Analysing identification issues in DSGE models

Nikolai Iskrev, Marco Ratto
Bank of Portugal, Lisbon, PORTUGAL
European Commission, Joint Research Centre, Ispra, ITALY

Preliminary and incomplete

March 10, 2010

Abstract

We present recent developments in the computational tools for analyzing identification in DSGE models. A growing interest is being addressed to identification issues in economic modeling (Canova and Sala, 2009; Komunjer and Ng, 2009; Iskrev, 2010). The aims of the present paper are two-fold. First, we present a new method for computing derivatives with respect to the deep parameters in linearized DSGE models. The availability of such derivatives provides substantial benefits for the quantitative analysis of such models, and, in particular, for the study of identification and the estimation of the model parameters. Closed form expressions for computing analytical derivatives with respect to the vector of deep parameters are presented in (Iskrev, 2010). This method makes an extensive use of sparse Kronecker-product matrices which are computationally inefficient, require a large amount of memory allocation, and are therefore

unsuitable for large-scale models. Our approach in this paper is to compute the derivatives with respect to each parameter separately. This leads to a system of generalized Sylvester equations, which can be solved efficiently and accurately using existing numerical algorithms. We show that this method leads to a dramatic increase in the speed of computations at virtually no cost in terms of accuracy.

The second objective is to present the ongoing development of the identification toolbox within the DYNARE framework. Such a toolbox includes the identification tests recently proposed by Iskrev and aims at integrating them with global sensitivity analysis methodologies (Ratto, 2008), to get useful insight about global identification properties.

Keywords: DSGE models, local identification, estimation, global identification, sensitivity analysis.

1 DSGE Models

This section provides a brief discussion of the class of linearized DSGE models and the restrictions they imply on the first and second order moments of the observed variables.

1.1 Structural model and reduced form

A DSGE model is summarized by a system $\boldsymbol{g}$ of m non-linear equations:

$$E_t\Big(\boldsymbol{g}(\hat{\boldsymbol{z}}_t, \hat{\boldsymbol{z}}_{t+1}, \hat{\boldsymbol{z}}_{t-1}, \boldsymbol{u}_t|\boldsymbol{\theta})\Big) = 0 \tag{1}$$

where $\hat{\boldsymbol{z}}_t$ is a $m-$dimensional vector of endogenous variables, $\boldsymbol{u}_t$ an n-dimensional random vector of structural shocks with $\mathrm{E}\,\boldsymbol{u}_t = \boldsymbol{0}$, $\mathrm{E}\,\boldsymbol{u}_t\boldsymbol{u}_t' = \boldsymbol{I}_n$ and $\boldsymbol{\theta}$ a $k-$dimensional vector of deep parameters. Here, $\boldsymbol{\theta}$ is a point in $\boldsymbol{\Theta} \subset \mathbb{R}^k$ and the parameter space $\boldsymbol{\Theta}$ is defined as the set of all theoretically admissible values of $\boldsymbol{\theta}$.

Currently, most studies involving either simulation or estimation of DSGE models use linear approximations of the original models. That is, the model is first expressed in terms of stationary variables, and then linearized around the steady-state values of these variables. Let $\hat{\boldsymbol{z}}_t$ be a $m-$dimensional vector of the stationary variables, and $\hat{\boldsymbol{z}}^*$ be the steady state value of $\hat{\boldsymbol{z}}_t$, such that $\boldsymbol{g}(\hat{\boldsymbol{z}}^*, \hat{\boldsymbol{z}}^*, \hat{\boldsymbol{z}}^*, 0|\boldsymbol{\theta}) = 0$. Once linearized, most DSGE models can be written in the following form

$$\boldsymbol{\Gamma}_0(\boldsymbol{\theta})\boldsymbol{z}_t = \boldsymbol{\Gamma}_1(\boldsymbol{\theta})\,\mathrm{E}_t\,\boldsymbol{z}_{t+1} + \boldsymbol{\Gamma}_2(\boldsymbol{\theta})\boldsymbol{z}_{t-1} + \boldsymbol{\Gamma}_3(\boldsymbol{\theta})\boldsymbol{u}_t \tag{2}$$

where $\boldsymbol{z}_t = \hat{\boldsymbol{z}}_t - \hat{\boldsymbol{z}}^*$. The elements of the matrices $\boldsymbol{\Gamma}_0$, $\boldsymbol{\Gamma}_1$, $\boldsymbol{\Gamma}_2$ and $\boldsymbol{\Gamma}_3$ are functions of $\boldsymbol{\theta}$.

There are several algorithms for solving linear rational expectations models (see for instance Blanchard and Kahn (1980), Anderson and Moore (1985), King and Watson (1998), Klein (2000), Christiano (2002), Sims (2002)).[1] Depending on the value of $\boldsymbol{\theta}$, there may exist zero, one, or many stable solutions.

[1] Although these algorithms use different representations of the linearized model and of the solution, it is not difficult to convert one representation into another. See the appendix in Anderson (2008) for some examples.

Assuming that a unique solution exists, it can be cast in the following form

$$\boldsymbol{z}_t = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{z}_{t-1} + \boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{u}_t \tag{3}$$

where the $m \times m$ matrix $\boldsymbol{A}$ and the $m \times n$ matrix $\boldsymbol{B}$ are functions of $\boldsymbol{\theta}$.

For a given value of $\boldsymbol{\theta}$, the matrices $\boldsymbol{A}$, $\boldsymbol{\Omega} := \boldsymbol{B}\boldsymbol{B}'$, and $\hat{\boldsymbol{z}}^*$ completely characterize the equilibrium dynamics and steady state properties of all endogenous variables in the linearized model. Typically, some elements of these matrices are constant, i.e. independent of $\boldsymbol{\theta}$. For instance, if the steady state of some variables is zero, the corresponding elements of $\hat{\boldsymbol{z}}^*$ will be zero as well. Furthermore, if there are exogenous autoregressive (AR) shocks in the model, the matrix $\boldsymbol{A}$ will have rows composed of zeros and the AR coefficients. As a practical matter, it is useful to separate the solution parameters that depend on $\boldsymbol{\theta}$ from those that do not. We will use $\boldsymbol{\tau}$ to denote the vector collecting the non-constant elements of $\hat{\boldsymbol{z}}^*$, $\boldsymbol{A}$, and $\boldsymbol{\Omega}$, i.e. $\boldsymbol{\tau} := [\boldsymbol{\tau}_z', \ \boldsymbol{\tau}_A', \ \boldsymbol{\tau}_\Omega']'$, where $\boldsymbol{\tau}_z$, $\boldsymbol{\tau}_A$, and $\boldsymbol{\tau}_\Omega$ denote the elements of $\hat{\boldsymbol{z}}^*$, $\text{vec}(\boldsymbol{A})$ and $\text{vech}(\boldsymbol{\Omega})$ that depend on $\boldsymbol{\theta}$.[2].

In most applications the model in (3) cannot be taken to the data directly since some of the variables in $\boldsymbol{z}_t$ are not observed. Instead, the solution of the DSGE model is expressed in a state space form, with transition equation

[2]The number of constants in the solution matrices may also depend on the solution algorithm one uses. For instance, to write the model in the form used by Sims (2002) procedure, one may have to include in $\boldsymbol{z}_t$ redundant state variables; this will increase the size of the solution matrices and the number of zeros in them. Removing the redundant states and excluding the constant elements from $\boldsymbol{\tau}$ is not necessary, but has practical advantages in terms of speed and numerical accuracy of the calculations

given by (3), and a measurement equation

$$\boldsymbol{x}_t = \boldsymbol{C}\boldsymbol{z}_t + \boldsymbol{D}\boldsymbol{u}_t + \boldsymbol{\nu}_t \tag{4}$$

where $\boldsymbol{x}_t$ is a l-dimensional vector of observed variables, $\boldsymbol{u}_t$ is a q-dimensional vector of exogenous variables, and $\boldsymbol{\nu}_t$ is a p-dimensional random vectors with $\mathrm{E}\,\boldsymbol{\nu}_t = \boldsymbol{0}$, $\mathrm{E}\,\boldsymbol{\nu}_t\boldsymbol{\nu}_t' = \boldsymbol{Q}$, where $\boldsymbol{Q}$ is $p \times p$ symmetric semi-positive definite matrix.

In the absence of a structural model it would, in general, be impossible to fully recover the properties of $\boldsymbol{z}_t$ from observing only $\boldsymbol{x}_t$. Having the model in (2) makes this possible by imposing restrictions, through (3) and (4), on the joint probability distribution of the observables. The model-implied restrictions on the first and second order moments of the $\boldsymbol{x}_t$ are discussed next.

1.2 Theoretical first and second moments

From (3)-(4) it follows that the unconditional first and second moments of $\boldsymbol{x}_t$ are given by

$$\mathrm{E}\,\boldsymbol{x}_t := \boldsymbol{\mu}_{\boldsymbol{x}} = \boldsymbol{s} \tag{5}$$

$$\mathrm{cov}(\boldsymbol{x}_{t+i}, \boldsymbol{x}_t') := \Sigma_{\boldsymbol{x}}(i) = \begin{cases} \boldsymbol{C}\Sigma_{\boldsymbol{z}}(0)\boldsymbol{C}' & \text{if } i = 0 \\ \boldsymbol{C}\boldsymbol{A}^i\Sigma_{\boldsymbol{z}}(0)\boldsymbol{C}' & \text{if } i > 0 \end{cases} \tag{6}$$

where $\Sigma_{\boldsymbol{z}}(0) := \mathrm{E}\,\boldsymbol{z}_t\boldsymbol{z}_t'$ solves the matrix equation

$$\Sigma_{\boldsymbol{z}}(0) = \boldsymbol{A}\Sigma_{\boldsymbol{z}}(0)\boldsymbol{A}' + \boldsymbol{\Omega} \tag{7}$$

Denote the observed data with $\boldsymbol{X}_T := [\boldsymbol{x}_1', \ldots, \boldsymbol{x}_T']'$, and let $\boldsymbol{\Sigma}_T$ be its covariance matrix, i.e.

$$\begin{aligned}
\boldsymbol{\Sigma}_T &:= \mathrm{E}\, \boldsymbol{X}_T \boldsymbol{X}_T' \\
&= \begin{pmatrix} \Sigma_{\boldsymbol{x}}(0), & \Sigma_{\boldsymbol{x}}(1)', & \ldots, & \Sigma_{\boldsymbol{x}}(T-1)' \\ \Sigma_{\boldsymbol{x}}(1), & \Sigma_{\boldsymbol{x}}(0), & \ldots, & \Sigma_{\boldsymbol{x}}(T-2)' \\ \ldots & \ldots & \ldots & \ldots \\ \Sigma_{\boldsymbol{x}}(T-1), & \Sigma_{\boldsymbol{x}}(T-2), & \ldots, & \Sigma_{\boldsymbol{x}}(0) \end{pmatrix}
\end{aligned} \tag{8}$$

Let $\boldsymbol{\sigma}_T$ be a vector collecting the unique elements of $\boldsymbol{\Sigma}_T$, i.e.

$$\boldsymbol{\sigma}_T := [\mathrm{vech}(\Sigma_{\boldsymbol{x}}(0))', \mathrm{vec}(\Sigma_{\boldsymbol{x}}(1))', ..., \mathrm{vec}(\Sigma_{\boldsymbol{x}}(T-1))']'$$

Furthermore, let $\boldsymbol{m}_T := [\boldsymbol{\mu}', \boldsymbol{\sigma}_T']'$ be a $(T-1)l^2 + l(l+3)/2$-dimensional vector collecting the parameters that determine the first two moments of the data. Assuming that the linearized DSGE model is determined everywhere in $\boldsymbol{\Theta}$, i.e. $\boldsymbol{\tau}$ is unique for each admissible value of $\boldsymbol{\theta}$, it follows that $\boldsymbol{m}_T$ is a function of $\boldsymbol{\theta}$. If either $\boldsymbol{u}_t$ is Gaussian, or there are no distributional assumptions about the structural shocks, the model-implied restrictions on $\boldsymbol{m}_T$ contain all information that can be used for the estimation of $\boldsymbol{\theta}$. The identifiability of $\boldsymbol{\theta}$ depends on whether that information is sufficient or not. This is the subject of the next section.

2 Identification

This section explains the role of the Jacobian matrix of the mapping from $\boldsymbol{\theta}$ to $\boldsymbol{m}_T$ for identification, as discussed in Iskrev (2010), and shows how it can be computed analytically, in a more efficient way with respect to Iskrev (2010).

2.1 The rank condition

The probability density function of the data contains all available sample information about the value of the parameter vector of interest $\boldsymbol{\theta}$. Thus, a basic prerequisite for making inference about $\boldsymbol{\theta}$ is that distinct values of $\boldsymbol{\theta}$ imply distinct values of the density function. This is known as the identification condition.

Definition 1. *Let $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^k$ be the parameter vector of interest, and suppose that inference about $\boldsymbol{\theta}$ is made on the basis of T observations of a random vector $\boldsymbol{x}$ with a known joint probability density function $f(\boldsymbol{X};\boldsymbol{\theta})$, where $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T]$. A point $\boldsymbol{\theta}_0 \in \boldsymbol{\Theta}$ is said to be globally identified if*

$$f(\boldsymbol{X};\tilde{\boldsymbol{\theta}}) = f(\boldsymbol{X};\boldsymbol{\theta}_0) \text{ with probability } 1 \Rightarrow \tilde{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 \tag{9}$$

for any $\tilde{\boldsymbol{\theta}} \in \boldsymbol{\Theta}$. If (9) is true only for values $\tilde{\boldsymbol{\theta}}$ in an open neighborhood of $\boldsymbol{\theta}_0$, then $\boldsymbol{\theta}_0$ is said to be locally identified.

In most applications the distribution of $\boldsymbol{X}$ is unknown or assumed to be Gaussian. Thus, the estimation of $\boldsymbol{\theta}$ is usually based on the first two moments of the data. If the data is not normally distributed, higher-order moments

may provide additional information about $\boldsymbol{\theta}$, not contained in the first two moments. Therefore, identification based on the mean and the variance of $\boldsymbol{X}$ is only sufficient but not necessary for identification with the complete distribution. Using the notation introduced in the previous section, we have the following result (see, e.g., Hsiao (1983) and the references therein)

Theorem 1. *Suppose that the data $\boldsymbol{X}_T$ is generated by the model* (3)-(4) *with parameter vector $\boldsymbol{\theta}_0$. Then $\boldsymbol{\theta}_0$ is globally identified if*

$$\boldsymbol{m}_T(\tilde{\boldsymbol{\theta}}) = \boldsymbol{m}_T(\boldsymbol{\theta}_0) \Leftrightarrow \tilde{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 \tag{10}$$

for any $\tilde{\boldsymbol{\theta}} \in \boldsymbol{\Theta}$. If (10) *is true only for values $\tilde{\boldsymbol{\theta}}$ in an open neighborhood of $\boldsymbol{\theta}_0$, the identification of $\boldsymbol{\theta}_0$ is local. If the structural shocks are normally distributed, then the condition in* (10) *is also necessary for identification.*

The condition in (10) requires that the mapping from the population moments of the sample - $\boldsymbol{m}_T(\boldsymbol{\theta})$, to $\boldsymbol{\theta}$ is unique. If this is not the case, there exist different values of $\boldsymbol{\theta}$ that result in the same value of the population moments, and the true value of $\boldsymbol{\theta}$ cannot be determined even with an infinite number of observations. In general, there are no known global conditions for unique solutions of systems of non-linear equations, and it is therefore difficult to establish the global identifiability of $\boldsymbol{\theta}$. Local identification, on the other hand, can be verified with the help of the following condition

Theorem 2. *Suppose that $\boldsymbol{m}_T$ is a continuously differentiable function of $\boldsymbol{\theta}$. Then $\boldsymbol{\theta}_0$ is locally identifiable if the Jacobian matrix $J(q) := \dfrac{\partial \boldsymbol{m}_q}{\partial \boldsymbol{\theta}'}$ has a full column rank at $\boldsymbol{\theta}_0$ for $q \leq T$. This condition is both necessary and sufficient when $q = T$ if $\boldsymbol{u}_t$ is normally distributed.*

This result follows from the implicit function theorem, and can be found, among others, in Fisher (1966) and Rothenberg (1971).[3] Note that, even though $J(T)$ having full rank is not necessary for local identification in the sense of Definition 1, it is necessary for identification from the first and second order moments. Therefore, when the rank of $J(T)$ is less than k, $\boldsymbol{\theta}_0$ is said to be unidentifiable from a model that utilizes only the mean and the variance of $\boldsymbol{X}_T$. A necessary condition for identification in that sense is that the number of deep parameters does not exceed the dimension of $\boldsymbol{m}_T$, i.e. $k \leq (T-1)l^2 + l(l+3)/2$.

The local identifiability of a point $\boldsymbol{\theta}_0$ can be established by verifying that the Jacobian matrix $J(T)$ has full column rank when evaluated at $\boldsymbol{\theta}_0$. Local identification at one point in $\boldsymbol{\Theta}$, however, does not guarantee that the model is locally identified everywhere in the parameter space. There may be some points where the model is locally identified, and others where it is not. Moreover, local identifiability everywhere in $\boldsymbol{\Theta}$ is necessary but not sufficient to ensure global identification. Nevertheless, it is important to know if a model is locally identified or not for the following two reasons. First, local identification makes possible the consistent estimation of $\boldsymbol{\theta}$, and is sufficient for the estimator to have the usual asymptotic properties (see Florens et al. (2008)). Second, and perhaps more important in the context of DSGE models is that with the help of the Jacobian matrix we can detect problems that are

[3] Both Fisher (1966) and Rothenberg (1971) makes the additional assumption that $\boldsymbol{\theta}_0$ is a regular point of $J(T)$, which means that if it belongs to an open neighborhood where the rank of the matrix does not change. Without this assumption the rank condition in Theorem 2 is only sufficient for local identification under normality. Although it is possible to construct examples where regularity does not hold (see Shapiro and Browne (1983)), typically the set of irregular points is of measure zero (see Bekker and Pollock (1986)).

a common cause for identification failures in these models. If, for instance, a deep parameter θ_j does not affect the solution of the model, it will be unidentifiable since its value is irrelevant for the statistical properties of the data generated by the model, and the first and second moments in particular. Consequently, $\frac{\partial \boldsymbol{m}_T}{\partial \theta_j}$ - the column of $J(T)$ corresponding to θ_j, will be a vector of zeros for any T, and the rank condition for identification will fail. Another type of identification failure occurs when two or more parameters enter in the solution in a manner which makes them indistinguishable, e.g. as a product or a ratio. As a result it will be impossible to identify the parameters separately, and some of the columns of the Jacobian matrix will be linearly dependent. An example of the first problem is the unidentifiability of the Taylor rule coefficients in a simple New Keynesian model pointed out in Cochrane (2007). An example of the second is the equivalence between the intertemporal and multisectoral investment adjustment cost parameters in Kim (2003). In these papers the problems are discovered by solving the models explicitly in terms of the deep parameters. That approach, however, is not feasible for larger models, which can only be solved numerically. As will be shown next, the Jacobian matrix in Theorem 2 is straightforward to compute analytically for linearized models of any size or complexity.

2.2 Computing the Jacobian matrix

The simplest method for computing the Jacobian matrix of the mapping from $\boldsymbol{\theta}$ to $\boldsymbol{m}_T$ is by numerical differentiation. The problem with this approach is that numerical derivatives tend to be inaccurate for highly non-linear func-

tions. In the present context this may lead to wrong conclusions concerning the rank of the Jacobian matrix and the identifiability of the parameters in the model. For this reason, Iskrev (2010) applied analytical derivatives, employing implicit derivation. As shown in Iskrev (2010), it helps to consider the mapping from $\boldsymbol{\theta}$ to $\boldsymbol{m}_T$ as comprising two steps: (1) a transformation from $\boldsymbol{\theta}$ to $\boldsymbol{\tau}$; (2) a transformation from $\boldsymbol{\tau}$ to $\boldsymbol{m}_T$. Thus, the Jacobian matrix can be expressed as

$$J(T) = \frac{\partial \boldsymbol{m}_T}{\partial \boldsymbol{\tau}'} \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\theta}'} \tag{11}$$

The derivation of the first term on the right-hand side is straightforward since the function mapping $\boldsymbol{\tau}$ into $\boldsymbol{m}_T$ is available explicitly (see the definition of $\boldsymbol{\tau}$ and equations (5)-(7)); thus the Jacobian matrix $J_1(T) := \frac{\partial \boldsymbol{m}_T}{\partial \boldsymbol{\tau}'}$ may be obtained by direct differentiation.

The elements of the second term $J_2(T) := \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\theta}'}$, the Jacobian of the transformation from $\boldsymbol{\theta}$ to $\boldsymbol{\tau}$, can be divided into three groups corresponding to the three blocks of $\boldsymbol{\tau}$: $\boldsymbol{\tau}_z$, $\boldsymbol{\tau}_A$ and $\boldsymbol{\tau}_\Omega$. In Iskrev (2010) it is assumed that $\hat{\boldsymbol{z}}^*$ is a known function of $\boldsymbol{\theta}$, implied by the steady state of the model, so that the derivative of $\boldsymbol{\tau}_z$ can be computed by direct differentiation. This is in general not true, since one can implement a non-linear DGSE model in packages like DYNARE, which provide the steady state computation and linearization even when the former is not available explicitly. Here we provide the extension to this case, by first noting that the 'static' model $\boldsymbol{g}^* = \boldsymbol{g}(\hat{\boldsymbol{z}}^*, \hat{\boldsymbol{z}}^*, \hat{\boldsymbol{z}}^*, 0|\boldsymbol{\theta}) = 0$ provides and implicit function between $\hat{\boldsymbol{z}}^*$ and $\boldsymbol{\theta}$. Therefore, $\frac{\partial \hat{\boldsymbol{z}}^*}{\partial \boldsymbol{\theta}'}$ can be computed exploiting the analytic derivatives of g^* with respect to $\hat{\boldsymbol{z}}^*$ and $\boldsymbol{\theta}$,

provided by the symbolic pre-processor of DYNARE:

$$\frac{\partial \hat{\boldsymbol{z}}^*}{\partial \boldsymbol{\theta}'} = -\Big(\frac{\partial \boldsymbol{g}^*}{\partial \hat{\boldsymbol{z}}^{*\prime}}\Big)^{-1} \cdot \frac{\partial \boldsymbol{g}^*}{\partial \boldsymbol{\theta}'} \tag{12}$$

and finally $\frac{\partial \boldsymbol{\tau}_z}{\partial \boldsymbol{\theta}'}$ is obtained by removing the zeros corresponding to the constant elements of $\hat{\boldsymbol{z}}^*$.

In order to properly compute the derivatives of $\boldsymbol{\tau}_A$ and $\boldsymbol{\tau}_\Omega$, the structural form (2) has to be re-written explicitly accounting for the dependency to $\hat{\boldsymbol{z}}^*$:

$$\boldsymbol{\Gamma}_0(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)\boldsymbol{z}_t = \boldsymbol{\Gamma}_1(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)\,\mathrm{E}_t\,\boldsymbol{z}_{t+1} + \boldsymbol{\Gamma}_2(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)\boldsymbol{z}_{t-1} + \boldsymbol{\Gamma}_3(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)\boldsymbol{u}_t \tag{13}$$

Also in this case, one can take advantage of the DYNARE symbolic preprocessor. The latter provides derivatives $\frac{\partial \boldsymbol{\Gamma}_i(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)}{\partial \boldsymbol{\theta}'}$ consistent with the form (13). However, since the dependence of $\hat{\boldsymbol{z}}^*$ to $\boldsymbol{\theta}$ is not known explicitly to the preprocessor, these derivatives miss the contribution of the steady state. Therefore, one has to exploit the computation of the Hessian, provided by DYNARE for the second order approximation of non-linear DSGE models. The Hessian gives the missing derivatives $\frac{\partial \boldsymbol{\Gamma}_i(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)}{\partial \hat{\boldsymbol{z}}^{*\prime}}$, allowing one to perform the correct derivation as:

$$\frac{\partial \boldsymbol{\Gamma}_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} = \frac{\partial \boldsymbol{\Gamma}_i(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}'} = \frac{\partial \boldsymbol{\Gamma}_i(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)}{\partial \boldsymbol{\theta}'} + \frac{\partial \boldsymbol{\Gamma}_i(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)}{\partial \hat{\boldsymbol{z}}^{*\prime}} \cdot \frac{\partial \hat{\boldsymbol{z}}^*}{\partial \boldsymbol{\theta}'} \tag{14}$$

The derivatives of $\boldsymbol{\tau}_A$ and $\boldsymbol{\tau}_\Omega$ can be obtained from the derivatives of $\mathrm{vec}(\boldsymbol{A})$ and $\mathrm{vech}(\boldsymbol{\Omega})$, by removing the zeros corresponding to the constant elements of $\boldsymbol{A}$ and $\boldsymbol{\Omega}$. In Iskrev (2010) the derivative of $\mathrm{vec}(\boldsymbol{A})$ is computed using the implicit function theorem. An implicit function of $\boldsymbol{\theta}$ and $\mathrm{vec}(\boldsymbol{A})$ is

provided by the restrictions the structural model (2) imposes on the reduced form (3). In particular, from (3) we have $\mathrm{E}_t\, \boldsymbol{z}_{t+1} = \boldsymbol{A}\boldsymbol{z}_t$, and substituting in (2) yields

$$(\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1\boldsymbol{A})\boldsymbol{z}_t = \boldsymbol{\Gamma}_2\boldsymbol{z}_{t-1} + \boldsymbol{\Gamma}_3\boldsymbol{u}_t \tag{15}$$

Combining the last equation with equation (3) gives to the following matrix equation

$$\boldsymbol{F}(\boldsymbol{\theta}, \mathrm{vec}(\boldsymbol{A})) := \Big(\boldsymbol{\Gamma}_0(\boldsymbol{\theta}) - \boldsymbol{\Gamma}_1(\boldsymbol{\theta})\boldsymbol{A}\Big)\boldsymbol{A} - \boldsymbol{\Gamma}_2(\boldsymbol{\theta}) = \mathbf{O} \tag{16}$$

Vectorizing (16) and applying the implicit function theorem gives

$$\frac{\partial \mathrm{vec}(\boldsymbol{A})}{\partial \boldsymbol{\theta}'} = -\left(\frac{\partial \mathrm{vec}(\boldsymbol{F})}{\partial \mathrm{vec}(\boldsymbol{A})'}\right)^{-1} \frac{\partial \mathrm{vec}(\boldsymbol{F})}{\partial \boldsymbol{\theta}'} \tag{17}$$

Closed-form expressions for computing the derivatives in (17) are provided in Iskrev (2010). Such a derivation requires the use of Kronecker products, implying a dramatic growth in memory allocation requirements and in computational time as the size of the model increases. The typical size of matrices to be handled in Iskrev (2010) is of $m^2 \times m^2$, which grows very rapidly with m. Here we propose an alternative method to compute derivatives, allowing to reduce both memory requirements and the computational time. Taking the derivative of (16) with respect to each θ_j, for $j = 1, \ldots, k$, one gets a set of k equations in the unknowns $\frac{\partial \boldsymbol{A}}{\partial \theta_j}$ of the form:

$$\boldsymbol{M}(\boldsymbol{\theta})\frac{\partial \boldsymbol{A}}{\partial \theta_j} + \boldsymbol{N}(\boldsymbol{\theta})\frac{\partial \boldsymbol{A}}{\partial \theta_j}\boldsymbol{P}(\boldsymbol{\theta}) = \boldsymbol{Q}_j(\boldsymbol{\theta}) \tag{18}$$

where

$$
\begin{aligned}
\boldsymbol{M}(\boldsymbol{\theta}) &= \Big(\boldsymbol{\Gamma}_0(\boldsymbol{\theta}) - \boldsymbol{\Gamma}_1(\boldsymbol{\theta})\boldsymbol{A}(\boldsymbol{\theta})\Big) \\
\boldsymbol{N}(\boldsymbol{\theta}) &= -\boldsymbol{\Gamma}_1(\boldsymbol{\theta}) \\
\boldsymbol{P}(\boldsymbol{\theta}) &= \boldsymbol{A}(\boldsymbol{\theta}) \\
\boldsymbol{Q}_j(\boldsymbol{\theta}) &= \frac{\partial \boldsymbol{\Gamma}_2}{\partial \theta_j} - \Big(\frac{\partial \boldsymbol{\Gamma}_0}{\partial \theta_j} - \frac{\partial \boldsymbol{\Gamma}_1}{\partial \theta_j}\boldsymbol{A}(\boldsymbol{\theta})\Big)\boldsymbol{A}(\boldsymbol{\theta})
\end{aligned}
$$

Equation (18) is a generalized Sylvester equation and can be solved using available algebraic solvers. For example, in DYNARE, this kind of equation is solved applying a QZ factorization for generalized eigenvalues of the matrices $\boldsymbol{M}(\boldsymbol{\theta})$ and $\boldsymbol{N}(\boldsymbol{\theta})$ and solving recursively the factorized problem. It is also interesting to note that the problems to be solved for different θ_j only differ in the right-hand side $\boldsymbol{Q}_j(\boldsymbol{\theta})$, allowing to perform the QZ factorization only once for all parameters in $\boldsymbol{\theta}$. In practice we replace here the single big algebraic problem of dimension $m^2 \times m^2$ of Iskrev (2010) with a set of k problems of dimension $m \times m$.

Using $\boldsymbol{\Omega} = \boldsymbol{B}\boldsymbol{B}'$, the differential of $\boldsymbol{\Omega}$ is given by

$$
\mathrm{d}\boldsymbol{\Omega} = \mathrm{d}\boldsymbol{B}\boldsymbol{B}' + \boldsymbol{B}\,\mathrm{d}\boldsymbol{B}' \tag{19}
$$

Having $\mathrm{d}\boldsymbol{\Omega}$ in terms of $\mathrm{d}\boldsymbol{B}$ is convenient since it shows how to obtain the derivative of $\boldsymbol{\Omega}$ from that of $\boldsymbol{B}$. Note that from equations (15) and (3) we have

$$
\Big(\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1\boldsymbol{A}\Big)\boldsymbol{B} = \boldsymbol{\Gamma}_3 \tag{20}
$$

and therefore

$$\mathrm{d}\boldsymbol{B} = \Big(\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1\boldsymbol{A}\Big)^{-1}\Big(\mathrm{d}\boldsymbol{\Gamma}_3 - (\mathrm{d}\boldsymbol{\Gamma}_0 - \mathrm{d}\boldsymbol{\Gamma}_1\boldsymbol{A} - \boldsymbol{\Gamma}_1\,\mathrm{d}\boldsymbol{A})\Big) \tag{21}$$

Thus, once $\frac{\partial \mathrm{vec}(\boldsymbol{A})}{\partial \boldsymbol{\theta}'}$ is available, it is straightforward to compute, first $\frac{\partial \mathrm{vec}(\boldsymbol{B})}{\partial \boldsymbol{\theta}'}$ and $\frac{\partial \mathrm{vech}(\boldsymbol{\Omega})}{\partial \boldsymbol{\theta}'}$, and then $\frac{\partial \boldsymbol{\tau}_A}{\partial \boldsymbol{\theta}'}$ and $\frac{\partial \boldsymbol{\tau}_\Omega}{\partial \boldsymbol{\theta}'}$.

2.2.1 Extension to second order derivatives

Computing second order derivatives of the model with respect to structural parameters can be performed recursively, starting from knowing second order derivatives of $\boldsymbol{\Gamma}_i$:

$$\begin{aligned}\frac{\partial^2 \boldsymbol{\Gamma}_i(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_l} &= \frac{\partial^2 \boldsymbol{\Gamma}_i(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*(\boldsymbol{\theta}))}{\partial\theta_j\partial\theta_l} = \frac{\partial^2 \boldsymbol{\Gamma}_i(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)}{\partial\theta_j\partial\theta_l} \\ &+ \Big(\frac{\partial}{\partial \hat{\boldsymbol{z}}^{*\prime}}\Big(\frac{\partial \boldsymbol{\Gamma}_i(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)}{\partial \hat{\boldsymbol{z}}^{*\prime}}\Big)' \cdot \frac{\partial \hat{\boldsymbol{z}}^*}{\partial\theta_j}\Big)' \cdot \frac{\partial \hat{\boldsymbol{z}}^*}{\partial\theta_l} + \frac{\partial \boldsymbol{\Gamma}_i(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)}{\partial \hat{\boldsymbol{z}}^{*\prime}} \cdot \frac{\partial^2 \hat{\boldsymbol{z}}^*}{\partial\theta_j\partial\theta_l}\end{aligned} \tag{22}$$

where $\frac{\partial^2 \boldsymbol{\Gamma}_i(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)}{\partial\theta_j\partial\theta_l}$ can be given by the DYNARE symbolic preprocessor and $\frac{\partial}{\partial \hat{\boldsymbol{z}}^{*\prime}}\Big(\frac{\partial \boldsymbol{\Gamma}_i(\boldsymbol{\theta}, \hat{\boldsymbol{z}}^*)}{\partial \hat{\boldsymbol{z}}^{*\prime}}\Big)'$ can be obtained from DYNARE third order approximation of non-linear DSGE models. Moreover, in order to compute $\frac{\partial^2 \hat{\boldsymbol{z}}^*}{\partial\theta_j\partial\theta_l}$, we need the implicit second order derivative from the implicit function $\boldsymbol{g}^* = \boldsymbol{g}(\hat{\boldsymbol{z}}^*, \hat{\boldsymbol{z}}^*, \hat{\boldsymbol{z}}^*, 0|\boldsymbol{\theta}) = 0$:

$$\frac{\partial^2 \hat{\boldsymbol{z}}^*}{\partial\theta_j\partial\theta_l} = -\Big(\frac{\partial \boldsymbol{g}^*}{\partial \hat{\boldsymbol{z}}^{*\prime}}\Big)^{-1} \cdot \Big(\frac{\partial^2 \boldsymbol{g}^*}{\partial\theta_j\partial\theta_l} + \boldsymbol{\gamma}^*\Big) \tag{23}$$

where each element γ_h^*, $h = 1, \ldots, m$, of the vector $\boldsymbol{\gamma}^*$ is given by:

$$\gamma_h^* = \Big(\frac{\partial}{\partial \hat{\boldsymbol{z}}^{*\prime}}\Big(\frac{\partial g_h^*}{\partial \hat{\boldsymbol{z}}^{*\prime}}\Big)' \cdot \frac{\partial \hat{\boldsymbol{z}}^*}{\partial\theta_j}\Big)' \cdot \frac{\partial \hat{\boldsymbol{z}}^*}{\partial\theta_l}$$

and both second order derivatives of $\boldsymbol{g}^*$ with respect to $\boldsymbol{\theta}$ and $\hat{\boldsymbol{z}}^*$ are needed from the DYNARE preprocessor.

Having obtained the second order derivatives of $\boldsymbol{\Gamma}_i$, we can take the second order derivatives of (16) with respect to θ_j and θ_l, for $j, l = 1, \ldots, k$, getting a set of k^2 equations in the unknowns $\frac{\partial^2 \boldsymbol{A}}{\partial\theta_l \partial\theta_j}$ again of the form of a generalized Sylvester equation:

$$\boldsymbol{M}(\boldsymbol{\theta})\frac{\partial^2 \boldsymbol{A}}{\partial\theta_l \partial\theta_j} + \boldsymbol{N}(\boldsymbol{\theta})\frac{\partial^2 \boldsymbol{A}}{\partial\theta_l \partial\theta_j}\boldsymbol{P}(\boldsymbol{\theta}) = \boldsymbol{Q}_{l,j}(\boldsymbol{\theta}) \tag{24}$$

where

$$\begin{aligned}\boldsymbol{Q}_{l,j}(\boldsymbol{\theta}) = &\frac{\partial \boldsymbol{Q}_j}{\partial\theta_l} \\ &- \Big(\frac{\partial \boldsymbol{M}(\boldsymbol{\theta})}{\partial\theta_l}\frac{\partial \boldsymbol{A}}{\partial\theta_j} + \frac{\partial \boldsymbol{N}(\boldsymbol{\theta})}{\partial\theta_l}\frac{\partial \boldsymbol{A}}{\partial\theta_j}\boldsymbol{P}(\boldsymbol{\theta}) + \boldsymbol{N}(\boldsymbol{\theta})\frac{\partial \boldsymbol{A}}{\partial\theta_j}\frac{\partial \boldsymbol{P}(\boldsymbol{\theta})}{\partial\theta_l}\Big)\end{aligned} \tag{25}$$

and

$$\begin{aligned}\frac{\partial \boldsymbol{M}(\boldsymbol{\theta})}{\partial\theta_l} &= \Big(\frac{\partial \boldsymbol{\Gamma}_0(\boldsymbol{\theta})}{\partial\theta_l} - \frac{\partial \boldsymbol{\Gamma}_1(\boldsymbol{\theta})}{\partial\theta_l}\boldsymbol{A}(\boldsymbol{\theta}) - \boldsymbol{\Gamma}_1(\boldsymbol{\theta})\frac{\partial \boldsymbol{A}(\boldsymbol{\theta})}{\partial\theta_l}\Big) \\ \frac{\partial \boldsymbol{N}(\boldsymbol{\theta})}{\partial\theta_l} &= -\frac{\partial \boldsymbol{\Gamma}_1(\boldsymbol{\theta})}{\partial\theta_l} \\ \frac{\partial \boldsymbol{P}(\boldsymbol{\theta})}{\partial\theta_l} &= \frac{\partial \boldsymbol{A}(\boldsymbol{\theta})}{\partial\theta_l} \\ \frac{\partial \boldsymbol{Q}_j(\boldsymbol{\theta})}{\partial\theta_l} &= \frac{\partial^2 \boldsymbol{\Gamma}_2}{\partial\theta_l \partial\theta_j} - \Big(\frac{\partial^2 \boldsymbol{\Gamma}_0}{\partial\theta_l \partial\theta_j} - \frac{\partial^2 \boldsymbol{\Gamma}_1}{\partial\theta_l \partial\theta_j}\boldsymbol{A}(\boldsymbol{\theta})\Big)\boldsymbol{A}(\boldsymbol{\theta}) \\ &\quad - \Big(\frac{\partial \boldsymbol{\Gamma}_0}{\partial\theta_j} - \frac{\partial \boldsymbol{\Gamma}_1}{\partial\theta_j}\boldsymbol{A}(\boldsymbol{\theta})\Big)\frac{\partial \boldsymbol{A}(\boldsymbol{\theta})}{\partial\theta_l} \\ &\quad + \frac{\partial \boldsymbol{\Gamma}_1}{\partial\theta_j}\frac{\partial \boldsymbol{A}(\boldsymbol{\theta})}{\partial\theta_l}\boldsymbol{A}(\boldsymbol{\theta})\end{aligned}$$

The problem (24) can be solved exactly in the same way as for first order

derivatives, still keeping the same QZ decomposition for matrices $\boldsymbol{M}$ and $\boldsymbol{N}$ for all $j, l = 1, \ldots, k$ and only changing the right hand side term $\boldsymbol{Q}_{l,j}$.

3 Computing derivatives: DYNARE implementation

We first summarize here the results and performance of the DYNARE implementation of the computation of first derivatives of DSGE models. The performed two types of checks: (i) consistency between the two analytical approaches and the numerical one (by perturbation); (ii) gain in computational time of the Sylvester equation solution with respect to the approach in Iskrev (2010). We considered a set of models of different size and complexity: Kim (2003), An and Schorfheide (2007), Levine et al. (2008), Smets and Wouters (2007), Ratto et al. (2009) and Ratto et al. (2010). The models of An and Schorfheide (2007) and Smets and Wouters (2007) are linearized DSGE models, and as such their DYNARE implementation already contains explicitly the steady state dependence on $\boldsymbol{\theta}$, thus not requiring the generalized form discussed in (14). On the other hand, the models of Kim (2003), Levine et al. (2008), Ratto et al. (2009) and Ratto et al. (2010) are fed to DYNARE in their full original non-linear form, thus allowing to test all elements of the proposed computational procedure.

The consistency of all different methods for computing derivatives is fulfilled in all models: in particular the maximum absolute difference between numerical derivatives and analytic ones was in the range $(10^{-6} - 10^{-9})$ across

model	Computing time (s)		model size (m)
	Sylvester	Iskrev (2010)	
Kim (2003)	0.0062	0.0447	4
An and Schorfheide (2007)	0.0075	0.054	5
Levine et al. (2008)	0.016	0.109	13
Smets and Wouters (2007)	0.183	5.9	40
Ratto et al. (2009)	1.6	907.6	107
Ratto et al. (2010)	11.1	∞	210

Table 1: Computational time required for the evaluation of first order analytic derivatives of models of growing size.

the different models, while the two analytic approaches are practically identical, in terms of numerical accuracy (maximum absolute difference in the range $(10^{-11} - 10^{-14})$). Concerning computational time, the gain of the approach proposed in this paper is evident looking at Table 1. The computational cost for the Iskrev (2010) approach becomes unsustainable for Ratto et al. (2009) and Ratto et al. (2010). Also note that we performed the tests with a 64-bit version of MATLAB, on a powerful HP ProLiant machine with 4 dual core processors (8 processors as a whole). This has a significant effect on the speed of the algorithm based on Kronecker products, linked to the multi-thread architecture of recent versions of MATLAB. Using only one single dual core processor for Smets and Wouters (2007), the computational cost doubles (11.24 s), while for Ratto et al. (2009) the computation of all derivatives lasted 47.5 minutes!

The present results show that, with the algorithms proposed in this paper, the evaluation of analytic is affordable also for DSGE models of medium/large scale, enabling to perform detailed identification analysis for such kind of models. This is discussed in the next Section.

4 Analyzing local identification of DSGE models: DYNARE implementation

We have discussed in Section 2 the main Theorem 2 for local identification of DSGE models as demonstrated by Iskrev (2010). We need to recall here another necessary condition discussed in Iskrev (2010):

Corollary 1. *The point $\boldsymbol{\theta}_0$ is locally identifiable only if the rank of $J_2 = \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\theta}'}$ at $\boldsymbol{\theta}_0$ is equal to k.*

The condition is necessary because the distribution of $\boldsymbol{X}_T$ depends on $\boldsymbol{\theta}$ only through $\boldsymbol{\tau}$, irrespectively of the distribution of $\boldsymbol{u}_t$. It is not sufficient since, unless all state variables are observed, $\boldsymbol{\tau}$ may be unidentifiable.

4.1 Identification analysis procedure

The procedure is based on Monte Carlo exploration of the space $\boldsymbol{\Theta}$ of model parameters. In particular, a sample from $\boldsymbol{\Theta}$ is made of many randomly drawn points from $\boldsymbol{\Theta}'$, where $\boldsymbol{\Theta} \in \boldsymbol{\Theta}'$ discarding values of $\boldsymbol{\theta}$ that do not imply a unique solution. The set $\boldsymbol{\Theta}'$ contains all values of $\boldsymbol{\theta}$ that are theoretically plausible, and may be constructed by specifying a lower and an upper bound for each element of $\boldsymbol{\theta}$. Such bounds are usually easy to come by from the economic meaning of the parameters. After specifying a distribution for $\boldsymbol{\theta}$ with support on $\boldsymbol{\Theta}'$, one can obtain points from $\boldsymbol{\Theta}$ by drawing from $\boldsymbol{\Theta}'$ and removing draws for which the model is either indetermined or does not have a solution. Conditions for existence and uniqueness are automatically checked by most computer algorithms for solving linear rational expectations

models, including of course DYNARE. The identifiability of each draw $\boldsymbol{\theta}_j$ is then established using the necessary and sufficient conditions discussed by Iskrev (2010):

- Finding that matrix J_2 is rank deficient at $\boldsymbol{\theta}_j$ implies that this particular point in $\boldsymbol{\Theta}$ is unidentifiable in the model.

- Finding that J_2 has full rank but $J(T)$ does not, means that $\boldsymbol{\theta}_j$ cannot be identified given the set of observed variables and the number of observations.

- On the other hand,if $\boldsymbol{\theta}$ is identified at all, it would typically suffice to check the rank condition for a small number of moments, since $J(q)$ is likely to have full rank for q much smaller than T. According to Theorem 2 this is sufficient for identification; moreover, the smaller matrix may be much easier to evaluate than the Jacobian matrix for all available moments. A good candidate to try first is the smallest q for which the order condition is satisfied, and then increase the number of moments if the rank condition fails;

- the DYNARE implementation showed here also analyzes the derivatives of the LRE form of the model ($J_{\boldsymbol{\Gamma}} = \frac{\partial \boldsymbol{\Gamma}_i}{\partial \boldsymbol{\theta}'}$), to check for 'trivial' non-identification problem, like two parameters always entering as a product in Γ_i matrices.

4.2 Weak identification analysis

The previous conditions are related to whether of not columns of $J(T)$ or J_2 are linearly dependent. Another typical avenue in DSGE models is weak identification. This can be tracked by checking conditions like $\frac{\partial \boldsymbol{\tau}}{\partial \theta_j} \approx \sum_{i \neq j} \alpha_i \frac{\partial \boldsymbol{\tau}}{\partial \theta_i}$ or $\frac{\partial \boldsymbol{m}_T}{\partial \theta_j} \approx \sum_{i \neq j} \alpha_i \frac{\partial \boldsymbol{m}_T}{\partial \theta_i}$, i.e. by checking multi-collinearity conditions among columns of $J(T)$ or J_2. In multi collinearity analysis, scaling issues in the Jacobian can matter significantly in interpreting results. In medium-large scale DSGE models there can be as many as thousands entries in $J(q)$ and J_2 matrices (as well as in corresponding $\boldsymbol{m}_q$ and $\boldsymbol{\tau}$ matrices). Each row of $J(q)$ and J_2 correspond to a specific moment or $\boldsymbol{\tau}$ element and there can be differences by orders of magnitude between the values in different rows. In this case, the multi-collinearity analysis would be dominated by the few rows with large elements, while it would be unaffected by all remaining elements. This can imply loss of 'resolution' in multi-collinearity indices, that can result to be too squeezed towards unity. Hence, while exact collinearity among columns would be invariant to the scaling of rows, an improper row scaling can make difficult to distinguish between weak and non-identification. Iskrev (2010) used the elasticities, so that the (j, i) element of the Jacobian is $\frac{\partial m_j}{\partial \theta_i} \frac{\theta_i}{m_j}$. This give the percentage change in the moment for 1% change in the parameter value. Here we re-scale each row of $J(q)$ and J_2 by its largest element in absolute value. In other words, assuming J_2 made of the two rows:

$$\begin{pmatrix} 0.1 & -0.5 & 2.5 \\ -900 & 500 & 200 \end{pmatrix}$$

multi-collinearity analysis will be performed on the scaled matrix:

$$\begin{pmatrix} 0.04 & -0.2 & 1 \\ -1 & 0.5556 & 0.2222 \end{pmatrix}$$

The effect of this scaling is that the order of magnitude of derivatives of any moment (or any $\boldsymbol{\tau}$ element) is the same. In other words, this grossly corresponds to an assumption that the model is equally informative about moments, thus implying equal weights across different rows of the Jacobian matrix.

4.3 DYNARE procedure

A new syntax is available in the β version of DYNARE. The simple keyword `identification(<options>=<values>);` triggers a Monte Carlo exploration described here, based on prior definitions and a list of observed variables entered by the user, using standard DYNARE syntax for setting-up an estimation. Current options are as follows:

- `prior_mc = <integer>` sets the number of Monte Carlo draws (default = 2000);
- `load_ident_files = 0`, triggers a new analysis generating a new sample from the prior space, while `load_ident_files = 1`, loads and displays a previously performed analysis (default = 0);
- `ar = <integer>` (default = 3), triggers the value for q in computing $J(q)$;

- `useautocorr`: this option triggers $J(q)$ in the form of auto-covariances and cross-covariances (`useautocorr = 0`), or in the form of auto-correlations and cross-correlations (`useautocorr = 1`). The latter form normalizes all $\boldsymbol{m}_q$ entries in $[-1, 1]$ which may be useful for comparability of derivatives of different elements of $J(q)$ (default = 0).

5 Examples

5.1 Kim (2003)

This paper demonstrated a functional equivalence between two types of adjustment cost specifications, coexisting in macroeconomic models with investment: intertemporal adjustment costs which involve a nonlinear substitution between capital and investment in capital accumulation, and multisectoral costs which are captured by a nonlinear transformation between consumption and investment. We reproduce results of Kim (2003), worked out analytically, applying the DYNARE procedure on the non-linear form of the model. The representative agent maximizes

$$\sum_{t=0}^{\infty} \beta^t \log C_t \tag{26}$$

subject to a national income identity and a capital accumulation equation:

$$(1-s)\Big(\frac{C_t}{1-s}\Big)^{1+\theta} + s\Big(\frac{I_t}{s}\Big)^{1+\theta} = (A_t K_t^{\alpha})^{1+\theta} \tag{27}$$

$$K_{t+1} = \left[\delta\left(\frac{I_t}{\delta}\right)^{1-\phi} + (1-\delta)K_t^{1-\phi}\right]^{\frac{1}{1-\phi}} \tag{28}$$

where $s = \frac{\beta\delta\alpha}{\Delta}$, $\Delta = 1 - \beta + \beta\delta$, $\phi(\geq 0)$ is the inverse of the elasticity of substitution between I_t and K_t and $\theta(\geq 0)$ is the inverse of the elasticity of transformation between consumption and investment. Parameter ϕ represents the size of intertemporal adjustment costs while θ is called the multisectoral adjustment cost parameter. Kim shows that in the linearized form of the model, the two adjustment cost parameter only enter through an 'overall' adjustment cost parameter $\Phi = \frac{\phi+\theta}{1+\theta}$, thus implying that they cannot be identified separately.

Here we assume that the Kim model is not analytically worked out to highlight this problem of identification. Instead, the analyst feeds the non-linear model (constraints and Euler equation) to DYNARE (also note that the adjustment costs are defined in such a way that the steady state is not affected by them). The identification analysis first tells that the condition number of the $J(q)$ and J_2 matrices is in the range $(10^{12}, 10^{16})$ across the entire Monte Carlo sample. Some numerical rounding errors in the computation of the analytic derivatives discussed in Section 2.2 imply that the rank condition test may or may not pass according to the tolerance for singularity. A much more severe check is performed analysing the multicorrelation coefficient across the columns of $J(q)$ and J_2. Absolute values of such correlation coefficients differ from 1 only by a tiny 10^{-15} across the entire Monte Carlo sample (namely the correlation is negative: -1), thus perfectly revealing the identification problem demonstrated analytically by Kim. We also checked that this result is invariant to row re-scaling, confirming the validity of our approach to better distinguish between weak identification and rank deficiency. This result shows that the procedure by Iskrev (2010) implemented

in DYNARE can help the analyst in detecting identification problems in all typical cases where such problems cannot easily worked out analytically. Perfect collinearity is detected both for J_2 and $J(q)$, implying that sufficient and necessary conditions for local identification are not fulfilled by this model.

It seems also interesting to show here the effect of the number of states fed to DYNARE on the results of the identification analysis. For simplicity of coding, Lagrange multipliers may be explicitly included in the model equations. In this case, one would have an additional equation for the Lagrange multiplier $\lambda_t = \frac{(1-s)^\theta}{(1+\theta)C_t^{(1+\theta)}}$, with λ_t entering the Euler equation. Under this kind of DYNARE implementation, and still assuming that only C_t and I_t can be observed, the multicollinearity test for $J(q)$ still provides correlation values which are virtually -1 for any q, thus confirming the identification problem. On the other hand, due to the specific effect of θ on λ_t, our identification tests would tell that θ and ϕ are separably identified in the model, provided that all states are observed. This exemplifies the nature of the *necessary* condition stated in Corollary 1.

In Figure 1 we show typical plots produced by DYNARE for multicollinearity tests. In the MC analysis performed, for each parameter value sampled from the prior distribution, a multi-collinearity measure is computed. This provides a MC sample of multi-collinearity measures for each parameter. Such samples are plotted in DYNARE in the form of box and whiskers plots. Boxplots are made of (i) a central box that indicates the width of the central quartiles of the empirical distribution in the MC sample (i.e. the width from the 25% to 75% quantiles); (ii) a red line indicating the median of the empirical distribution; (iii) whiskers are lines that indi-

cate the 'tail' of the distribution, and extend up to a maximum width of 1.5 times the width of the central [25%, 75%] box; (iv) MC points falling outside the maximum whiskers width, are taken as 'outliers' and plotted as circles. Such 'outliers' indicate a small subset of values of multi-collinearity coefficients that are very different form the bulk of the MC sample. In the box and whiskers plots of Figure 1 we can see that, when λ_t is included in the model, the sample of multi-collinearity coefficients of J_2 for ϕ and θ is centered around a value 0.98, near but not equal to one, and a number of 'outliers' with small correlation is detected. This kind of plot reflects the *necessary* nature of Corollary 1 and usually indicate some possible weak identification problems. The bottom graph, showing the box and whiskers plots of $J(q)$, clearly shows the collinearity problems of ϕ and θ, given that λ_t is not observed.

5.2 An and Schorfheide (2007)

The model An and Schorfheide (2007), linearized in log-deviations from steady state, reads:

$$y_t = \mathbb{E}_t[y_{t+1}] + g_t - \mathbb{E}_t[g_{t+1}] - 1/\tau \cdot (R_t - \mathbb{E}_t[\pi_{t+1}] - \mathbb{E}_t[z_{t+1}]) \tag{29}$$

$$\pi_t = \beta \mathbb{E}_t[\pi_{t+1}] + \kappa(y_t - g_t) \tag{30}$$

$$R_t = \rho_R R_{t-1} + (1-\rho_R)\psi_1 \pi_t + (1-\rho_R)\psi_2(\Delta y_t - z_t) + \varepsilon_{R,t} \tag{31}$$

$$g_t = \rho_g g_{t-1} + \varepsilon_{g,t} \tag{32}$$

$$z_t = \rho_z z_{t-1} + \varepsilon_{z,t} \tag{33}$$

where y_t is GDP in efficiency units, π_t is inflation rate, R_t is interest rate, g_t is government consumption and z_t is change in technology. The model is completed with three observation equations for quarterly GDP growth rate (YGR_t), annualized quarterly inflation rates (INF_t) and annualized nominal interest rates (INT_t):

$$YGR_t = \gamma^Q + 100 * (y_t - y_{t-1} + z_t) \tag{34}$$

$$INFL_t = \pi^A + 400\pi_t \tag{35}$$

$$INT_t = \pi^A + r^A + 4\gamma^Q + 400R_t \tag{36}$$

where $\beta = \frac{1}{1+r^A/400}$.

The rank condition tests for rank deficiencies in $J(q)$ and J_2 are passed by the list model parameters. In Figure 2 we show the box and whiskers plots for multicollinearity for this model: the model parameters on the x-axes are ranked in decreasing order of weakness of identification, i.e. the parameters at the left are those most likely to be weakly identified. Multicollinearity in the model does not signal any problem. On the other hand, the plot for moments indicate that weak identification problems may occur for specially for ψ_1 and ψ_2. The check pairwise correlations is also performed, as shown in Figure 3. There is no extremely large pairwise correlation pattern, however it is interesting to note the links between ψ_1, ψ_2 and ρ_R. Moreover, auto-correlations of exogenous shocks are linked to the corresponding shock standard deviation. This is a quite typical outcome, since the variance of an autocorrelated shock depends on its persistence through the relation $\sigma^2/(1-\rho^2)$, which affects the moments magnitude.

5.3 Smets and Wouters (2007)

All parameters estimated in Smets and Wouters (2007) pass the rank conditions of Iskrev (2010) (Figure 4). Multi-collinearity analysis (Figure 5) and pairwise correlation analysis (Figures 6-8) suggest possible weak identification issues for moments, while in the model no problem is highlighted. Parameters in the left part of Figure 5 are most likely to be weakly identified. Constraining them to, e.g., their prior mean is most likely to affect only slightly estimation results, due to the possibility of the model parameterization to compensate this constraint by opportunely adjusting other parameters collinear to them. This can be the case for *crpi* (r_π the weight of inflation in the Taylor rule) and *cry* (r_y: the weight of output in the Taylor rule). These two parameters are also quite significantly correlated (Figure 8). Also interesting is to notice in Figure 7 correlations between *csigl* (σ_l) and *cprobw* (ξ_w: Calvo parameter for wages) and between *csigma* (σ_c: inverse of elasticity of substitution) and *chabb* (λ: habit persistence). The latter couple, however, does not seem to be specially affected by weak identification problems. Finally, similar correlation patterns as in An and Schorfheide (2007) for parameters in exogenous shocks can be seen in Figure 6,

5.4 Ratto et al. (2009)

All parameters estimated in Ratto et al. (2009) pass the rank conditions of Iskrev (2010) (Figure 9). Multi-collinearity analysis (Figure 10) and pairwise correlation analysis (Figure 11) suggest possible weak identification issues. Parameters in the left part of Figure 10 are most likely to be weakly identified.

For example, this happens for $WRLAGE$ (real wage rigidity) or $GAMWE$ (nominal wage rigidity). These two parameters have large multi-collinearity also for J_2 (top graph in Figure 10), meaning that even with available information for all states, weak identification would be present there. A significant pairwise correlation is also detected for ($WRLAGE$, $GAMWE$), both in $J(q)$ and J_2, explaining the weak identification result. Similarly to Kim (2003), model linearization seems to mitigate separable effects of those two parameters. Finally, the usual strong pairwise correlations between the standard deviation of exogenous shocks and their persistence were detected.

5.5 Ratto et al. (2010)

All parameters estimated in Ratto et al. (2010) pass the rank conditions of Iskrev (2010) (Figure 12). Multi-collinearity analysis (Figure 13-14) gives very similar results as Ratto et al. (2009) concerning weak identification issues.

6 Conclusions

We proposed a new approach for computing analytic derivatives of linearized DSGE models. This method proved to dramatically improve the speed of computation with respect to Iskrev (2010), virtually without any loss in accuracy. Furthermore, we implemented in DYNARE the local identification procedure proposed by Iskrev (2010) and tested it on a number of estimated DSGE model in the literature. In general, all DSGE models pass the necessary and sufficient condition for local identification. The most interest-

ing aspect to be analyzed in detail is therefore weak identification. Multicollinearity coefficients seem a useful measure for weak identification and pairwise correlation analysis can highlight pairs of parameters which act in a very similar way. One thing about the multicollinearity analysis is that sometimes it may be misleading about weak identification. This is because if the moments are very sensitive to a parameter, this may partially offset the strong multicollinearity. Basically the weak identification is an interaction of the two things: the sensitivity and the multicollinearity. The parameter σ_C in Smets and Wouters (2007) is a good example of that: it is overall better identified than its multicollinearity would suggest because the derivative of the moments with respect to σ_C is large (relative to the value of σ_C). We noticed that the multi-collinearity analysis for this parameter is very sensitive to scaling of the Jacobian: not applying any scaling, our analysis would flag σ_C as one of the most prone to weak identification, while with the scaling applied here or in Iskrev (2010) this is not the case. So, with the analysis based on the Jacobian it can be difficult to measure the overall result about weak identification. Another caveat is that the model in not equally informative about all moments, so they may have to be weighted differently.

In addition to these caveats, we can see a number of possible lines of improvement of current procedure:

- improve the mapping of weak identification, highlighting regions in the prior space where such problems are most sensible;
- deepen the analysis of multi-collinearity structure, to possibly highlight systematic patterns across the entire prior space: the existence of

such patterns may suggest ways to re-parameterize the model to make identification stronger.

Finally, some procedure to inspect global identification features would be of great importance. Research is in progress in this direction.

References

An, S. and F. Schorfheide (2007). Bayesian analysis of DSGE models. *Econometric Reviews 26*(2-4), 113–172. DOI:10.1080/07474930701220071.

Anderson, G. (2008, March). Solving linear rational expectations models: A horse race. *Computational Economics 31*(2), 95–113.

Anderson, G. and G. Moore (1985). A linear algebraic procedure for solving linear perfect foresight models. *Economics Letters 17*(3), 247–252. available at http://ideas.repec.org/a/eee/ecolet/v17y1985i3p247-252.html.

Bekker, P. A. and D. S. G. Pollock (1986, March). Identification of linear stochastic models with covariance restrictions. *Journal of Econometrics 31*(2), 179–208. available at http://ideas.repec.org/a/eee/econom/v31y1986i2p179-208.html.

Blanchard, O. J. and C. M. Kahn (1980, July). The solution of linear difference models under rational expectations. *Econometrica 48*(5), 1305–11. available at http://ideas.repec.org/a/ecm/emetrp/v48y1980i5p1305-11.html.

Canova, F. and L. Sala (2009, May). Back to square one: identification issues in DSGE models. *Journal of Monetary Economics 56*(4).

Christiano, L. J. (2002). Solving dynamic equilibrium models by a method of undetermined coefficients. *Computational Economics 20*(1-2).

Cochrane, J. H. (2007, September). Identification with taylor rules: A critical review. NBER Working Papers 13410, National Bureau of Economic Research, Inc.

Fisher, F. (1966). *The identification problem in econometrics*. McGraw-Hill.

Florens, J.-P., V. Marimoutou, and A. Péguin-Feissolle (2008). *Econometric Modelling and Inference*. Cambridge.

Hsiao, C. (1983, June). Identification. In Z. Griliches and M. D. Intriligator (Eds.), *Handbook of Econometrics*, Volume 1 of *Handbook of Econometrics*, Chapter 4, pp. 223–283. Elsevier.

Iskrev, N. (2010). Local identification in DSGE models. *Journal of Monetary Economics 57*, 189–202.

Kim, J. (2003, February). Functional equivalence between intertemporal and multisectoral investment adjustment costs. *Journal of Economic Dynamics and Control 27*(4), 533–549.

King, R. G. and M. W. Watson (1998, November). The solution of singular linear difference systems under rational expectations. *International Economic Review 39*(4), 1015–26.

Klein, P. (2000, September). Using the generalized schur form to solve a multivariate linear rational expectations model. *Journal of Economic Dynamics and Control 24*(10), 1405–1423. available at http://ideas.repec.org/a/eee/dyncon/v24y2000i10p1405-1423.html.

Komunjer, I. and S. Ng (2009). Dynamic identification of DSGE models. unpublished manuscript.

Levine, P., J. Pearlman, and R. Pierse (2008). Linear-quadratic approximation, external habit and targeting rules. *Journal of Economic Dynamics and Control 32*(10), 3315 – 3349.

Ratto, M. (2008). Analysing DSGE models with global sensitivity analysis. *Computational Economics 31*, 115–139.

Ratto, M., W. Roeger, and J. in 't Veld (2009). QUEST III: An estimated open-economy DSGE model of the euro area with fiscal and monetary policy. *Economic Modelling 26*(1), 222 – 233.

Ratto, M., W. Roeger, and J. in 't Veld (2010, January). Using a DSGE model to look at the recent boom-bust cycle in the US. European Economy. Economic Papers 397, European Commission, Brussels.

Rothenberg, T. J. (1971, May). Identification in parametric models. *Econometrica 39*(3), 577–91. available at http://ideas.repec.org/a/ecm/emetrp/v39y1971i3p577-91.html.

Shapiro, A. and M. Browne (1983, June). On the investigation of local identifiability: A counterexample. *Psychometrika 48*(2), 303–304.

Sims, C. (2002). Solving rational expectations models. *Computational Economics 20*, 1–20.

Smets, F. and R. Wouters (2007, June). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *The American Economic Review 97*(3), 586–606.

7 Figures

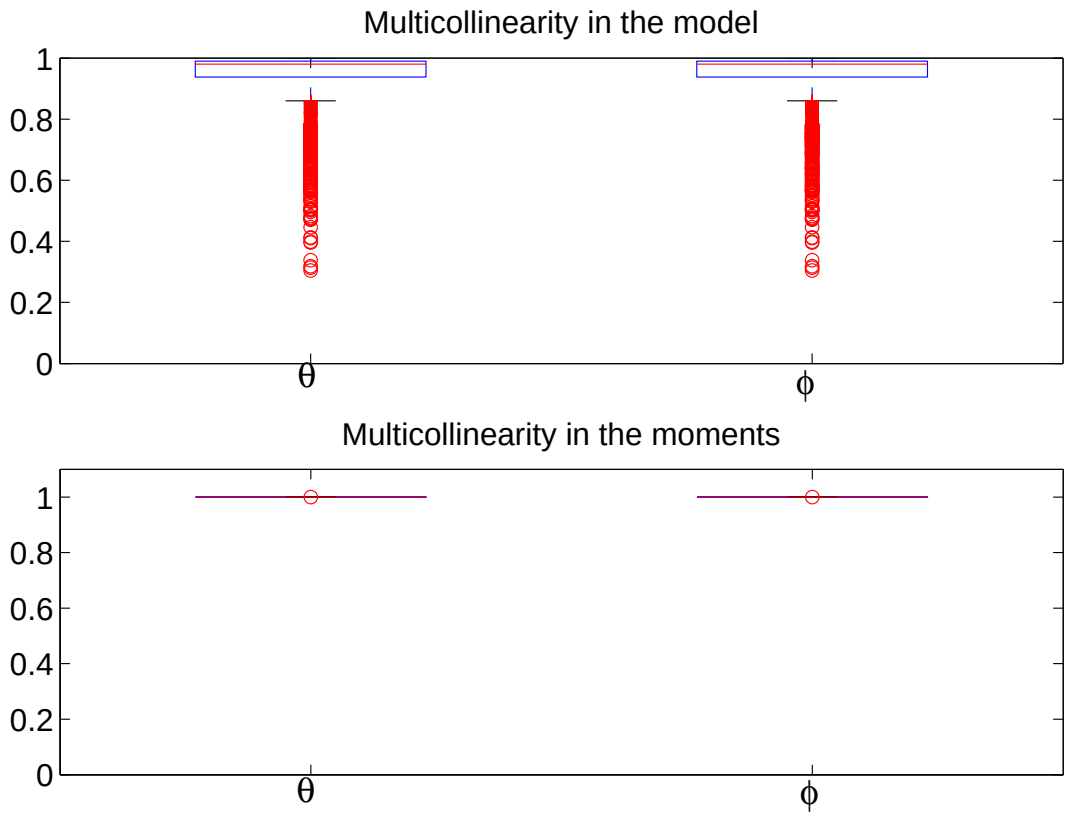

Figure 1: DYNARE Boxplots for identification analysis of the Kim model).

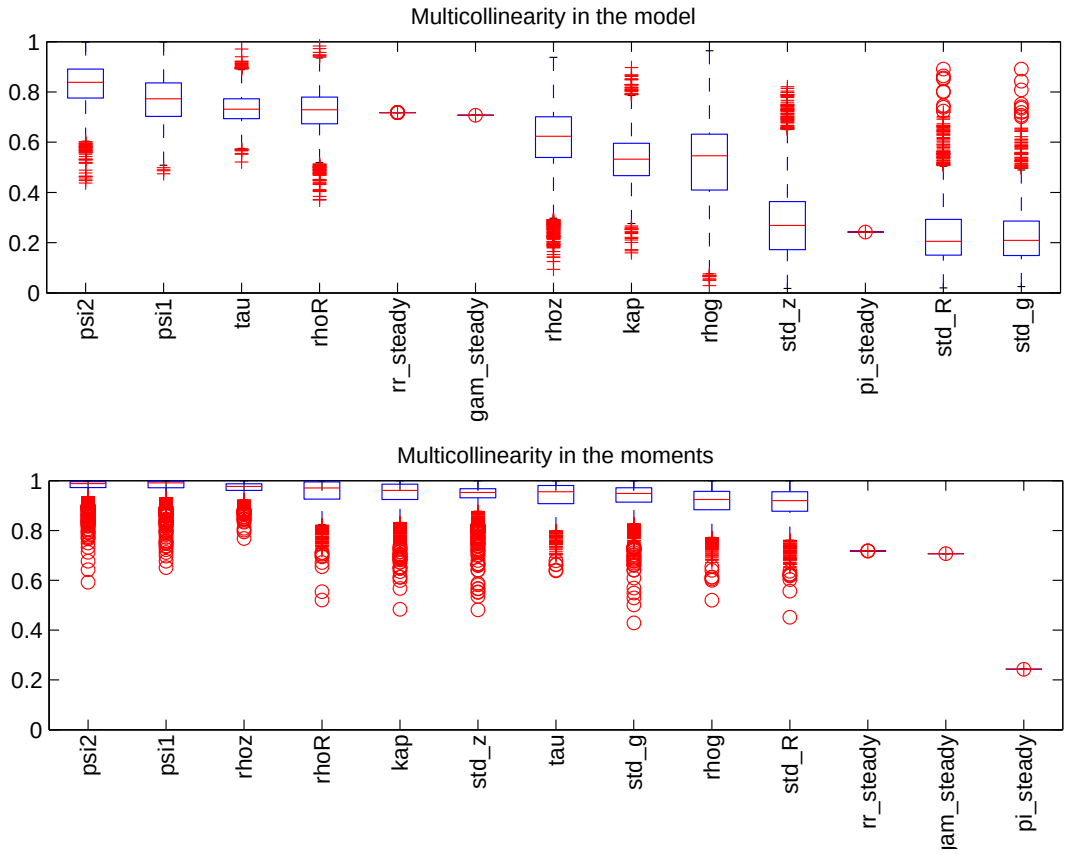

Figure 2: DYNARE Boxplots for identification analysis of the An and Schorfheide (2007) model.

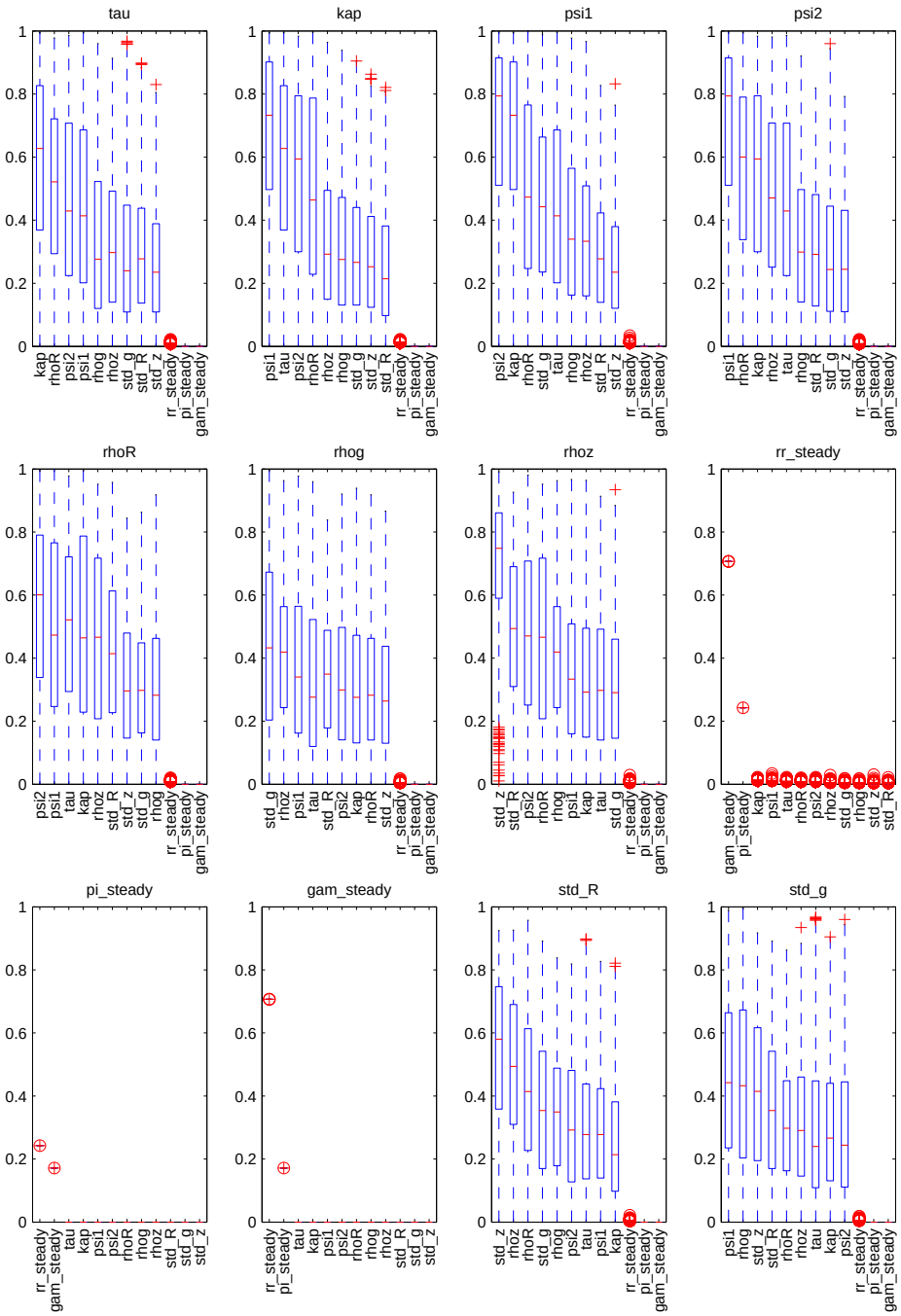

Figure 3: DYNARE Boxplots for pairwise correlations in $J(q)$ columns for the An and Schorfheide (2007) model.

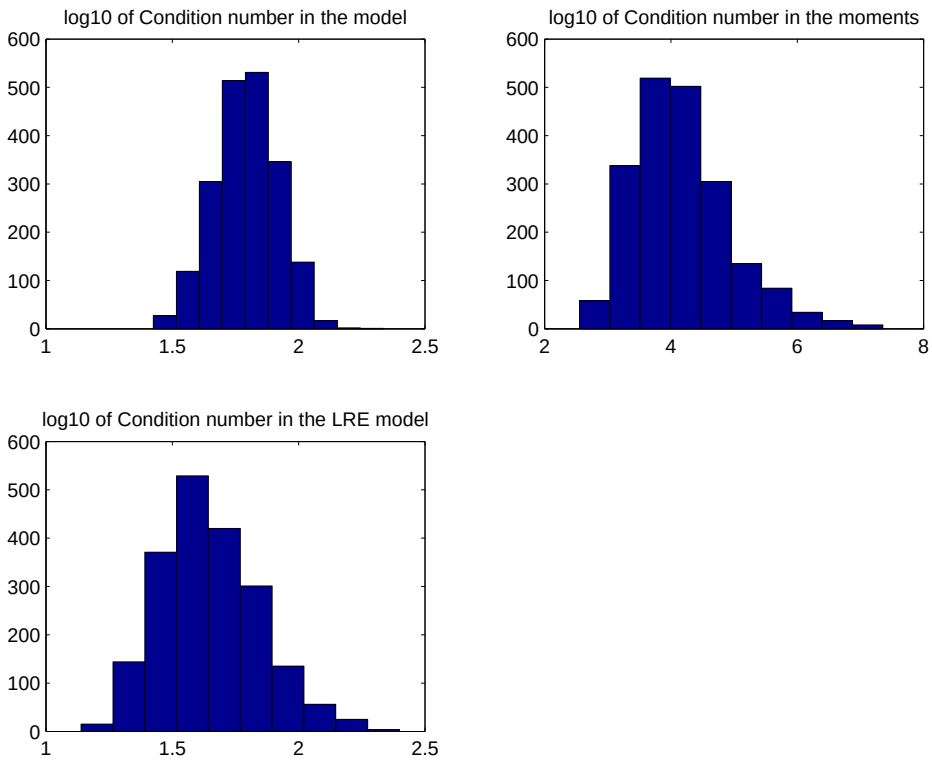

Figure 4: Distributions of condition numbers of J_2, $J(q)$, J_Γ for the Smets and Wouters (2007) model.

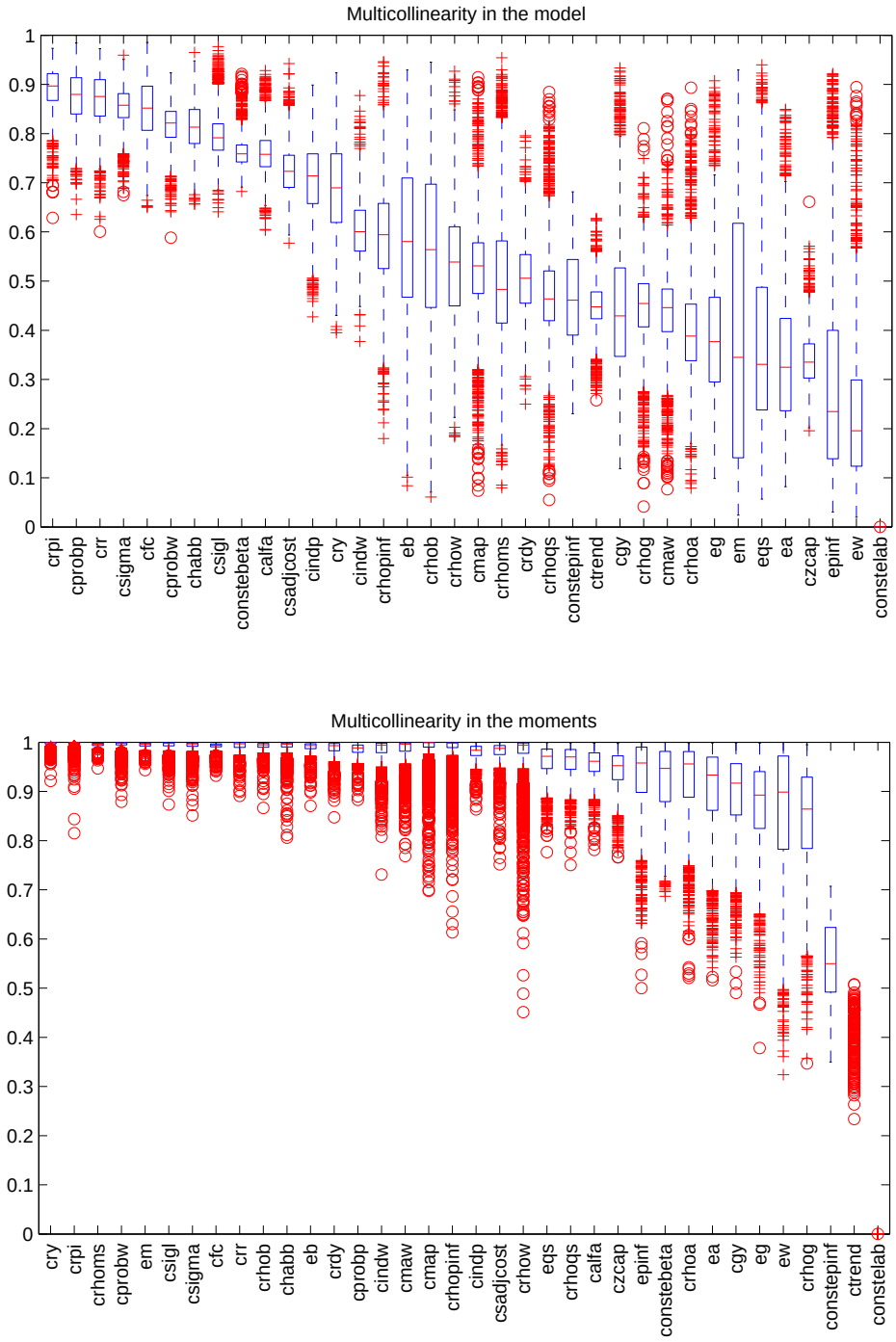

Figure 5: DYNARE Boxplots for identification analysis of the Smets and Wouters (2007) model.

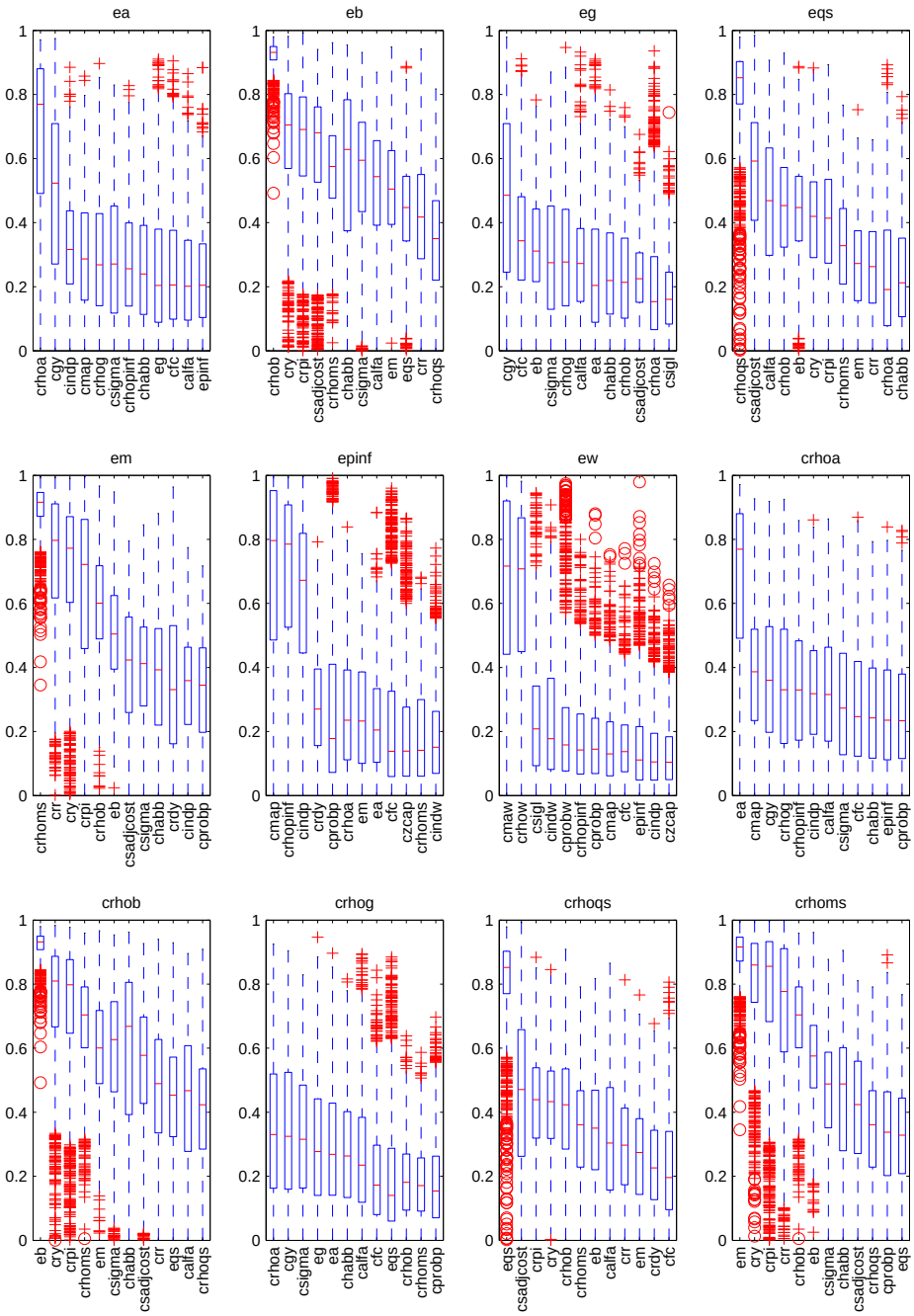

Figure 6: DYNARE Boxplots for pairwise correlations in $J(q)$ columns for the Smets and Wouters (2007) model.

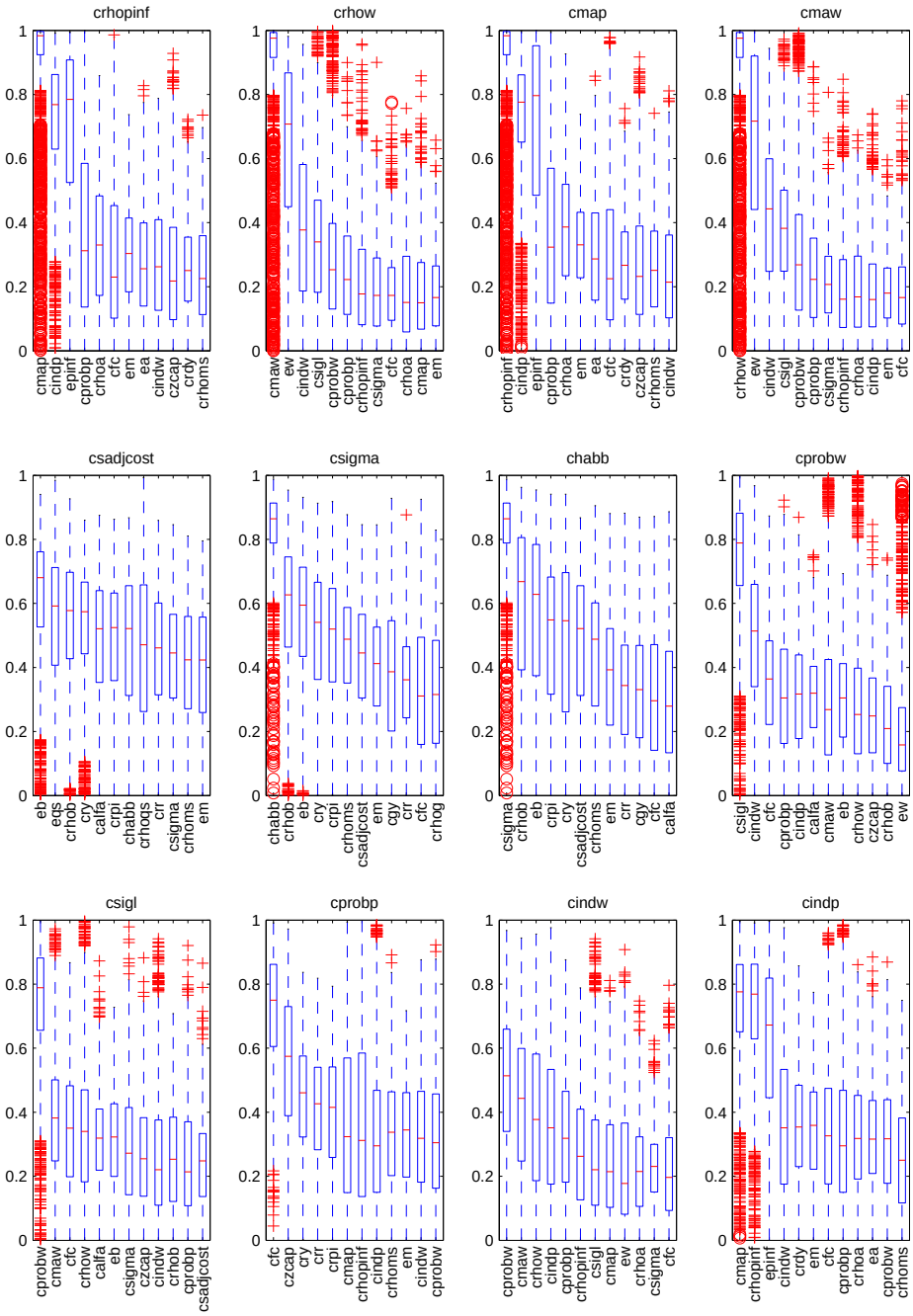

Figure 7: DYNARE Boxplots for pairwise correlations in $J(q)$ columns for the Smets and Wouters (2007) model.

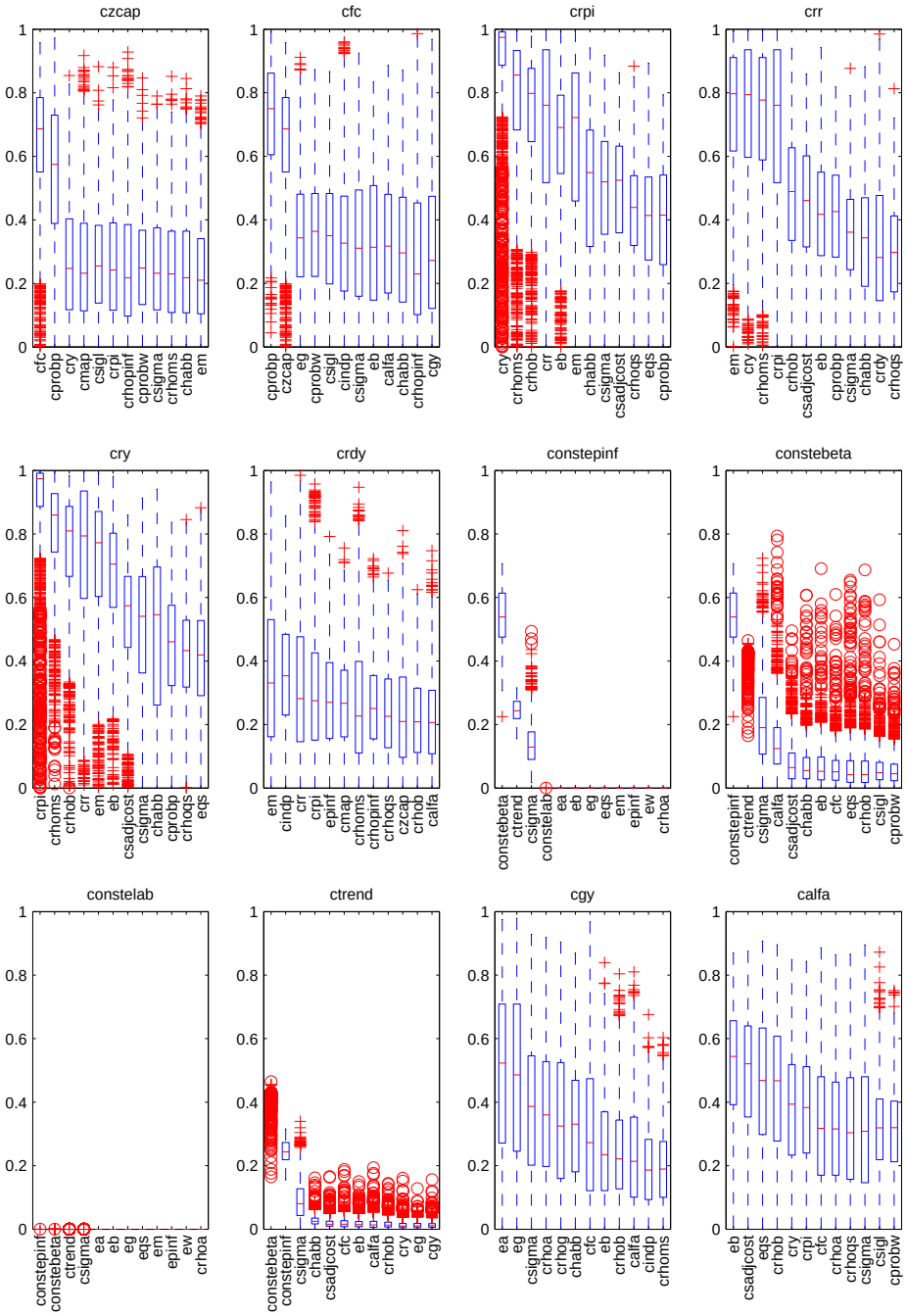

Figure 8: DYNARE Boxplots for pairwise correlations in $J(q)$ columns for the Smets and Wouters (2007) model.

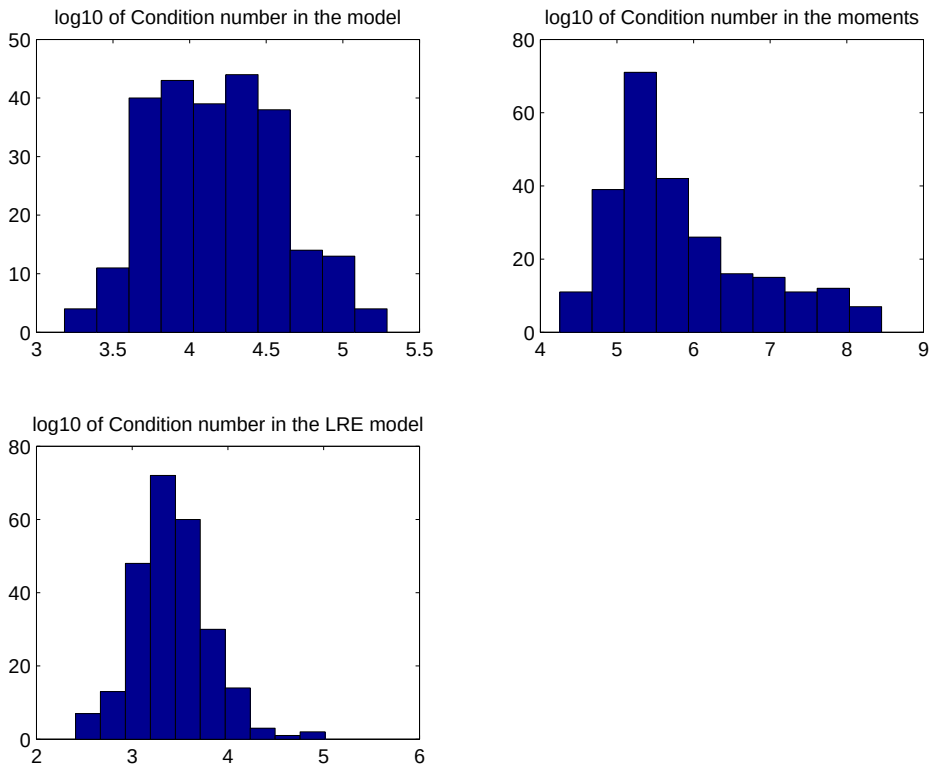

Figure 9: Distributions of condition numbers of J_2, $J(q)$, J_Γ for the Ratto et al. (2009) model.

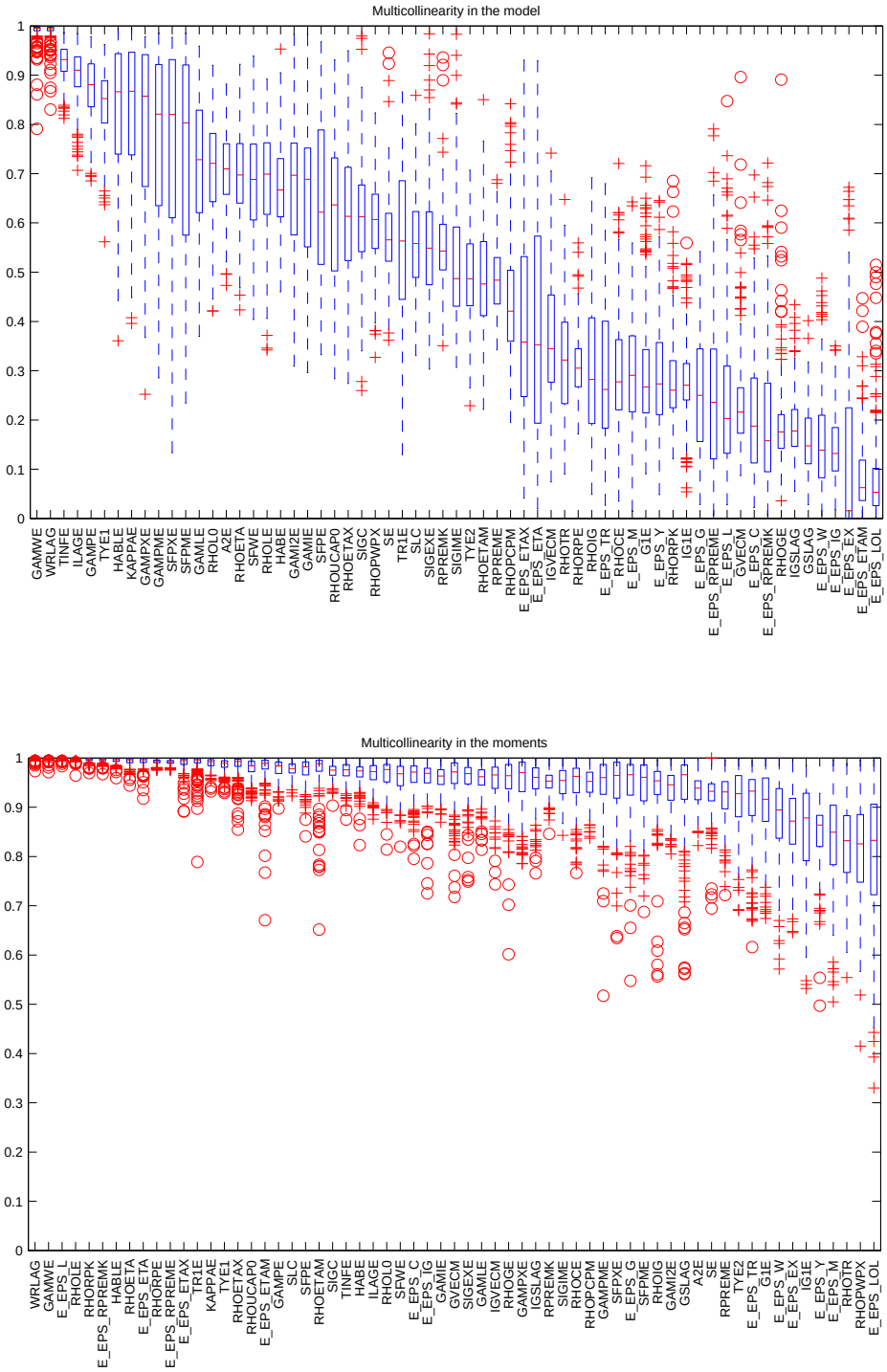

Figure 10: DYNARE Boxplots for identification analysis of the Ratto et al. (2009) model.

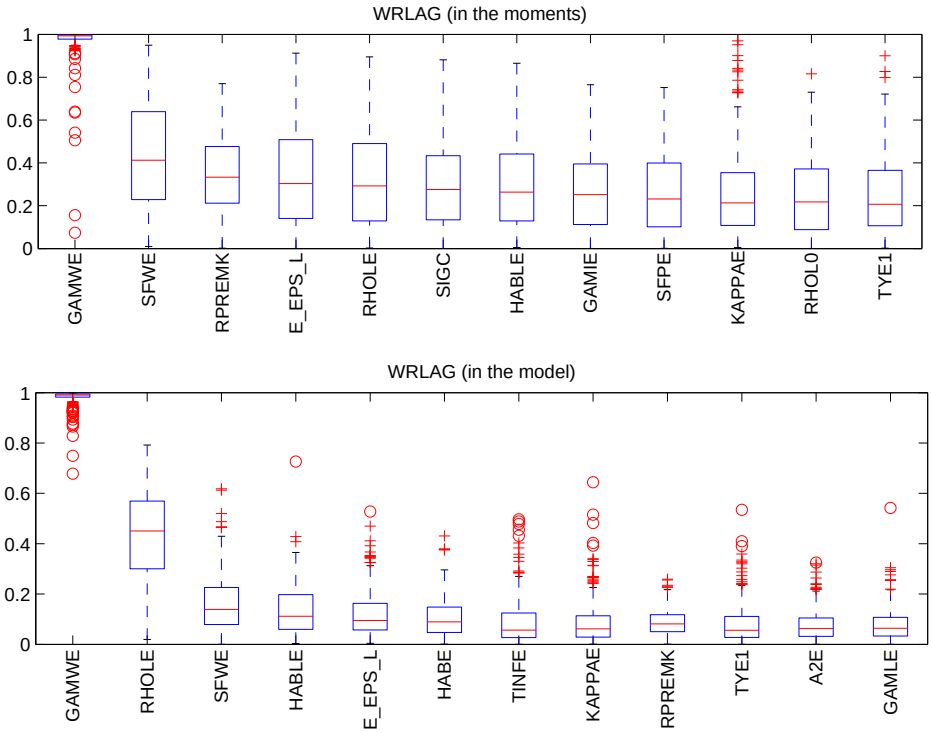

Figure 11: DYNARE Boxplots for most relevant pairwise correlations in $J(q)$ columns (top graph) and J_2 (bottom graph) for the Ratto et al. (2009) model.

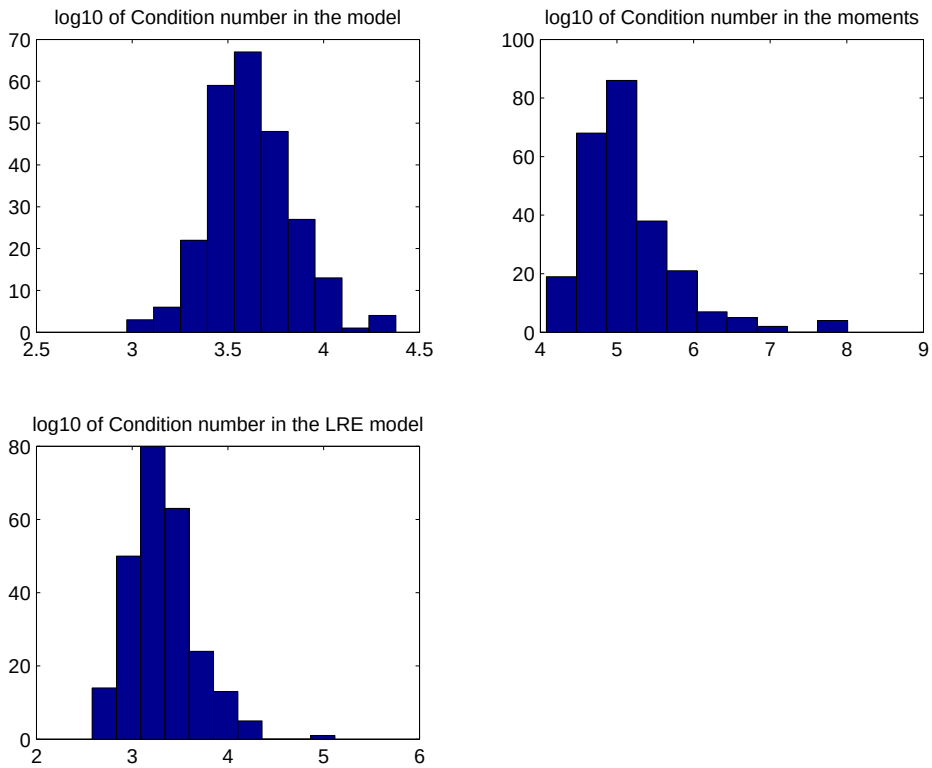

Figure 12: Distributions of condition numbers of J_2, $J(q)$, J_Γ for the Ratto et al. (2010) model.

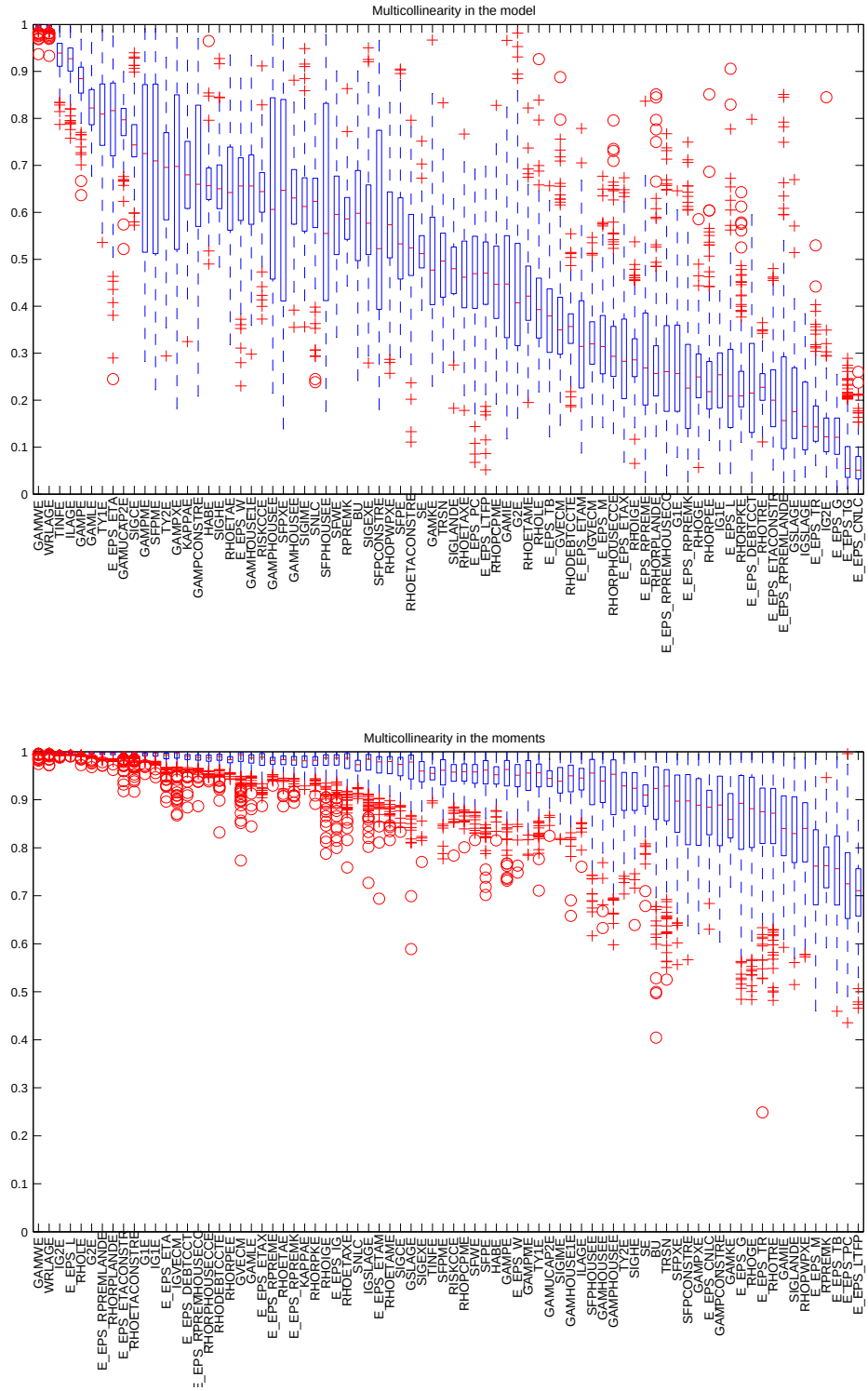

Figure 13: DYNARE Boxplots for identification analysis of the Ratto et al. (2010) model.

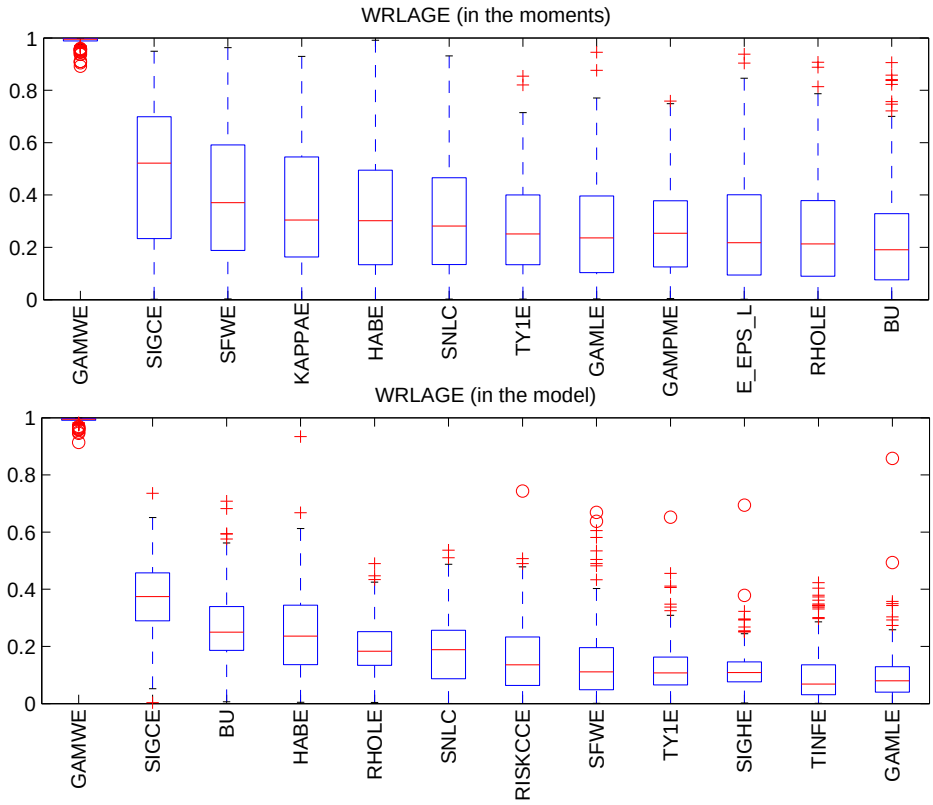

Figure 14: DYNARE Boxplots for most relevant pairwise correlations in $J(q)$ columns (top graph) and J_2 (bottom graph) for the Ratto et al. (2010) model.