min.	0.046	0.001	0.112	0.008
Case 3	0.887	0.001	0.620	0.221
Max.	0.987	0.987	0.943	0.982
Median	0.888	0.383	0.766	0.428
Min.	0.057	0.003	0.004	0.009
Case 4	0.887	0.001	0.001	0.221
Max.	0.986	0.977	0.788	0.972
Median	0.931	0.395	0.248	0.479
Min.	0.749	0.002	0.007	0.023
Case 5	0.887	0.862	0.620	0.221
Max.	0.989	0.989	0.986	0.987
Median	0.873	0.406	0.906	0.563
Min.	0.046	0.001	0.122	0.008

The first row of each block presents the true parameter values. In the other three rows the minimum, the median and the maximum of the parameters generating values for the objective function which are in the 0.1 percentile of the distribution around zero are reported. Percentiles are obtained by uniformly randomizing the parameters within the ranges presented in Fig. 2. In cases 1–4 only monetary shocks are used; in case 5 monetary and technology shocks are used to construct the distance function. ζ_p (ζ_w) is the price (wage) stickiness parameter; γ_p (γ_w) is the price (wage) indexation parameter.

4. Identification and the sample objective function

This section examines the distortions added to parameter estimates in real life situations, where the sample, rather than the population distance function, is available. While there is evidence in the literature regarding the properties of minimum distance estimators when small samples induce information deficiencies, to the best of our knowledge, no one has investigated the properties of these estimators when population deficiencies are present.

We are primarily interested in two issues: (a) quantifying the distortions that population problems generate when samples of the size typically used in macroeconomics are employed and (b) highlight the properties of the estimates of parameters with and without problematic population identification features. The analysis focuses on the responses to monetary policy shocks, since these are the disturbances that have received most attention in the literature. We simulate 500 time series of size $T = 80, 160, 500, 5000$ from the solution, estimate an unrestricted six-variable VAR with six lags, identify the monetary policy shock, compute impulse responses and bootstrap confidence bands. The lag length of the model is chosen to make sure that population impulse responses are recoverable, thus avoiding distortions due to the potential non-invertibility of the small scale VAR employed. The estimated response uncertainty at each horizon is used to build a diagonal weighting matrix $\hat{\Omega}$, whose entries are inversely proportional to the uncertainty in the estimates. While this is not the optimal choice, it is standard in the literature (see Christiano et al., 2005). We give the routine the best chance to recover the true parameters by correctly identifying monetary shocks—a column of the covariance matrix of VAR residuals is rotated until the theoretical contemporaneous impacts of the monetary shocks is found—and by starting estimation always at the true parameter vector. Table 4 presents a summary of the estimation results: for each T, it reports the true parameter vector, the mean estimates, the standard errors, the 5% and 95% percentiles of the estimated distribution and the vector of average percentage biases. Standard errors are computed numerically across simulations. The entries of the table do not depend on the minimization algorithm—we have checked this using a number of local Matlab routines and a global minimizer, such as simulated annealing—or on the nature of the objective function—estimation results are independent of whether the log or the level of the function is minimized.

4.1. The results

The estimation outcomes are generally poor. Nevertheless, a few features of Table 4 deserve further discussion. First, the mean estimate of $\eta, \nu, \psi, \zeta_w, \gamma_w, \varepsilon_w, \lambda_\pi$ does not depend much on the sample size. Second, the standard errors and the biases
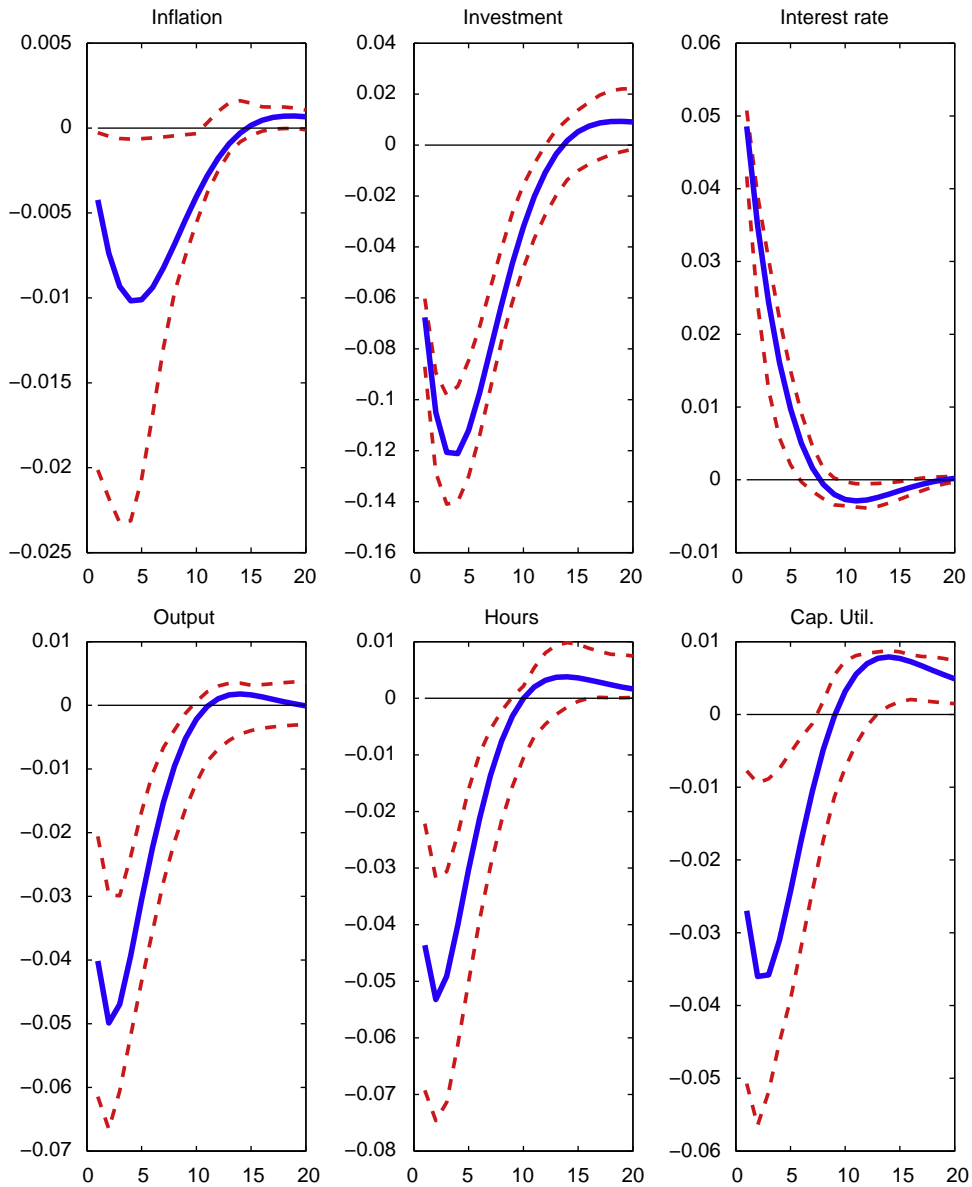
Fig. 3. Impulse responses for parameter configurations generating distances within the 0.1 percentile of the distribution of the distance functions. The bold lines represent responses obtained with parameterization "case 1" in Table 2. The region of impulse responses falling within the bounds is delimited by the dashed lines.

decrease somewhat with the sample size but, even with $T = 500$, they are far from negligible for many parameters. Moreover, increasing the sample to $T = 5000$ makes little difference for their magnitude. Even more interesting is to look at the distribution of parameter estimates (see Fig. 5 for $T = 500$): many are bimodal with peaks at the boundaries of the parameter space. The constraints imposed to ensure the existence of a unique solution or the economic interpretability of the results make standard asymptotic approximations unreliable (this is why we reported numerical standard errors in the table). Third, as one can see in Fig. 6, which plots the empirical distributions of VAR-based responses (dashed lines) and of model-based responses (solid lines) computed again for $T = 500$, responses computed with improbable parameter values lie within the confidence bands of the VAR estimates. Hence, the practice of showing that model's responses, computed using the estimated parameters, are within the confidence bands of responses estimated from the data is uninformative when identification problems exist. Finally, mean estimates are not only statistically but also economically different from those of the true DGP. For example, they typically imply a much stronger response of interest rates to both inflation and output and a lower labor supply elasticity. There is no possibility to detect such interpretation problems by looking at the estimated responses or at the magnitude of the minimized objective function.
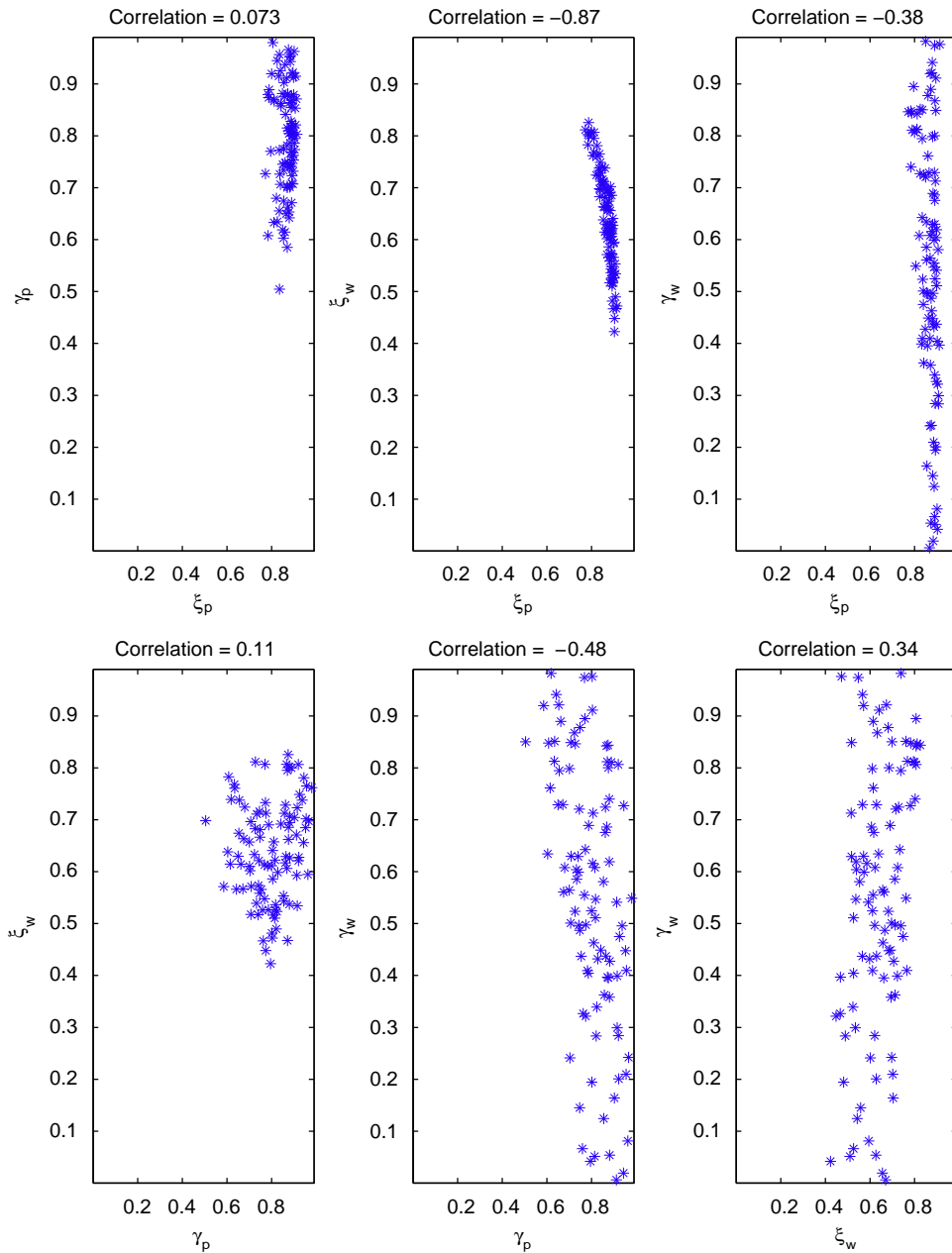
Fig. 4. Bivariate regions for $(\zeta_p, \zeta_w, \gamma_p, \gamma_w)$ producing distances within the 0.1 percentile of the distribution of the distance functions.

While not very encouraging, the results of Table 4 are somewhat optimistic about the performance of impulse response matching estimators when population informational deficiencies are important. They are optimistic because the lag length of the VAR is appropriately selected; the sign and size of the contemporaneous impact of monetary shocks correctly identified; and estimation started at the true values. Had we started from random initial conditions and/or added identification and lag length uncertainty, the distortions would have been larger.

4.2. Are there alternatives?

Given the results of Table 4, one may wonder whether there are estimation approaches which are robust to the type of identification problems we consider. Impulse response matching is a special case of a general class of extremum estimators (see Newey and McFadden, 1994) and, under certain conditions, asymptotic methods to compute estimates which are robust to certain types of identification problems exist (see, among others, Stock and Wright, 2000; Kleibergen and

Table 4
Small sample estimates of the parameters: monetary policy shocks.

True	$T = 80$		$T = 160$		$T = 500$		$T = 5000$	
	Mean [5% 95%]	Bias, std. err.	Mean [5% 95%]	Bias, std. err.	Mean [5% 95%]	Bias, std. err.	Mean [5% 95%]	Bias, std. err.
$\eta = 0.209$	0.16 [0.10 0.49]	−21.7, 0.14	0.18 [0.10 0.54]	−15.57, 0.15	0.20 [0.10 0.67]	−3.27, 0.17	0.18 [0.10 0.40]	−13.12, 0.10
$h = 0.448$	0.57 [0.20 0.80]	25.75, 0.25	0.53 [0.20 0.80]	18.9, 0.24	0.50 [0.20 0.80]	12.3, 0.19	0.46 [0.36 0.57]	1.89, 0.06
$\chi = 6.30$	3.1 [2 9.56]	−50.86, 2.23	4.11 [2 10]	−34.69, 2.54	5.61 [2.46 10]	−10.9, 2.44	6.53 [4.91 8.46]	3.65, 1.03
$\phi = 3.014$	2.34 [1 6]	−22.34, 1.80	2.58 [1 6]	−14.38, 1.68	2.9 [1 6]	−3.9, 1.49	2.96 [2.07 3.73]	−1.7, 0.48
$v = 2.145$	3.83 [0.50 5]	78.32, 1.85	3.36 [0.50 5]	56.6, 2.08	3.08 [0.50 5]	43.5, 2.11	2.51 [0.50 5]	16.95, 1.81
$\psi = 0.564$	0.54 [0.10 0.80]	−3.70 0.27	0.61 [0.10 0.80]	7.44, 0.27	0.60 [0.10 0.80]	6.13, 0.28	0.58 [0.13 0.80]	3.31, 0.25
$\zeta_p = 0.887$	0.85 [0.70 0.97]	−4.16, 0.09	0.86 [0.72 0.95]	−2.76, 0.07	0.88 [0.81 0.95]	−0.49, 0.05	0.89 [0.86 0.94]	0.25, 0.03
$\gamma_p = 0.862$	0.90 [0.11 0.99]	3.89, 0.26	0.91 [0.28 0.99]	5.32, 0.22	0.90 [0.57 0.99]	4.72, 0.15	0.90 [0.79 0.99]	3.87, 0.07
$\zeta_w = 0.620$	0.50 [0.01 0.95]	−19.05, 0.27	0.55 [0.01 0.90]	−10.83, 0.28	0.56 [0.03 0.84]	−9.08, 0.26	0.56 [0.15 0.75]	−9.43, 0.20
$\gamma_w = 0.221$	0.83 [0.01 0.99]	273.39, 0.35	0.80 [0.01 0.99]	260.75, 0.37	0.69 [0.01 0.99]	210, 0.43	0.68 [0.01 0.99]	209.88, 0.38
$\varepsilon_w = 1.2$	1.63 [0.40 2]	36.09, 0.61	1.64 [0.40 2]	37.08, 0.61	1.7 [0.40 2]	41.4, 0.54	1.36 [0.48 2]	12.95, 0.60
$\lambda_y = 0.234$	0.30 [0.01 0.50]	26.44, 0.23	0.36 [0.01 0.50]	55.56, 0.20	0.38 [0.01 0.50]	61.5, 0.17	0.27 [0.13 0.45]	15.94, 0.09
$\lambda_\pi = 1.454$	2.83 [1.01 5]	94.49, 1.82	2.26 [1.01 5]	54.95, 1.62	1.76 [1.01 5]	21, 1.13	1.35 [1.01 2.05]	−7.26, 0.36
$\lambda_r = 0.779$	0.72 [0.40 0.92]	−7.36, 0.16	0.75 [0.56 0.88]	−3.75, 0.10	0.77 [0.69 0.85]	−0.94 ,0.05	0.77 [0.74 0.81]	−1.05, 0.02

T is the sample size. Mean, 90th percentile, standard errors and biases are obtained using 500 replications of each experiment. η is capital share; h controls habit persistence; χ is the inverse of the elasticity of investment with respect to Tobin's q; ψ is the sensitivity of capacity utilization to interest rate; ϕ is the risk aversion coefficient; v is the inverse elasticity of labor supply; ζ_p (ζ_w) is the price (wage) stickiness parameter; γ_p (γ_w) is the price (wage) indexation parameter; ε_w is steady state wage markup; $\lambda_r, \lambda_\pi, \lambda_y$ are monetary policy parameters.

Mavroeidis, 2008). Unfortunately, this alternative estimation theory does not apply here—it would require the use of the continuously updating estimator for the weighting matrix (Hansen et al., 1996), an unlikely choice in applications of the impulse response matching methodology. To shed some light on the issue, we employ a method similar in spirit to the one proposed by Rosen (2006). We first simulate the distribution of the distance function $S(\hat{\kappa}(\omega_0), \omega_0)$ at the minimum for $T = 500$, concentrating out 10 of the 14 parameters in θ_1 (and collecting them in the vector ω), where

$$S(\hat{\kappa}(\omega_0), \omega_0) = \arg \min_{\tilde{\kappa}} \|\hat{f}(y_t, [\kappa_0 \quad \omega_0]) - f(Y, [\tilde{\kappa}(\omega_0), \omega_0])\|_\Omega.$$

$\kappa = [\zeta_p, \gamma_p, \zeta_w, \gamma_w]$ contains the parameters of interest and ω_0 are true values for the remaining parameters. The information contained in the objective function for one simulated set of data is then inverted and the values κ producing an objective function that does not exceed the 5% critical value of the simulated distribution are retained. The resulting sets are similar to the S-sets proposed by Stock and Wright (2000), but are obtained from an empirical, rather than from an asymptotic distribution. In Table 5 a few percentiles of the marginal distribution for each element in κ are presented. These intervals are a bit tighter than those in Table 4, but still quite large. The most problematic parameters are γ_p and γ_w. For the latter, values ranging from 0 to 1 are indistinguishable with 500 observations. Hence, substantial biases and erratic inference are likely to be produced even with this most robust approach.

4.3. Conclusion

Identification deficiencies create considerable inferential problems. Estimates are biased; the biases do not disappear as the sample grows; consistency is hard to obtain, and (asymptotic) standard errors are uninterpretable. But, perhaps more importantly, when weak and partial identification problems are present in population, wrong economic interpretations and wrong policy advice are possible, even when the model correctly represents the DGP of the data.

5. Detecting the source of the problems

Applied investigators can use a number of diagnostics to detect identification problems and to understand what features of the model economy may be responsible for them. For example, versions of the graphical and the simulation analyses employed in Section 3 can give useful hints about parameters with population identification problems. This analysis should be routinely performed prior to estimation in a neighborhood of reasonably calibrated parameter values, much in the same spirit of Bayesian prior predictive analysis (Geweke, 2005, p. 262), and lots of information could be gathered this way.

5.1. Theoretical considerations

Exploration of the various mapping, linking the structural parameters to the sample objective function, can help to understand sources of identification deficiencies. To see this, consider the following problem:

$$\min_\theta [f(\alpha(\theta_0)) - f(g(\theta))]\Omega[f(\alpha(\theta_0)) - f(g(\theta))]' \tag{13}$$
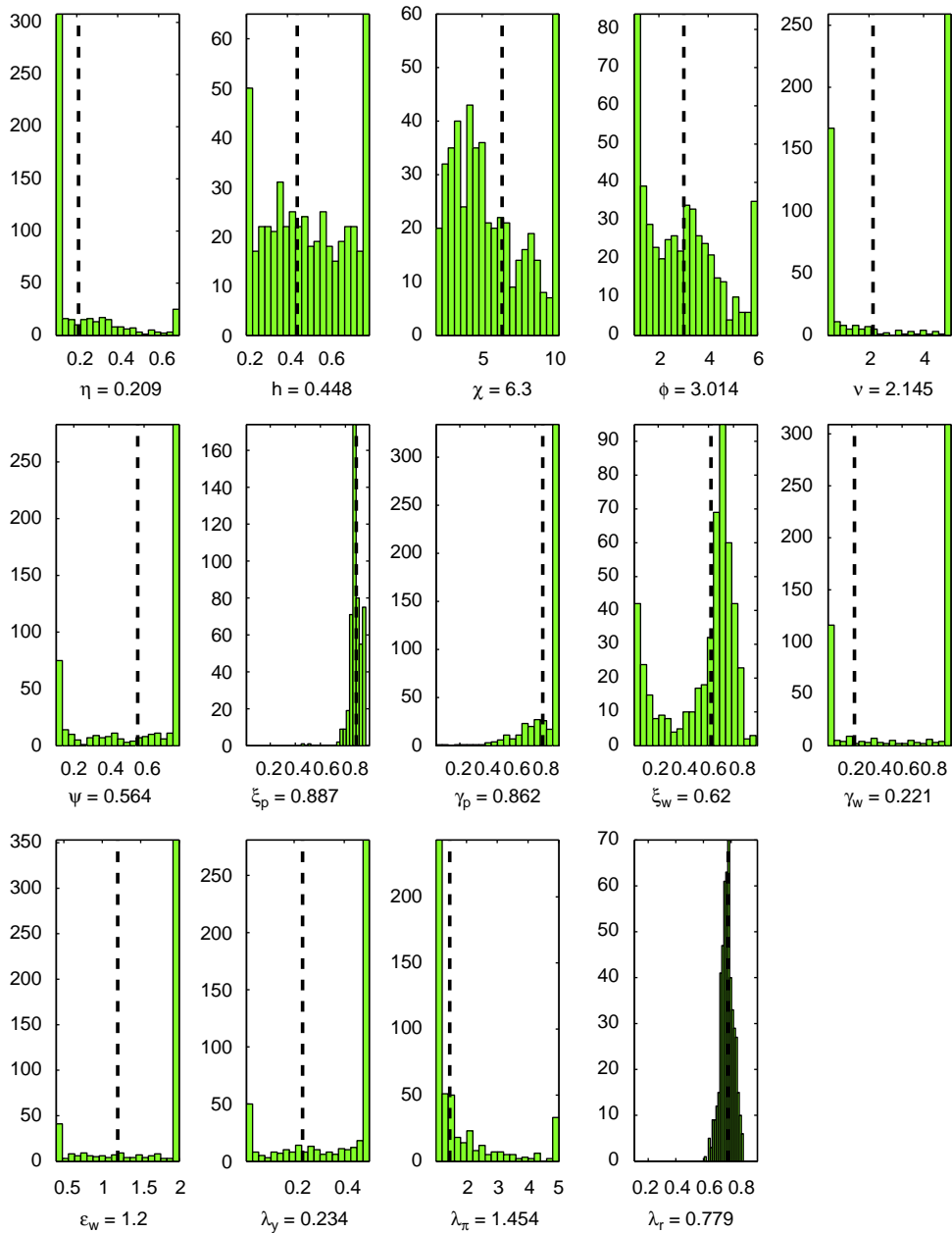
Fig. 5. Histograms of the estimated parameter, $T = 500$; 500 replications.

where $f(\cdot)$ is a $m \times 1$ vector of impulse responses, $\alpha(\theta_0)$ a $m_1 \times 1$ vector of VAR coefficients obtained from the data, $g(\theta)$ a $m_1 \times 1$ vector of VAR coefficients obtained from the model solution, once a $k \times 1$ vector of structural parameters θ is selected, and Ω a weighting matrix. We assume that $f(\cdot)$ and $g(\cdot)$ are differentiable, at least once. Taking a first order Taylor expansion of $f(g(\theta))$ around $f(g(\theta_0))$ leads to

$$\theta - \theta_0 = (G'F'\Omega FG)^{-1}G'F'\Omega(f(\alpha) - f(\alpha_0)) \tag{14}$$

where $F = \partial f/\partial \alpha|_{\alpha_0}$, $G = \partial g/\partial \theta|_{\theta_0}$ and $f(\alpha_0) = f(g(\theta_0))$. Hence, to translate information about $f(\alpha)$ into information about θ, we need to compute $(G'F'\Omega FG)^{-1}$ and this is possible only if the rank of $(G'F'\Omega FG)$ is k. If this condition is not met, at least one $\theta_i, i = 1, \ldots, k$ cannot be identified. If the rank is full, but some of the eigenvalues $(G'F'\Omega FG)$ are small, partial and weak identification problems emerge. Since G and F are the derivatives of the solution and of the moment mappings, the properties of $G'G$ and $F'F$ indicate at what stage information deficiencies may emerge. Iskrev (2007) has applied similar ideas to check for identification deficiencies in a model estimated with likelihood techniques. The conditions he states are similar and essentially require that the G mapping is well behaved.
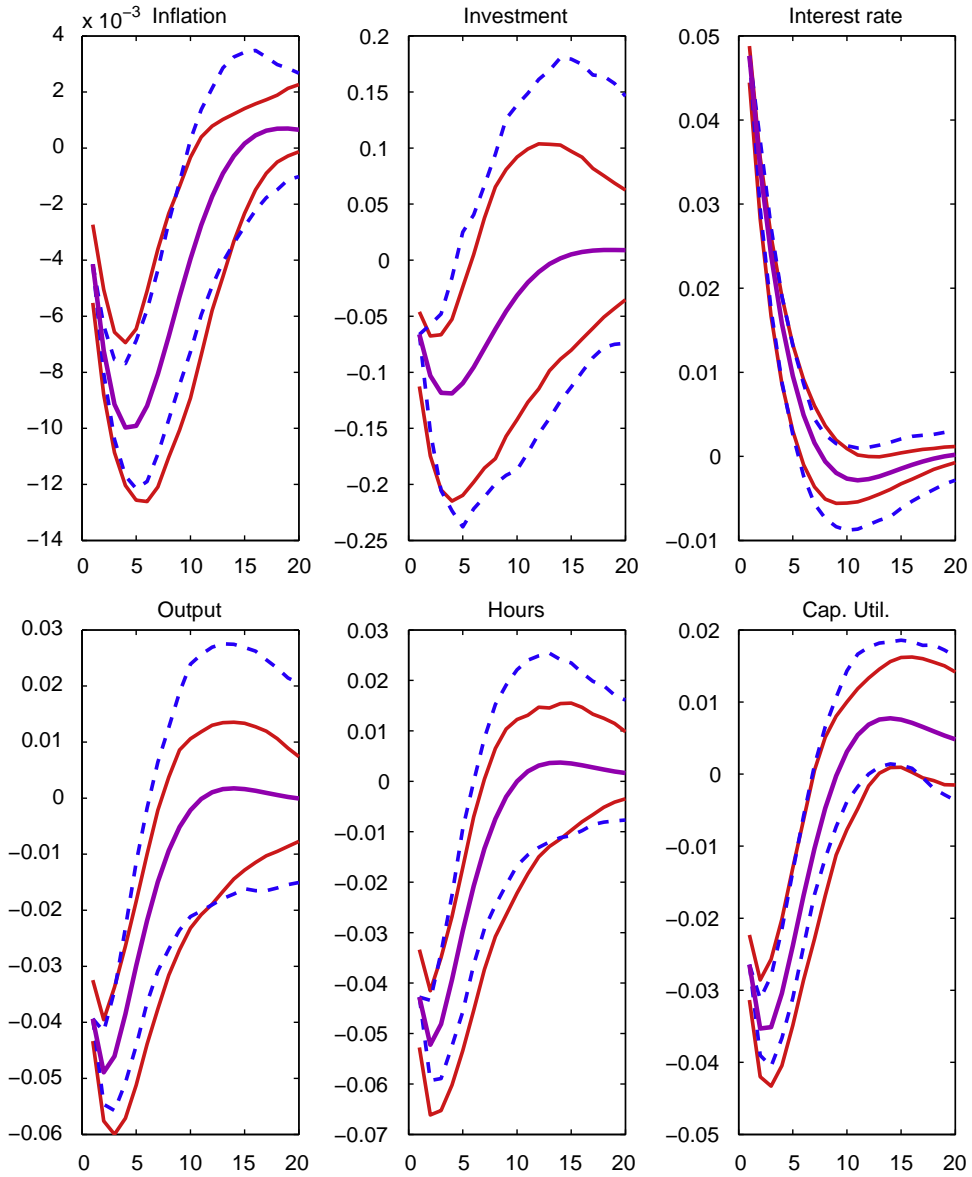
Fig. 6. Estimated impulse responses, $T = 500$. Solid: confidence bands from the model. Dashed: confidence bands from the VAR. Bold: true.

Table 5
Parameter estimates, alternative approach

	ζ_p	γ_p	ζ_w	γ_w
True value	0.887	0.862	0.620	0.221
5%	0.810	0.460	0.510	0.050
95%	0.920	0.970	0.790	0.930

The parameter estimates are obtained inverting the objective function for $T = 500$, after concentrating out the remaining parameters in $\theta_1 \zeta_p$ (ζ_w) is the price (wage) stickiness parameter; γ_p (γ_w) is the price (wage) indexation parameter.

Perhaps, a more informative way to see how F and G convey identification information about the θ is to split the minimization problem into two stages. That is, first, find the θ that minimizes the distance between the VAR parameters in the data, α, and in the model, $g(\theta)$, and, second, find the reduced form parameters α that make impulse responses in the data, $f(\alpha)$ and in the model $(f(g(\theta)))$ close. The first minimization problem tells us about deficiencies in the solution

mapping; the second about deficiencies in the moment mapping. In this case

$$\theta - \theta_0 = (G'\Omega_1 G)^{-1} G'\Omega_1 (\alpha - g(\alpha_0))$$
$$= (G'\Omega_1 G)^{-1} G'\Omega_1 (F'\Omega_2 F)^{-1} F'\Omega_2 (f(\alpha) - f(\alpha_0)) \tag{15}$$

where Ω_1 is a weighting matrix of dimension $m_1 \times m_1$ and Ω_2 is a weighting matrix of dimension $m \times m$. When both these matrices are the identity, information about θ can be obtained from $f(\alpha)$ if and only if $G'G$ and $F'F$ are invertible. Therefore, the rank and the magnitude of the eigenvalues of these two matrices are of crucial interest to assess the identifiability of the parameters of a model.

When only an estimate $\hat{\alpha}$ is available, the deviation of $f(\hat{\alpha})$ from $f(g(\theta))$ can be divided into three parts: one due to the solution mapping, one to the moment mapping and one to the data mapping. To focus on population identification issues in this case, one can compute F and G using the α obtained by calibrating the θ's to reasonable values (as we have done in Section 3) and then perform sensitivity analysis varying θ. This latter step is unlikely to be necessary in practice since estimates of the parameters obtained with structural methods rarely deviate from those used in the calibration literature—optimization routines are often constrained to stay around these values and "strange" estimates are looked upon with suspicion in the profession. A comparison of the conclusions obtained with fixed and estimated α, on the other hand, informs us about the extent of small sample identification failures.

5.2. Practical issues

Methods to test for the rank of a matrix exist in the literature (see e.g. Cragg and Donald, 1997), and they can be used to investigate under-identification pathologies. Here, we focus on ways to check for the size of the eigenvalues of a matrix since weak and partial identification problems are more difficult to detect a priori and likely to be more relevant in practice.

Anderson (1984, p. 475) has shown that estimates of the eigenvalues of a matrix when properly scaled have an asymptotic standard normal distribution. Therefore, the null hypothesis of full rank can be tested against the alternative of rank deficiency by examining whether the smallest eigenvalues of $F'F$ and $G'G$ are statistically different from zero. Since the magnitude of the eigenvalues may depend on the units of measurement, Anderson also suggests to test the null that the sum of the smallest k' eigenvalues to the average of all k eigenvalues is large. This ratio is also asymptotically normally distributed with zero mean and unit variance when properly scaled, and the alternative accounts for the possibility that none of the first k' eigenvalues is exactly zero but that all of them are small. While neither test is directly applicable to our setup—they require that the relevant matrix be consistently estimated under the null and the alternative, which is impossible to do in our case—one could nevertheless use the insights of these tests to diagnose identification anomalies.

The concentration statistics $\mathscr{C}_{\theta_0}(i) = \int_{j\neq i}(g(\theta) - g(\theta_0)) d\theta / \int (\theta - \theta_0) d\theta, i = 1, 2 \ldots$ can also be used. Stock et al. (2002) showed that this statistics synthetically measures the curvature of the objective function around some θ_0 and it is related to the non-centrality parameter of the χ^2 used in testing the hypothesis that the objective function at the optimum is zero. Large values of $\mathscr{C}_{\theta_0}(i)$ imply that it is easy to reject the null if the objective function is not zero; small values, that displacements from the null are difficult to detect. Critical values for $\mathscr{C}_{\theta_0}(i)$ for models which are nonlinear in the parameters do not exist. Still, one could use the values Stock et al. produced in the linear case to get an indication of potential problems.

Is it possible to use (asymptotic) estimates of the standard errors of the parameters to diagnose identification problems? In our opinion, standard errors are not useful for a number of reasons. As Table 4 shows, relatively small standard errors can coexist with severe identification problems. In addition, since identification is a multivariate problem, standard errors are silent as far as ridges in the objective function or near observational equivalence problems are concerned. Moreover, standard errors do not tell us whether problems emerge in population or because the data are uninformative. Finally, detection of identification problems should naturally precede parameter estimation. Estimated standard errors are unlikely to help in this enterprise.

5.3. An application

We apply these ideas to the model of Sections 3 and 4 to demonstrate how can one check for the presence of local information deficiencies and to understand where they come from. The conclusions obtained are very similar to those produced by the graphical and simulation analyses. The (14×14) matrix $G'G$ produced by the responses to monetary policy shocks has the largest eigenvalue that represents 99.9% of the trace. The same is true for the matrix $G'F'FG$: the sum of the smallest 13 eigenvalues is only 0.001% of the trace. Hence, consistent with the findings of Section 2, the solution mapping is the root of the problems.

One can go one step further and ask what features of the model produce such a poorly behaved solution mapping in response to monetary policy shocks. To do this, we numerically examine the mapping between structural parameters and the coefficients of the law of motion of the states—call this matrix G_s. It turns out that the largest eigenvalue of $G'_s G_s$ represent 99.9% of the trace and that the remaining 13 eigenvalues are tiny. Hence, it is the fact that the law of motion of the states is insensitive to changes in many of the structural parameters that generates information deficiencies.

What can one do in this situation? Since the pathologies are intrinsic to the way shocks are transmitted to the endogenous variables and not the result of particular choices made by applied investigators, they can be solved only through model respecification. This term is used here in a general sense to include reparameterization of certain primitive functions, which are now routinely used without much debate (e.g. utility functions), choice of different pivotal points around which to (log-)linearize the solution and use of higher order approximations to the solution. For example, log-linearizing certain new-Keynesian models around the flexible price equilibrium (rather than the steady state) may help to break the multidimensional ridge in price and wage stickiness and indexation found in this paper. On the other hand, Fernandez-Villaverde and Rubio-Ramirez (2005) and An and Schorfheide (2007) have shown that second order approximations produce likelihood functions with better curvature than first order approximations. Such a conclusion, however, does not necessarily carry over to distance functions constructed using impulse responses (see Canova and Sala, 2005, for evidence on this issue).

6. Conclusions and practical suggestions

Liu (1960) and Sims (1980) have argued that traditional simultaneous equations models were hopelessly under-identified and that identification of an economic structure was often achieved not because there was sufficient information in the data but because researchers wanted it to be so—limiting the number of variables in an equation or eschewing a numbers of equations from the model.

Since then models have dramatically evolved; microfoundations have been added; and general equilibrium features taken into account. Still, it appears that a large class of popular DSGE structures are only very weakly identified; observational equivalence is widespread; and reasonable estimates are obtained not because the model and the data are informative but because auxiliary restrictions make the likelihood of the data (or a portion of it) informative. In these situations, structural parameter estimation amounts to sophisticated calibration and this makes model evaluation and economic inference hard. Since DSGE models have never being built with an eye to the identification of their parameters, finding that they are subject to information deficiencies is far from surprising. Perpetuating such a practice, however, can only lead to sterile outcomes.

A study of identification issues like ours, besides ringing a warning bell about the potential problems existing in using DSGE models for inference and policy exercises, is useful in practice only to the extent it gives applied researchers a strategy to detect problems and means to avoid them in practice. Providing such a set of tools is generally complicated since the relationship between the structural parameters and the objective function is nonlinear; the mapping is unknown and problems typically multidimensional.

This paper provides hints on how to detect model vs. data based problems and to understand the reasons for their existence. We summarize our suggestions as a list of non-exhaustive steps we recommend applied investigators to check before attempting structural estimation. First, exploration of the properties of the population objective function should logically precede model estimation. This exploration can be conducted graphically, plotting transformations of the objective function; numerically, using simulation exercises; or algebraically, examining the rank of the matrix of derivatives of the various mappings linking the structural parameters to the population objective function. In the widely used model considered in this paper, all diagnostics give the same conclusions. In others, they may generate complementary information about parameters with problematic identification features. Second, small sample Monte Carlo exercises can help to understand the type of problems one is likely to encounter in practice and measure the incremental distortions induced by the use of the sample version of the objective function. Third, one should be aware that the smaller is the informational content of the objective function, the larger is the chance that identification problems will be present. Thus, when possible, all the implications of the model should be considered. Also, while for identification purposes likelihood methods are generally preferable, even the likelihood function may not cure all problems. Furthermore, improper use of prior restrictions may cause researchers to oversee important informational deficiencies.

Fourth, methods to measure the discrepancy of the model to the data that are robust to identification issues do exist in the literature (see e.g. Canova and Paustian, 2007). However, their (semiparametric) nature, only allow to obtain interval estimates of certain parameters. Thus, if policy considerations require the availability of structural point estimates, and if identification problems persist when all the implications of the model are used, one should go back to the drawing board and rethink about model specification and parameterization. At this stage the use of simplified versions of the model may give some economic intuition for why identification problems emerge, as could the use of several limited information objective functions. Fifth, while popular in the literature, we advise against the use of mixed calibration–estimation approaches, in particular, when partial population identification problems exist. Bringing more data to the analysis will definitively help when identification problems are data based. However, when problems are theory based, it is not obvious what are the advantages of such a strategy and minor mistakes in the setting of calibrated parameters may lead researchers to draw inappropriate conclusions about how the economy works. Furthermore, when micro-based estimates are only vaguely related to macro-based quantities because of heterogeneities and aggregation issues, specification problems could become important. If one insists on using a mixed calibration–estimation approach to analyze economic issues, it is advisable to conduct extensive sensitivity analysis to examine how economic conclusions change when calibrated parameters are varied within a reasonable interval.

Finally, we think it is unfortunate that the identification problems that current macro-models face have not received sufficient consideration in academics and that the agenda has mainly centered on matching facts rather than making sure that the mechanisms proposed to explain the facts are identifiable. We also think that it is unfortunate that large-scale models with numerous frictions are used in policy circles when these models have been subject neither to extensive identification checks nor to proper evaluation analysis. When observational equivalence is present, it is impossible to interpret outcomes in terms of what features are important and what are not. Similarly, when weak and partial identification problems are widespread, distortions spread and policy conclusions become debatable. DSGE models have helped to bridge academic investigations and practical policy analyses. It is important to recognize their current limitations to make knowledge about the economic phenomena accumulate on solid ground.

References

Anderson, T., 1984. An Introduction to Multivariate Statistical Analysis. Wiley, New York.

An, S., Schorfheide, F., 2007. Bayesian analysis of DSGE models. Econometric Reviews 26, 113–172.

Beyer, A., Farmer, R., 2004. On the indeterminacy of new Keynesian economics. ECB Working Paper 323.

Canova, F., forthcoming. What explains the Great Moderation in the US? A structural analysis. Journal of the European Economic Association.

Canova, F., Gambetti, L., 2007. Do expectations matter? The Great Moderation revisited. Available at: ⟨http://www.econ.upf.edu/crei/canova⟩.

Canova, F., Paustian, M., 2007. Measurement with some theory: using sign restrictions to evaluate business cycle models. Available at: ⟨http://www.econ.upf.edu/crei/canova⟩.

Canova, F., Sala, L., 2005, Back to square one: identification issues in DSGE models. ECB Working Paper 583.

Chari, V.V., Kehoe, P., McGrattan, E., 2008. Are structural VARs with long run restrictions useful for developing business cycle theory. Journal of Monetary Economics 55, 1337–1456.

Choi, I., Phillips, P.C., 1992. Asymptotic and finite sample distribution theory for IV estimators and tests in partially identified structural equations. Journal of Econometrics 51, 113–150.

Christiano, L., Eichenbaum, M., Evans, C., 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. Journal of Political Economy 113, 1–45.

Christiano, L., Eichenbaum, M., Vigfusson, R., 2006. Assessing structural VARs. NBER Macroeconomic Annual 21, 1–106.

Cragg, J.G., Donald, S.G., 1997. Inferring the rank of a matrix. Journal of Econometrics 76 (1–2), 223–250.

Dedola, L., Neri, S., 2007. What does a technology shock do? A VAR analysis with model-based sign restrictions. Journal of Monetary Economics 54, 512–549.

Del Negro, M., Schorfheide, F., Smets, F., Wouters, R., 2006. On the fit of new Keynesian models. Journal of Business and Economic Statistics 25, 143–162.

Dufour, J.M., 2003. Identification, weak instruments and statistical inference in econometrics. Canadian Journal of Economics 36, 767–808.

Fernandez-Villaverde, J., Rubio-Ramirez, J., 2005. Estimating dynamic equilibrium economies: linear versus nonlinear likelihood. Journal of Applied Econometrics 20, 891–910.

Fernandez-Villaverde, J., Rubio-Ramirez, J., Sargent, T., Watson, M., 2007. The ABC (and D) of structural VARs. American Economic Review 97, 1021–1026.

Gertler, M., Sala, L., Trigari, A., 2008. An estimated monetary DSGE model with unemployment and staggered nominal wage bargaining. Journal of Money, Credit and Banking 40, 1713–1764.

Geweke, J., 2005. Contemporary Bayesian Econometrics and Statistics. Wiley, NJ.

Hansen, L., Heaton, J., Yaron, A., 1996. Finite sample properties of some alternative GMM estimators. Journal of Business and Economic Statistics 14, 262–280.

Hsiao, C., 1983. Identification. In: Griliches, Z., Intrilligator, M. (Eds.), Handbook of Econometrics, vol. I. Elsevier, Amsterdaam, pp. 224–283.

Ireland, P., 2004. Technology shocks in the new Keynesian model. Review of Economics and Statistics 86, 923–936.

Iskrev, N., 2007. How much do we learn from the estimation of DSGE models? A case study of identification issues in a new Keynesian business cycle model. Manuscript, University of Michigan.

Kleibergen, F., Mavroeidis, S., 2008. Weak instrument robust tests in GMM and the new Keynesian Phillips curve. Manuscript, Brown University.

Koopmans, T., Reiersol, O., 1950. The identification of structural characteristics. Annals of Mathematical Statistics XXI, 165–181.

Liu, T., 1960. Underidentification, structural estimation, and forecasting. Econometrica 28, 855–865.

Moon, R., Schorfheide, F., 2007. A Bayesian look at partially identified models. Manuscript, University of Pennsylvania.

Newey, W., McFadden, D., 1994. Large sample estimates and hypothesis testing. In: Engle, R., McFadden, D. (Eds.), The Handbook of Econometrics, vol. 4, North-Holland, Amsterdam.

Pesaran, H., 1981. Identification of rational expectations models. Journal of Econometrics 16, 375–398.

Poirier, D., 1998. Revising beliefs in nonidentified models. Econometric Theory 14, 183–209.

Rosen, A., 2006. Confidence sets for partially identified parameters that satisfy a finite number of moment inequalities. Manuscript, University College London.

Rothenberg, T., 1971. Identification in parametric models. Econometrica 39, 577–591.

Rotemberg, J., Woodford, M., 1997. An optimization based econometric framework for the evaluation of monetary policy. NBER Macroeconomic Annual 12, 297–346.

Sargent, T., 1978. Estimation of dynamic labor demand schedules under rational expectations. Journal of Political Economy 86, 1009–1044.

Sims, C., 1980. Macroeconomic and reality. Econometrica 48, 1–48.

Smets, F., Wouters, R., 2003. An estimated dynamic stochastic general equilibrium models of the euro area. Journal of the European Economic Association I, 1123–1175.

Stock, J., Wright, J., 2000. GMM with weak identification. Econometrica 68, 1055–1096.

Stock, J., Wright, J., 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.