

Global identification of linearized DSGE models

ANDRZEJ KOCIĘCKI
Narodowy Bank Polski

MARCIN KOLASA
Narodowy Bank Polski and SGH Warsaw School of Economics

This paper introduces a computational framework to analyze global identification of linearized DSGE models. A formal identification condition is established that relies on the restrictions linking the observationally equivalent state space representations and on the inherent constraints imposed by the model solution on the deep parameters. This condition is next used to develop an algorithm that checks global identification by searching for observationally equivalent model parametrizations. The algorithm is efficient as the identification conditions it employs shrink considerably the space of candidate deep parameter points and the model does not need to be solved at each of these points. The working of the algorithm is demonstrated with two examples.

KEYWORDS. Global identification, DSGE models, state-space representation.

JEL CLASSIFICATION. C13, C51, E32.

1. INTRODUCTION

Dynamic stochastic general equilibrium (DSGE) models have developed into useful tools for macroeconomic analysis. A growing number of policy making institutions, and central banks in particular, use them not only for designing counterfactual experiments, but also for assessing the current stance of the economy and forecasting. The latter application has been supported by growing evidence that estimated medium-sized DSGE models can be competitive with time series models and expert judgment (Del Negro and Schorfheide (2013)).

It has been well understood that DSGE models can suffer from serious identification deficiencies, making estimation and economic inference problematic. Early examples of simple unidentified DSGE models include Kim (2003), Beyer and Farmer (2007), or Cochrane (2011). Canova and Sala (2009) brought this issue to attention of a wider

Andrzej Kocięcki: andrzej.kociecki@nbp.pl

Marcin Kolasa: marcin.kolasa@nbp.pl

Part of this paper was written while Marcin Kolasa was visiting the Columbia University. The authors would like to thank three anonymous referees for very detailed and constructive comments. This paper also benefited from discussions with Ivana Komunjer, Denis Tkachenko, Daniel Waggoner, and Tao Zha. Finally, the authors acknowledge comments received from the participants to the NBP Summer Workshop in Warsaw, CEF conference in Oslo, IAAE conference in London, EEA-ESEM Congress in Toulouse and NBP Workshop “Identification in Macroeconomics” in Warsaw. The views expressed herein are those of the authors and not necessarily those of Narodowy Bank Polski.

group of applied modelers by developing simple diagnostic tools for detecting problems with identification in more empirically-oriented DSGE models estimated by impulse response matching. Their findings allowed them to state in this context that “observational equivalence, partial and weak identification problems are widespread.” A more formal analysis, applicable also to likelihood-based methods, is offered by [Iskrev \(2010\)](#). He established conditions for local identification based on the rank of the Jacobian matrix that maps the deep parameters of a DSGE model to its implied first and second moments of observable variables. [Komunjer and Ng \(2011\)](#) drew on control theory and spectral analysis. They derived their local identification conditions based on the rank of the appropriately defined Jacobian matrix that uses the restrictions between the observationally equivalent state space systems. Another important theoretical contribution is by [Qu and Tkachenko \(2012\)](#), who establish the rank conditions for local identification using the spectral density matrix that maps from deep model parameters to functions defined in a Banach space.

Local identification is a necessary condition for existence of well-behaved estimators, and hence is “to be or not to be” for econometricians. However, one might argue that what really matters for economists is whether there exists another point in the parameter space, possibly distinctly far from the original one, that results in the same probability distribution. This question relates to the problem of global identification, which still remains a hardly explored area in the context of DSGE models. To our knowledge, there are only three formal attempts to handle global identification in the literature to date. [Fukac, Waggoner, and Zha \(2007\)](#) considered models with no latent variables that can be solved analytically. [Morris \(2014\)](#) extended this analysis to somewhat richer models, but still relies on rather stringent assumptions that allowed him to write the model as a vector autoregression with one lag. Finally, [Qu and Tkachenko \(2017\)](#) worked in the frequency domain and offered a more general framework to check global identification by assessing the Kullback–Leibler distance between two parametrizations of a DSGE model.

In this paper, we offer an alternative theoretical analysis of global identification of DSGE models. There are two ingredients of our approach. The first one relies on the results from [Komunjer and Ng \(2011\)](#), who show that observationally equivalent state space representations, where the observational equivalence is defined in terms of the spectral density, are related by a special similarity transformation. Second, we derive the inherent constraints that are imposed by the model solution on the structural parameters to establish the formal and operational condition for their global identification. We merge these two insights to develop an algorithm that checks global identification by searching for observationally equivalent model parametrizations. The algorithm is efficient as the identification conditions it employs shrink considerably the space of candidate deep parameter points. To apply the algorithm, we need to solve the DSGE model only for the parameter value for which we are checking identification, but not for any other alternative parametrizations. This makes our procedure not only fast, but probably also relatively accurate, given that potential numerical inaccuracy is introduced each time a DSGE model is solved. In this respect, our identification procedure can be contrasted with [Qu and Tkachenko \(2017\)](#)—the only currently available alternative attempt to check global identification that is applicable to a broad class of DSGE models.

Our approach has several further important features that make it particularly attractive. As already mentioned, it has a global rather than local flavor. Moreover, unlike the previous literature on local identification (Iskrev (2010), Komunjer and Ng (2011), Qu and Tkachenko (2012)), our method does not rely on evaluating the rank of the matrices that are obtained numerically. Problems that may arise with this approach are acknowledged by Iskrev (2010) and Komunjer and Ng (2011); see also Canova, Ferroni, and Matthes (2014). In contrast, our approach is based on solving the system of nonlinear equations, and hence is less prone to the aforementioned problems. On the downside, and unlike Qu and Tkachenko (2017), the current version of our framework does not allow for indeterminacy.

The rest of this paper is structured as follows. Section two lays out the structure of a typical DSGE model and its solution. Section 3 discusses the equivalent state space representations and the related concept of global identification. In Section 4, we work out the formal condition for global identification. Section 5 presents our algorithm for checking global identification. In Section 6, we demonstrate the usefulness of our algorithm with a simple a-theoretical state-space model, which can also be analyzed analytically. In Section 7, the algorithm is applied to a variant of the widely analyzed DSGE model of An and Schorfheide (2007). Section 8 concludes.

2. DSGE MODEL

A DSGE model is a system of nonlinear equations involving conditional expectations. While solving this type of models using global methods is in principle possible, it can be prohibitively time consuming unless the number of state variables is very small. In consequence, most studies use local approximations of the original models. In particular, likelihood-based estimation that requires calculating the model solution at each optimization step is usually done with linearized models.¹

Once linearized, most DSGE models can be cast in the following form:

$$\Gamma_0(\theta) \begin{bmatrix} s_t \\ p_t \end{bmatrix} = \Gamma_1(\theta) E_t \begin{bmatrix} s_{t+1} \\ p_{t+1} \end{bmatrix} + \Gamma_2(\theta) s_{t-1} + \Gamma_3(\theta) \varepsilon_t, \quad (1)$$

where s_t is an $n \times 1$ vector of states, p_t is a $q \times 1$ vector of policy variables, matrices $\Gamma_0(\theta)$, $\Gamma_1(\theta)$, $\Gamma_2(\theta)$, and $\Gamma_3(\theta)$ are explicit functions of deep model parameters collected in an $m \times 1$ vector $\theta \in \Theta \subseteq \mathbb{R}^m$, and $\varepsilon_t \sim i.i.d. N(0, \Sigma(\theta))$ is a $k \times 1$ vector of shocks, where $\Sigma(\theta)$ is a $k \times k$ symmetric positive definite matrix for every $\theta \in \Theta$. In our notation, $\varepsilon_t = [(\varepsilon_t^s)' (\varepsilon_t^m)']'$ collects innovations to k^s structural shocks ε_t^s and k^m measurement errors ε_t^m so that the last k^m columns in $\Gamma_3(\theta)$ are zero. We do not impose any additional rank restrictions on $\Gamma_0(\theta)$, $\Gamma_1(\theta)$, $\Gamma_2(\theta)$, and $\Gamma_3(\theta)$, except that the underlying matrix pencil is regular (see, e.g., King and Watson (1998); Klein (2000)) so that the model is well formulated.²

¹See Fernández-Villaverde, Rubio-Ramírez, and Santos (2006) for a discussion on how second-order approximation errors affect the likelihood function.

²The form of equation (1) is the same as in Iskrev (2010) and compatible with the Dynare package (Adjemian et al. (2011)), except that we single out variables showing up in lags (but possibly also at the

Let y_t denote an $r \times 1$ vector of observable variables. Then the measurement equations, linking observables to the model variables, can be written as

$$y_t = H(\theta) \begin{bmatrix} s_t \\ p_t \end{bmatrix} + J(\theta)\varepsilon_t, \quad (2)$$

where $H(\theta)$ is an $r \times (n + q)$ matrix and $J(\theta)$ is an $r \times k$ matrix, both of which may (explicitly) depend on θ . In a typical formulation, matrix $J(\theta)$ loads measurement errors, and hence its first k^s columns are zero.

In what follows, we deal with the model parametrizations that imply unique stable solutions, that is, the parameter space Θ excludes those values of θ that result in indeterminacy or for which no stable solution exists. Such a solution of (1), obtainable for example, from Dynare, can be written as

$$s_t = A(\theta)s_{t-1} + B(\theta)\varepsilon_t, \quad (3)$$

$$p_t = F(\theta)s_{t-1} + G(\theta)\varepsilon_t, \quad (4)$$

where $A(\theta)$ is an $n \times n$ matrix, $B(\theta)$ is an $n \times k$ matrix, $F(\theta)$ is a $q \times n$ matrix, and $G(\theta)$ is a $q \times k$ matrix, all of which implicitly depend on deep model parameters θ .³ For future reference, we will call (4) the policy function and $F(\theta)$, $G(\theta)$ as policy matrices.

If we decompose $H(\theta)$ into blocks corresponding to the state and policy variables $H(\theta) = [H^s(\theta) H^p(\theta)]$, then using the model solution (3) and (4) allows us to rewrite measurement equation (2) as

$$y_t = C(\theta)s_{t-1} + D(\theta)\varepsilon_t, \quad (5)$$

where matrices $C(\theta)$ and $D(\theta)$ of dimension $r \times n$ and $r \times k$, respectively, are defined as

$$C(\theta) = H^s(\theta)A(\theta) + H^p(\theta)F(\theta), \quad (6)$$

$$D(\theta) = H^s(\theta)B(\theta) + H^p(\theta)G(\theta) + J(\theta). \quad (7)$$

Consequently, the law of motion for observable variables y_t has a representation in the state space form given by transition equation (3) and measurement equation (5). For future reference, such a representation will be called the ABCD-representation. The term is not accidental and indicates that we are in the world of the A, B, C, and D's explored by [Fernández-Villaverde et al. \(2007\)](#).

Note that some of the matrices defined above depend on the deep model parameters explicitly, while others usually only implicitly. We will refer to the former group, represented by $\Gamma_0(\theta)$, $\Gamma_1(\theta)$, $\Gamma_2(\theta)$, $\Gamma_3(\theta)$, $\Sigma(\theta)$, $H(\theta)$, and $J(\theta)$, as semistructural parameters.

current or future periods). We call these variables as states and refer to the remaining ones as policy variables. In the Appendix, we give more details on this classification and show how the widely cited model of [Smets and Wouters \(2007\)](#) can be easily cast into the form required by our analysis. Note that the chosen representation of a DSGE model is general in the sense that it is easily transformable from those used by popular solution algorithms, for example, [Blanchard and Kahn \(1980\)](#), [Anderson and Moore \(1985\)](#), [Uhlig \(1999\)](#), [Klein \(2000\)](#), [King and Watson \(2002\)](#), or [Sims \(2002\)](#). See also [Anderson \(2008\)](#).

³If we allowed for indeterminacy, equations (3) and (4) would additionally need to include sunspot shocks; see [Lubik and Schorfheide \(2003\)](#).

3. OBSERVATIONALLY EQUIVALENT ABCD-REPRESENTATIONS

An important part of our identification analysis is based on the approach and results from [Komunjer and Ng \(2011\)](#). Since we closely follow their framework, we will be as brief as possible and refer the readers to this paper for necessary details. First of all, we confine ourselves to the ABCD-representations with the number of observables greater or equal to the number of shocks, that is, $k \leq r$.⁴ This covers the most interesting special case of the square model (when $k = r$). Let us formalize the assumption concerning stability.

ASSUMPTION 1. *For every $\theta \in \Theta$ and for any $z \in \mathbb{C}$ (a set of complex numbers), $\det(zI_n - A(\theta)) = 0$ implies $|z| < 1$.*

Let us next define the transfer function as $H(z; \theta) = D(\theta) + C(\theta)(zI_n - A(\theta))^{-1}B(\theta)$. This allows us to write the following.

ASSUMPTION 2. *For every $\theta \in \Theta$, $|z| > 1$ implies $\text{rank}(H(z; \theta)) = k$, that is, the transfer function has full column rank.*

To state the final assumption, we need to define the observability matrix $O = [C(\theta)'; A(\theta)'C(\theta)'; A(\theta)'^2C(\theta)'; \dots; A(\theta)^{n-1}C(\theta)']'$ and the controllability matrix $K = [B(\theta); A(\theta)B(\theta); A(\theta)^2B(\theta); \dots; A(\theta)^{n-1}B(\theta)]$. Then we have our last assumption.

ASSUMPTION 3. *For every $\theta \in \Theta$, matrices O and K have full column and row rank, respectively, that is, $\text{rank}(O) = \text{rank}(K) = n$.*

A full discussion and interpretation of the above assumptions can be found in [Komunjer and Ng \(2011\)](#). Now we are in a position to define the observational equivalence. Let us define the spectral density of the ABCD-representation as $\Phi(z; \theta) = H(z; \theta)\Sigma(\theta)H'(z^{-1}; \theta)$, where z^{-1} corresponds to the backward shift in the transfer function. Then we have the following definition.

DEFINITION 1. θ and $\bar{\theta}$ are observationally equivalent (written as $\theta \sim \bar{\theta}$) if $\Phi(z; \theta) = \Phi(z; \bar{\theta})$ for all $z \in \mathbb{C}$.

The key result in [Komunjer and Ng \(2011\)](#) is that, in some circumstances, $\theta \sim \bar{\theta}$ if and only if the resultant $A(\bar{\theta})$, $B(\bar{\theta})$, $C(\bar{\theta})$, $D(\bar{\theta})$, $\Sigma(\bar{\theta})$ are related to $A(\theta)$, $B(\theta)$, $C(\theta)$, $D(\theta)$, $\Sigma(\theta)$ by the so-called similarity transformation, which is defined by the following theorem.

⁴Accommodating the case when the number of shocks is greater than the number of observables would require using innovations representation of the state space model; see [Komunjer and Ng \(2011\)](#). In this case, the simple link between matrices A , B , C , D , and semistructural parameters, which we develop in the next section and on which our approach hinges, breaks down.

THEOREM 1 (Komunjer and Ng (2011)). *Let $k \leq r$, $\varepsilon_t \sim i.i.d. N(0, \Sigma(\theta))$ with $\Sigma(\theta)$ positive definite for every $\theta \in \Theta$, and Assumptions 1 to 3 hold. Then $\theta \sim \bar{\theta}$ if and only if (1) $A(\bar{\theta}) = TA(\theta)T^{-1}$, (2) $B(\bar{\theta}) = TB(\theta)U$, (3) $C(\bar{\theta}) = C(\theta)T^{-1}$, (4) $D(\bar{\theta}) = D(\theta)U$, (5) $\Sigma(\bar{\theta}) = U^{-1}\Sigma(\theta)U'^{-1}$, for some nonsingular $n \times n$ matrix T and nonsingular $k \times k$ matrix U .*

Note that, of the four matrices characterizing the DSGE model solution (3) and (4), only $A(\theta)$ and $B(\theta)$ appear in Theorem 1. In particular, lack of restrictions linking $F(\bar{\theta})$ to $F(\theta)$ and $G(\bar{\theta})$ to $G(\theta)$ will have some consequences on the way we develop our identification framework for the deep model parameters in the next section.

If we assume that shocks ε_t are independent, that is, $\Sigma(\theta) = I_k$,⁵ then we have the useful corollary.

COROLLARY 1. *Let $k \leq r$ and Assumptions 1 to 3 hold. Moreover, assume $\varepsilon_t \sim i.i.d. N(0, I_k)$. Then $\theta \sim \bar{\theta}$ if and only if (1) $A(\bar{\theta}) = TA(\theta)T^{-1}$, (2) $B(\bar{\theta}) = TB(\theta)V$, (3) $C(\bar{\theta}) = C(\theta)T^{-1}$, (4) $D(\bar{\theta}) = D(\theta)V$, for some nonsingular $n \times n$ matrix T and orthogonal $k \times k$ matrix V .*

The concept of observational equivalence is inherently related to global identification.

DEFINITION 2. The ABCD-representation is globally identified (from the spectral density) at $\theta \in \Theta$ if and only if $\theta \sim \bar{\theta} \Rightarrow \theta = \bar{\theta}$.

Since $\varepsilon_t \sim i.i.d. N(0, \Sigma(\theta))$ and under Assumption 1, we stay in the stationary Gaussian environment. Hence, the underlying distribution of observables is fully characterized by their unconditional mean and autocovariances. If the former is zero, then Definition 2 is equivalent to the standard notion of global identification expressed in terms of probability distribution; see, for example, [Rothenberg \(1971\)](#). However, in some economic applications the measurement equation of the ABCD-representation may contain a linear deterministic component that depends on (a subvector of) θ . This directly contributes to the unconditional mean of the observables. In such a case, which we do not consider in this paper, the first moments convey additional information concerning identifiability and Definition 2 should be modified accordingly. See Definition 3 in [Qu and Tkachenko \(2017\)](#) for some clarification on how to do this.

4. IDENTIFICATION OF DEEP PARAMETERS

From now on, we use the generic simplifying notation $X := X(\theta)$, where X is a matrix that depends on θ . Analogously, when referring to other points in the deep parameter space, we write $\bar{X} := X(\bar{\theta})$. Bearing in mind Theorem 1 and Definition 2, checking global identification at θ boils down to checking if there exist $\bar{\theta} \neq \theta$ that results in an ABCD-representation with \bar{A} , \bar{B} , \bar{C} , \bar{D} , and $\bar{\Sigma}$ such that $\bar{A} = TAT^{-1}$; $\bar{B} = TBU$; $\bar{C} = CT^{-1}$;

⁵If shocks are independent, then it is always possible to normalize matrix $\Sigma(\theta)$ to identity by writing the model given by (1) such that the standard deviations of shocks show up in matrix $\Gamma_3(\theta)$. This is what we do in our example studied in Section 7.

$\bar{D} = DU$; $\bar{\Sigma} = U^{-1}\Sigma U'^{-1}$ for some nonsingular matrices T and U . It would be straightforward to do it if \bar{A} , \bar{B} , \bar{C} , \bar{D} , and $\bar{\Sigma}$ were all explicit functions of $\bar{\theta}$, that is, if the model solution was analytical. Then we would just need to solve the above system of equations for T , U , and $\bar{\theta}$. If the model cannot be solved analytically, as it is mostly the case, using Theorem 1 to check identification becomes impractical since we would need to compute numerically the model solution for every candidate $\bar{\theta}$.

To circumvent this problem, we use the links between the model's semistructural parameters $\Gamma_0, \Gamma_1, \Gamma_2, \Gamma_3, H, J$, the model solution given by matrices A, B, F, G , and the remaining two objects of the ABCD-representation, that is, matrices C and D . Two of such links are already available as equations (6) and (7). To find other inherent relations, we exploit the method of undetermined coefficients used, for example, by Uhlig (1999). To this end, let us partition Γ_0 and Γ_1 in (1) into blocks corresponding to the state and policy variables

$$\begin{bmatrix} \Gamma_0^s & \Gamma_0^p \\ \Gamma_1^s & \Gamma_1^p \end{bmatrix} \begin{bmatrix} s_t \\ p_t \end{bmatrix} = \begin{bmatrix} \Gamma_1^s & \Gamma_1^p \end{bmatrix} E_t \begin{bmatrix} s_{t+1} \\ p_{t+1} \end{bmatrix} + \Gamma_2 s_{t-1} + \Gamma_3 \varepsilon_t. \quad (8)$$

Plugging the model solution (3)–(4) into (8) twice and using $E_t \varepsilon_{t+1} = 0$ yields

$$(\Gamma_0^s A + \Gamma_0^p F - \Gamma_1^s A^2 - \Gamma_1^p F A - \Gamma_2) s_{t-1} = (\Gamma_1^s A B + \Gamma_1^p F B + \Gamma_3 - \Gamma_0^s B - \Gamma_0^p G) \varepsilon_t. \quad (9)$$

Equation (9) has to hold for all s_{t-1} and ε_t , hence the coefficient matrices on these two terms must be equal to zero.⁶ This gives us two matrix equation restrictions:

$$\Gamma_0^s A + \Gamma_0^p F - \Gamma_1^s A^2 - \Gamma_1^p F A = \Gamma_2, \quad (10)$$

$$\Gamma_1^s A B + \Gamma_1^p F B - \Gamma_0^s B + \Gamma_3 = \Gamma_0^p G. \quad (11)$$

Suppose we check global identification at some $\theta \in \Theta$. Given our restrictions on Θ (determinacy region), we have a unique set A, B, F , and G that must be consistent with (10) and (11). From equations (6) and (7), we also get unique C and D . Similarly, to each observationally equivalent $\bar{\theta}$ there correspond unique $\bar{A}, \bar{B}, \bar{F}, \bar{G}, \bar{C}$, and \bar{D} that must conform to equations (10), (11), (6), and (7). We also know from Theorem 1 that the observationally equivalent ABCD-representations must be related to each other with the similarity transformation, and this is the other key ingredient of our procedure. If we write restrictions (10), (11), (6), and (7) for $\bar{\theta}$, that is, add overbars to all matrices showing up in these four sets of equations, and next merge them with the restrictions from Theorem 1 in a way that eliminates $\bar{A}, \bar{B}, \bar{C}$, and \bar{D} , we arrive at the following system of equations:

$$\bar{\Gamma}_0^s T A T^{-1} + \bar{\Gamma}_0^p \bar{F} - \bar{\Gamma}_1^s T A^2 T^{-1} - \bar{\Gamma}_1^p \bar{F} T A T^{-1} = \bar{\Gamma}_2, \quad (12)$$

$$\bar{\Gamma}_1^s T A B U + \bar{\Gamma}_1^p \bar{F} T B U - \bar{\Gamma}_0^s T B U + \bar{\Gamma}_3 = \bar{\Gamma}_0^p \bar{G}, \quad (13)$$

$$C T^{-1} = \bar{H}^s T A T^{-1} + \bar{H}^p \bar{F}, \quad (14)$$

⁶An implicit assumption justifying this claim is that the dimension (in linear algebra sense) of ε_t is k and that of s_t is n . The former is satisfied since Σ is positive definite and the latter holds by Assumption 3.

$$DU = \bar{H}^s T B U + \bar{H}^p \bar{G} + \bar{J}, \quad (15)$$

$$\bar{\Sigma} = U^{-1} \Sigma U'^{-1}. \quad (16)$$

These equations jointly define all parameter vectors $\bar{\theta}$ that are observationally equivalent to a given θ .

Note that, for given θ , matrices A , B , C , D , and Σ are all known, while matrices $\bar{\Gamma}_0$, $\bar{\Gamma}_1$, $\bar{\Gamma}_2$, $\bar{\Gamma}_3$, \bar{H} , \bar{J} , and $\bar{\Sigma}$ are analytically linked to $\bar{\theta}$. Hence, using some θ of interest, checking identification at this point amounts to searching for $\bar{\theta} \in \Theta$, nonsingular T and U , as well as \bar{F} and \bar{G} , that solve the system of equations (12) to (16). If we can find a solution such that $\bar{\theta} \neq \theta$, the DSGE model is not globally identified at θ . The efficiency of such a procedure when applied even to large DSGE models follows from the fact that, unlike Qu and Tkachenko (2017), we do not have to solve the model for each candidate deep parameter point to check whether a model is globally identified. Instead, we automatically connect the deep parameters with the model solution through the system of nonlinear equations (12) to (16). As a result, our procedure is faster to implement, and probably also numerically more accurate, because each time the model is solved, a potential numerical inaccuracy is introduced.

Before proceeding, it is worth discussing the role of matrices \bar{F} and \bar{G} in our identification conditions. They show up in equations (12)–(16) because the original problem (1)–(2) features policy variables p_t , while the ABCD-representation does not. Consequently, we cannot eliminate these matrices using Theorem 1, as we did in the case of \bar{A} , \bar{B} , \bar{C} , and \bar{D} . But we know that any unique stable solution for p_t must be of form (4), and by plugging it into (1)–(2) we obtained implicit restrictions on \bar{F} and \bar{G} so that they are consistent with the original problem for given $\bar{\theta}$. These restrictions are sufficient to make sure that, when a solution to the system of equations (12)–(16) is found, the obtained \bar{F} and \bar{G} (as well as \bar{A} , \bar{B} , and $\bar{\Sigma}$, via their links to A , B , and Σ transmitted by T and U according to Theorem 1) describe the unique stable solution for $\bar{\theta}$. This is because, by Assumption 1, matrix $\bar{A} = T A T^{-1}$ is stable so that unstable solutions are ruled out, while uniqueness results from the restrictions on the parameter space Θ to which $\bar{\theta}$ must belong. As we will show later, it is very often possible to solve for F and G as explicit functions of matrices A , B , C , D , and semistructural parameters. This will speed up the algorithm that we describe in the next section.

5. ALGORITHM CHECKING IDENTIFICATION

In this section, we develop an algorithm that checks global identification using the conceptual framework described above.

5.1 General setup

For a typical DSGE model, the system given by equations (12)–(16) is nonlinear in $\bar{\theta}$, and hence numerical methods have to be applied to solve it. Using the fact that, by construction, the system is satisfied for $\bar{\theta} = \theta$, $\bar{F} = F$, $\bar{G} = G$, $T = I_n$, and $U = I_k$, one can generate the starting values for a numerical solver by randomizing around this point. Naturally,

the more diffuse distribution is used in the randomization, the longer it might take to obtain a solution, but the more likely it is to eventually find alternative parameter values that are far from θ at which we check identification.

A simple implementation of our procedure to check global identification can be described in form of the following algorithm.

ALGORITHM 1. *To search for observationally equivalent vectors of deep parameters of a DSGE model defined by equations (1) and (2):*

1. *Choose a point $\theta \in \Theta$ at which global identification is to be checked.*
2. *Evaluate Σ . Solve the model to obtain matrices A , B , F , and G . Calculate matrices C and D from equations (6) and (7).*
3. *Generate a vector of starting values $[(\bar{\theta}^0)' : \text{vec}(\bar{F}^0)' : \text{vec}(\bar{G}^0)' : \text{vec}(T^0)' : \text{vec}(U^0)']' = [\theta' : \text{vec}(F)' : \text{vec}(G)' : \text{vec}(I_n)' : \text{vec}(I_k)']' + \epsilon$, where $\epsilon \sim N(0, \text{diag}(\Xi))$ is a randomizing distribution and $\Xi = [v_i^2]_{i=1}^{m+qn+qk+n^2+k^2}$.*
4. *Using the starting values from step 3, solve numerically the system of equations (12)–(16) for $\bar{\theta}$, T , U , \bar{F} , and \bar{G} . If no solution such that $\bar{\theta} \in \Theta$ can be found, go back to step 3.⁷*
5. *If $\bar{\theta} \neq \theta$, stop and conclude that the model is not globally identified at θ . If $\bar{\theta} = \theta$, go back to step 3 or stop if the total number of solutions found has reached M .*

The algorithm is parameterized by two objects: Ξ and M . The first one characterizes the randomizing distribution used to generate the starting values for the numerical routine that is applied to solve equations (12)–(16). The second controls the number of performed model solutions that we consider sufficient to judge if the model is globally identified. When global identification fails due to lack of local identification, $M = 1$ is sufficient for any nonnegligible scale of randomization Ξ . However, if the model is locally identified, larger values of M might be needed to detect potential problems with global identification, especially if the elements in Ξ are small, and hence the starting values lie close to the solution $\bar{\theta} = \theta$. Too large values in Ξ might in turn produce starting values so far away from any solution to the system of equations (12)–(16) that the numerical solver fails to find any, and step 4 has to be repeated many times.

More generally, the optimal values of Ξ and M are model-specific and some fine-tuning for particular applications might be needed. However, we found the following procedure for calibrating Ξ to work very well in our experiments. As a first pass, use relatively large scale of randomization v_i for all deep model parameters (i.e., for $i = 1, \dots, m$)

⁷Recall that our analysis deals only with those model parametrizations that imply a unique stable solution, which imposes restrictions on the parameter space Θ . In practice, and in particular for model in which the determinacy region is cumbersome to define, it may be more convenient to leave Θ unrestricted while solving equations (12)–(16), that is, assume $\Theta = \mathbb{R}^m$, and then go back to step 3 if the obtained solution for $\bar{\theta}$ does not satisfy the Blanchard–Kahn conditions.

so that $\bar{\theta}^0$ may be far away from θ ,⁸ but do not randomize the starting values for the remaining unknowns (i.e., set $v_i = 0$ for $i > m$). If no observationally equivalent parameter is found, experiment with other parameterizations of the randomizing distribution by gradually increasing v_i 's for $i > m$. As regards M , offering any concrete recommendation is much more tricky since the probability of finding $\bar{\theta} \neq \theta$ (if it exists) clearly depends on the choice of Ξ ; see Section 6 for illustration. It is also reasonable to expect that larger values of M might be needed for larger dimensions of the parameter space involved. In all examples considered in this paper, we set $M = 1000$, and, for the calibration of Ξ described above, this made the probability of failing to detect identification problems, if such exist, very small. Naturally, there is no guarantee that this number will be sufficiently high in other applications, especially in bigger models.

Several steps in Algorithm 1 require application of numerical methods, which necessarily result in numerical approximation errors. We perform all calculations in Matlab R2015a, which uses double precision numbers. To solve the system of equations (12)–(16), we use the Matlab routine `fsolve`, setting the tolerance levels for the size of step `TolX` and change in the value of function `TolFun` at a very low value of 10^{-20} as such high precision did not significantly affect the speed of computations compared to more conventional and less stringent values.⁹ To judge if $\bar{\theta}$ equals θ , we compare them up to the fifth decimal place, which is far more accurate than required by any meaningful economic application.

It needs to be stressed once again that the recommendations formulated above about the tuning constants and tolerance levels are based on our personal experience gained by experimenting with the models presented as examples below, and they may not necessarily work equally well with other applications.

5.2 Useful special cases

A couple of further remarks that might be helpful to specialize or simplify Algorithm 1 are in order. First of all, if shocks are independent so that the model conforms to Corollary 1, we suggest to parameterize an orthogonal V by $V = 2(I_k + X)^{-1} - I_k$, where X is a $k \times k$ skew symmetric matrix, that is, it satisfies $X + X' = 0$. Second, as we have explained before, while checking global identification using Algorithm 1, we have to solve the system of equations (12)–(16) not only for $\bar{\theta}$, T , and U , but also for \bar{F} and \bar{G} . We now

⁸For example, in the models presented in the rest of this paper, setting $v_i = 1$ for $i = 1, \dots, m$ implies the range of variation in $\bar{\theta}^0$ that well covers all reasonable values of the deep model parameters.

⁹Note that matrices A , B , C , and D that enter Algorithm 1 as parameters can usually be obtained only numerically. Therefore, if the tolerance levels `TolX` and `TolFun` are too small, `fsolve` may fail to find a solution $\bar{\theta} \neq \theta$ even if the model is not identified. However, in such cases it is likely that the routine will have problems to find the solution $\bar{\theta} = \theta$ as well. Therefore, while setting the tolerance levels it is recommended to make sure that this solution is easily found for a small degree of randomization, that is, for $\bar{\theta}^0$ that is close to θ . Notwithstanding these concerns, our experiments with the models presented in the rest of this paper suggest that problems related to inappropriate calibration of the tolerance levels are unlikely to emerge unless one sets them at extremely low levels.

show how it is possible to eliminate these two matrices from the list of unknowns by exploiting the model structure summarized in equation (1).¹⁰

We first consider only those of the model equilibrium conditions that do not involve taking expectations of policy variables. This can be often sufficient to recover policy matrices F and G from the model semistructural parameters and matrices A and B , and hence to get rid of \bar{F} and \bar{G} from our algorithm. Let us write such a subset of l equations in (1) as

$$\begin{bmatrix} \Phi_0^s & \Phi_0^p \end{bmatrix} \begin{bmatrix} s_t \\ p_t \end{bmatrix} = \Phi_1^s E_t s_{t+1} + \Phi_2 s_{t-1} + \Phi_3 \varepsilon_t. \quad (17)$$

Plugging the model solution (3) into (17) and using $E_t s_{t+1} = A s_t$ yields

$$\Phi_0^p p_t = (\Phi_2 - \Phi_0^s A + \Phi_1^s A^2) s_{t-1} + (\Phi_3 + \Phi_1^s AB - \Phi_0^s B) \varepsilon_t. \quad (18)$$

If Φ_0^p has full column rank q , then we can select q equations in (18) such that the corresponding block of Φ_0^p is nonsingular. If we use tilde to denote the appropriate blocks of matrices defined by this selection, then juxtaposing (18) and (4) we obtain

$$F = (\tilde{\Phi}_0^p)^{-1} (\tilde{\Phi}_2 - \tilde{\Phi}_0^s A + \tilde{\Phi}_1^s A^2), \quad (19)$$

$$G = (\tilde{\Phi}_0^p)^{-1} (\tilde{\Phi}_3 + \tilde{\Phi}_1^s AB - \tilde{\Phi}_0^s B). \quad (20)$$

Of course, it still may be the case that, even if $l > q$, the rank of Φ_0^p is less than q . In practice, this can happen if at least one of the policy variables does not enter the subsystem (17). Then, without loss of generality, we can write $\Phi_0^p = [\hat{\Phi}_0^p \ 0]$, where $\hat{\Phi}_0^p$ has full column rank $\hat{q} < q$, and accordingly partition $p_t = [(\hat{p}_t)' (\check{p}_t)']'$. We can retrieve the first \hat{q} rows of F and G , that is, the policy functions for \hat{p}_t , using the same arguments that led to (19) and (20).

Importantly, the missing rows of F and G (i.e., those corresponding to \check{p}_t) can still often be recovered using additional model equations in (1), and distinct from (17), with the following structure:

$$\begin{bmatrix} \Psi_0^s & \hat{\Psi}_0^p & \check{\Psi}_0^p \end{bmatrix} \begin{bmatrix} s_t \\ \hat{p}_t \\ \check{p}_t \end{bmatrix} = \Psi_1^s E_t s_{t+1} + \hat{\Psi}_1^p E_t \hat{p}_{t+1} + \Psi_2 s_{t-1} + \Psi_3 \varepsilon_t \quad (21)$$

that is, those equilibrium conditions where the subset of policy variables \hat{p}_t enter in expectations, but the complementary subset \check{p}_t does not. Provided that $\check{\Psi}_0^p$ has full column rank $q - \hat{q}$, we can essentially repeat the above reasoning. That is, using (21), we obtain the rows of F and G corresponding to \check{p}_t by plugging into it $\hat{p}_t = \hat{F} s_{t-1} + \hat{G} \varepsilon_t$ and $E_t \hat{p}_{t+1} = \hat{F} A s_t + \hat{F} B \varepsilon_t$, where \hat{F} and \hat{G} denote the policy functions for \hat{p}_t recovered at the previous stage, and proceeding as before.

¹⁰Naturally, matrices \bar{F} and \bar{G} could be eliminated just by computing them directly using one of the routines solving dynamic linear models with rational expectations. However, that would defeat the purpose of Algorithm 1, which is to avoid solving the model for each candidate parameter vector.

The procedure derived above can be used in Algorithm 1 so that it does not need to search for all elements in \bar{F} and \bar{G} , and sometimes the set of unknowns can be reduced just to $\bar{\theta}$, T , and U . To give more intuition, in the Supplementary Material (Kocięcki and Kolasa (2018)), we show how entire policy matrices F and G can be recovered in the Smets and Wouters (2007) model, using the reasoning developed above or, alternatively, by working directly with the model equilibrium conditions on an equation-by-equation basis.

An alternative or complementary way of eliminating matrices F and G to the one exploiting the model equilibrium conditions is to use restrictions (6)–(7) associated with the measurement equations, for which we need matrix H^p to have full column rank. If H is just a selection matrix, this condition states that all policy variables must be treated as observable. This is the case, for example, in the canonical New Keynesian model of Clarida, Gertler, and Galí (1999) or its closely related variant by An and Schorfheide (2007) that we consider in Section 7, where the observable variables are output, inflation, and the interest rate. It is also worth noting that full column rank of H^p can sometimes be achieved simply by using some equations in (1) to eliminate unobservable policy variables.

6. EXAMPLE: SIMPLE A-THEORETICAL STATE-SPACE SYSTEM

Before we demonstrate how our method can be applied to a fully-fledged DSGE model, it is instructive to consider a simple a-theoretical case, in which the mapping from structural parameters to the ABCD-representation can be derived analytically. This will allow us to see how global identification problems may arise and how the framework we propose can be useful in detecting them. As an example, we use a model that results in a similar state-space system as discussed in Section 3.1 of Schorfheide (2011).

More specifically, consider the following artificial 3-equation DSGE model:

$$s_{1,t} = \alpha_1^2 s_{1,t-1} + \varepsilon_t, \quad (22)$$

$$s_{2,t} = (1 - \alpha_1^2)(s_{1,t-1} + s_{2,t-1}) - \alpha_2 p_t, \quad (23)$$

$$p_t = \alpha_2 E_t p_{t+1} - (\alpha_1^2 \alpha_2 - 1) s_{1,t}, \quad (24)$$

where the vector of states is $s_t = [s_{1t} \ s_{2t}]'$ and the only policy variable is p_t . There is only one stochastic shock $\varepsilon_t \sim N(0, 1)$ so that $\Sigma = 1$. The vector of deep model parameters is $\theta = [\alpha_1 \ \alpha_2]'$, with each of the two elements assumed to lie in the unit interval, that is, $0 < \alpha_i < 1$ for $i = 1, 2$.

The system of equations (22)–(24) is already written in the form that is conformable with (1), where

$$\Gamma_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \alpha_2 \\ \alpha_1^2 \alpha_2 - 1 & 0 & 1 \end{bmatrix}, \quad \Gamma_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha_2 \end{bmatrix},$$

$$\Gamma_2 = \begin{bmatrix} \alpha_1^2 & 0 \\ 1 - \alpha_1^2 & 1 - \alpha_1^2 \\ 0 & 0 \end{bmatrix}, \quad \Gamma_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

We assume that the observable variable is a simple sum of the two states, that is, $y_t = s_{1t} + s_{2t}$, and there is no measurement error so that

$$H = [1 \quad 1 \quad 0], \quad J = 0.$$

Given the restrictions on θ , there exists a unique stable solution. Bearing in mind Section 2, the relevant matrices are

$$A = \begin{bmatrix} \alpha_1^2 & 0 \\ 1 - \alpha_1^2 - \alpha_1^2 \alpha_2 & 1 - \alpha_1^2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ -\alpha_2 \end{bmatrix}, \quad C = [1 - \alpha_1^2 \alpha_2 \quad 1 - \alpha_1^2],$$

$$D = 1 - \alpha_2, \quad F = [\alpha_1^2 \quad 0], \quad G = 1.$$

Since in this example we know the mapping from θ to the ABCD-representation, we can check global identification simply by solving analytically the five restrictions included in Theorem 1.¹¹ According to this solution, which we present in the Supplementary Material (Kocięcki and Kolasa (2018)), there exists one observationally equivalent parametrization to θ , that is, $\bar{\theta} = [(1 - \alpha_1^2)^{\frac{1}{2}} \alpha_2]'$. Hence, of the two model parameters, α_2 is globally identified while α_1 is not, unless $\alpha_1 = 2^{-\frac{1}{2}}$. Summing up, the model is not globally identified at almost every θ .¹²

Suppose now that the analytical solution, that is, the mapping from structural parameters α_1 and α_2 to the coefficients in the A , B , C , and D matrices, is not available. How can Algorithm 1 be useful to an applied researcher? We illustrate it by showing how the algorithm's output depends on the choice of the starting values. More specifically, let us check identification at $\theta = [0.3 \ 0.2]$ so that we know that $\bar{\theta} = [0.91^{\frac{1}{2}} \ 0.2]$ implies exactly the same likelihood for any possible observable variable y_t . We consider a grid for the starting values of the deep model parameters $\bar{\theta}^0$ defined as a Cartesian square of a set $\{0.001 \{\frac{i}{10}\}_{i=1}^9 \ 0.999\}$. For each point on the grid, we solve the system of equations (12)–(16) 10,000 times using independent draws from the standard multivariate normal distribution to randomize the starting values for the remaining unknowns, that is, the elements of \bar{F} , \bar{G} , T , and U .

The outcomes are presented in the first panel of Figure 1 in the form of a heat map, where the temperature is defined as the frequency of finding the alternative solution $\bar{\theta} = [0.91^{\frac{1}{2}} \ 0.2]$. Not surprisingly, the closer are the starting values $\bar{\theta}^0$ to $\bar{\theta}$, the higher is the chance that we will find this point. While the probability of detecting global identification failure if the system of equations (12)–(16) is solved just once stands well below 0.5, even if the starting values are close to $\bar{\theta}$, the chances of missing it after many (say,

¹¹Taking into account the restrictions on Θ , one can easily check that Assumptions 1 to 3 hold.

¹²Proceeding similar to Schorfheide (2011), lack of global identification in the example considered in this section can also be easily established by noting that y_t is a restricted AR(2) process

$$(1 - \alpha_1^2 L)(1 - (1 - \alpha_1^2)L)y_t = (1 - \alpha_2)\varepsilon_t,$$

where L is a lag operator such that $L^s x_t = x_{t-s}$ for any variable x and $s = 0, 1, \dots$. By switching the values of the two AR roots, we can obtain an observationally equivalent AR(2) process.

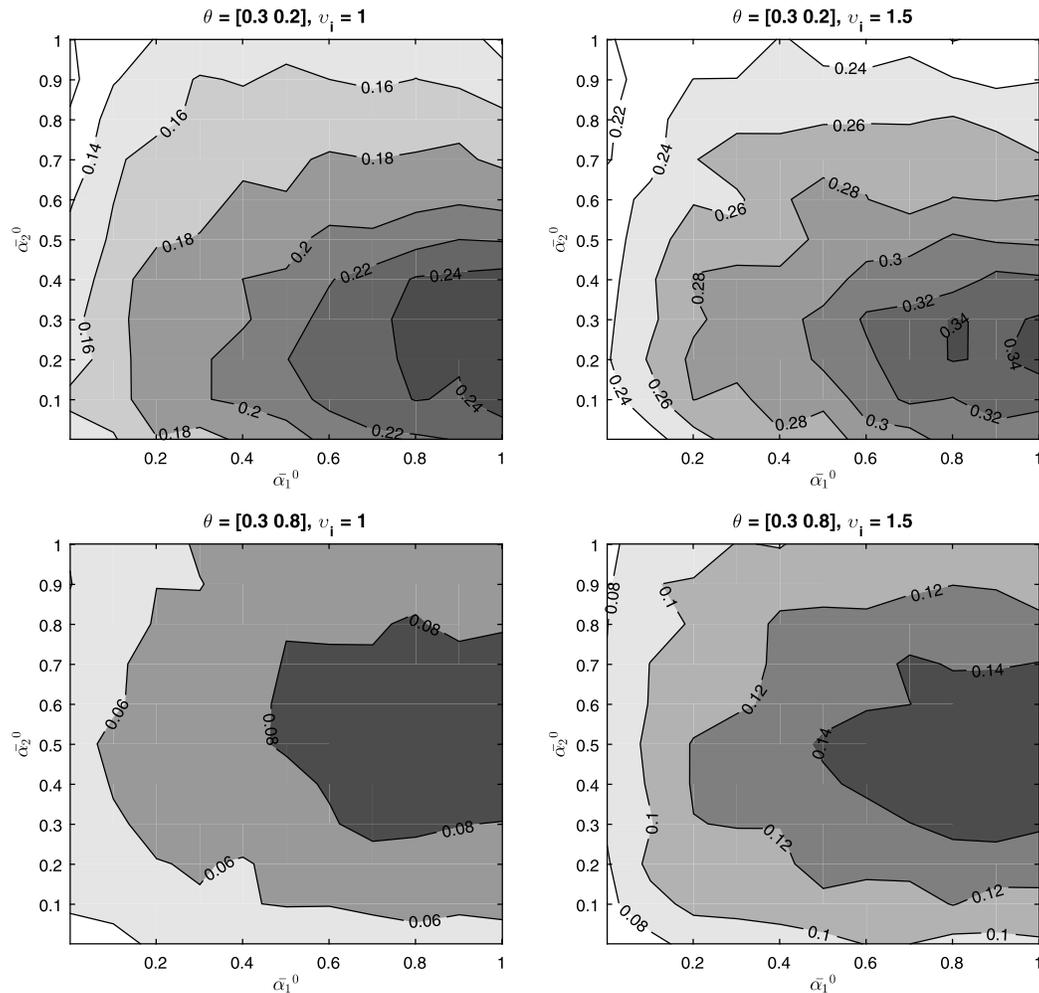


FIGURE 1. Identification of simple a-theoretical state-space system. Notes: The figures illustrate the frequency of finding the alternative solution $\bar{\theta} \neq \theta$ with Algorithm 1 applied to the simple a-theoretical state-space system, depending on the starting values $\bar{\theta}^0$. The outcomes are presented for two values of the benchmark parameter vector θ (in rows), and two variants of the dispersion in the randomizing distribution Ξ (in columns) used to generate the starting values for the remaining unknowns in system (12)–(16), that is, \bar{F}^0 , \bar{G}^0 , T^0 , and U^0 . For each variant and each point on the grid for $\bar{\theta}^0$, the frequencies are estimated using 10.000 independent draws from the randomizing distribution.

$M = 1000$) tries are negligibly small. The second panel in the first row repeats this exercise with the dispersion of the randomizing distribution increased by 50%. The frequencies are significantly larger for all starting values $\bar{\theta}^0$. The downside of higher dispersion is that it may take longer for the numerical solver to find any solution, so it is not necessarily the case that larger v_i 's allow to check global identification in a shorter time. In the second row of Figure 1, we replicate our results for an alternative value of $\theta = [0.3 \ 0.8]$,

which differs from the one previously used in the value of α_2 . The frequencies of finding $\bar{\theta} = [0.91^{\frac{1}{2}} \ 0.8]$ are now smaller, so it will usually take more time for Algorithm 1 to run until we conclude that the model is not identified.¹³ Clearly, the time after which our procedure detects global identification failure depends on the point at which we check it.

Overall, this simple example shows us that running Algorithm 1 on a DSGE model before estimating it can give the researcher valuable information about possible problems and caveats that he or she will face. In particular, if Bayesian methods are applied, one can expect high sensitivity of the results to the prior assumptions and multimodality in the posterior distribution, the latter calling for a careful choice of the sampling algorithm.

7. EXAMPLE: AN-SCHORFHEIDE MODEL

Our next example is a fully-fledged, but still small DSGE model by [An and Schorfheide \(2007\)](#), AS henceforth, modified to allow for correlation between government spending and productivity as in [Herbst and Schorfheide \(2016\)](#). Identification of the original AS model was also examined locally by [Komunjer and Ng \(2011\)](#) so it provides a natural object to highlight the main features of our approach.

7.1 Model summary

When written in log-linearized form, the model is given by the following equations:

$$z_t = \rho_z z_{t-1} + \rho_{zg} g_{t-1} + \sigma_z \varepsilon_{z,t}, \quad (25)$$

$$g_t = \rho_g g_{t-1} + \rho_{gz} z_{t-1} + \sigma_g \varepsilon_{g,t}, \quad (26)$$

$$x_t = E_t x_{t+1} + g_t - E_t g_{t+1} - \frac{1}{\tau} (R_t - E_t \pi_{t+1} - E_t z_{t+1}), \quad (27)$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa (x_t - g_t), \quad (28)$$

$$R_t = \rho_m R_{t-1} + (1 - \rho_m) [\psi_1 \pi_t + \psi_2 (x_t - g_t)] + \sigma_m \varepsilon_{m,t}. \quad (29)$$

There are three endogenous variables in the model: detrended output x_t , inflation π_t , and the interest rate R_t . They are driven by two exogenous AR(1) processes for productivity growth z_t and government spending g_t , with innovations $\varepsilon_{z,t}$ and $\varepsilon_{g,t}$, respectively, and by an i.i.d. monetary policy shock $\varepsilon_{m,t}$. All of these i.i.d. innovations are assumed to be mutually uncorrelated so that we can normalize $\Sigma = I_3$. The 13-dimensional vector of parameters is $\theta = [\tau \ \beta \ \kappa \ \psi_1 \ \psi_2 \ \rho_z \ \rho_{zg} \ \rho_g \ \rho_{gz} \ \rho_m \ \sigma_z \ \sigma_g \ \sigma_m]'$.

¹³To see why this is the case, it is instructive to use the analytical form of T derived in the Supplementary Material. Evaluating it at $\theta = [0.3 \ 0.2]$ gives $T = \begin{bmatrix} 1.23 & 1.13 \\ -0.23 & -0.12 \end{bmatrix}$, while for $\theta = [0.3 \ 0.8]$ we have $T = \begin{bmatrix} 4.60 & 4.51 \\ -3.60 & -3.51 \end{bmatrix}$. Clearly, the latter is further away from the identity matrix around which we center randomization, so it takes a larger number of draws to find the solution $\hat{\theta} \neq \theta$.

The model can be cast in the form given by (1), with states $s_t = [z_t g_t R_t]'$, policy variables $p_t = [x_t \pi_t]'$, shocks $\varepsilon_t = [\varepsilon_{z,t} \varepsilon_{g,t} \varepsilon_{m,t}]'$, and matrices Γ_0 , Γ_1 , Γ_2 , and Γ_3 given by

$$\Gamma_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & \frac{1}{\tau} & 1 & 0 \\ 0 & \kappa & 0 & -\kappa & 1 \\ 0 & (1 - \rho_m)\psi_2 & 1 & (\rho_m - 1)\psi_2 & (\rho_m - 1)\psi_1 \end{bmatrix},$$

$$\Gamma_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\tau} & -1 & 0 & 1 & \frac{1}{\tau} \\ 0 & 0 & 0 & 0 & \beta \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\Gamma_2 = \begin{bmatrix} \rho_z & \rho_{zg} & 0 \\ \rho_{gz} & \rho_g & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \rho_m \end{bmatrix}, \quad \Gamma_3 = \begin{bmatrix} \sigma_z & 0 & 0 \\ 0 & \sigma_g & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_m \end{bmatrix}.$$

The vector of observable variables is $y_t = [R_t x_t \pi_t]'$ and there are no measurement errors, which means that $H = [0_{3 \times 2} \ I_3]$ and $J = [0_{3 \times 3}]$ for all θ .

Our benchmark parametrization θ of the AS model is the same as used by [Komunjer and Ng \(2011\)](#) and presented in Table 1.

7.2 Local identification issues

In the original AS model, government spending and productivity are assumed to be independent, that is, $\rho_{zg} = \rho_{gz} = 0$. We know from the earlier literature that this setup is locally unidentified; see [Komunjer and Ng \(2011\)](#) or [Qu and Tkachenko \(2012\)](#). Applying Algorithm 1 confirms these findings. For instance, if we reparameterize the Taylor rule

TABLE 1. Benchmark parametrization of the AS model.

Parameter	Value	Parameter	Value
τ	2	ρ_{zg}	0
β	0.9975	ρ_{gz}	0
κ	0.33	ρ_m	0.75
ψ_1	1.5	σ_z	0.003
ψ_2	0.125	σ_g	0.006
ρ_z	0.9	σ_m	0.002
ρ_g	0.95		

such that $\psi_1 = 2$, $\psi_2 = -0.67$, $\rho_m = 0.688$, and $\sigma_m = 0.0018$, we get exactly the same likelihood (and moments) for observables as for the benchmark calibration from Table 1. To make the model identified, we need to fix (i.e., treat as known) at least one parameter from the set $[\psi_1 \ \psi_2 \ \rho_m \ \sigma_m]$.

Even though these results are not new, let us highlight an important strength of Algorithm 1, also in this local context. Komunjer and Ng (2011) show that if they reduce the degree of inertia in productivity and government spending shocks (i.e., ρ_z and ρ_g) to about 0.1, their method fails to confirm identification of the model, even if one of the Taylor rule parameters is restricted. More generally, their method is sensitive to the tolerance level they set while performing the matrix rank tests. As we signaled in footnote 9, our algorithm seems to be more robust to numerical difficulties. In particular, if we consider low (but nonzero) shock inertia, our identification analysis leads to the same conclusions as for the benchmark parameter set.

7.3 Global identification issues

The results presented above can be obtained with a minimum degree of randomization of the initial values in step 3 of Algorithm 1, with different starting values implying different $\bar{\theta}$. This indicates that the detected identification problems are local, consistently with the previous literature. Now we move to the problem of global identification in the AS model.

So far we have considered the case in which no correlation between productivity and government expenditures was allowed, and now we relax this assumption. Let us consider the benchmark parameter vector from Table 1, except that $\rho_{zg} = 0.1$, where a positive value of this coefficient can be thought of as a reduced-form way of capturing the effects of publicly financed research or infrastructure on productivity. When we apply Algorithm 1 to this benchmark parameterization, no observationally equivalent parameter vector can be found, even if none of the deep model parameters are treated as known. While this outcome does not formally prove that the model is globally identified at the considered point, it tells us that allowing for spillovers from public spending to productivity can be an alternative way of ensuring (at least local) identification of the AS model. This result may look counterintuitive as one would expect that increasing the number of unknown parameters should worsen rather than alleviate identification problems. However, it is easy to see that if there are no feedback effects from government to productivity shocks ($\rho_{zg} = 0$), the former do not affect inflation or the interest rate in the AS model. Hence, allowing for nonzero ρ_{zg} breaks the disconnect between moments related to these two observable variables and movements in government spending, which helps identify other parameters.

We next allow for a two-way correlation between productivity and government spending, that is, both ρ_{zg} and ρ_{gz} are no longer restricted to equal zero, and use Algorithm 1 to check global identification of this extension to the original AS model. Table 2 documents the outcomes for two illustrative points θ , differing only in the value of ρ_{gz} . We consider two negative values for this parameter, which can be interpreted as a simple way of capturing countercyclical fiscal policy. The remaining elements of θ are set

TABLE 2. Global identification failure in AS model with correlated government spending and productivity.

Observationally equivalent pairs	Set 1		Set 2	
	θ	$\bar{\theta}$	θ	$\bar{\theta}$
τ	2	2	2	2
β	0.9975	0.8492	0.9975	0.9372
κ	0.33	0.3817	0.33	0.3510
ψ_1	1.5	1.4338	1.5	1.4756
ψ_2	0.125	0.1698	0.125	0.1415
ρ_z	0.9	0.9047	0.9	0.9020
ρ_g	0.95	0.9453	0.95	0.9480
ρ_{zg}	0.1	0.1506	0.1	0.1184
ρ_{gz}	-0.08	-0.0517	-0.075	-0.0625
ρ_m	0.75	0.75	0.75	0.75
σ_z	0.003	0.0032	0.003	0.0031
σ_g	0.006	0.0043	0.006	0.0052
σ_m	0.002	0.002	0.002	0.002

to their baseline values from Table 1, except that, as before, $\rho_{zg} = 0.1$. For each of these two benchmark values of θ , Algorithm 1 finds one observationally equivalent $\bar{\theta}$ from the determinacy region.

Several observations are in order. First, the identification failure concerns a large subset of parameters: only τ , ρ_m , and σ_m are globally identified. Relative to the benchmark θ 's, their observationally equivalent twins $\bar{\theta}$'s are characterized by larger spillovers from government spending to productivity and smaller (in the absolute value) spillovers in the opposite direction, a steeper Phillips curve, and bigger relative weight of output in the monetary policy rule. Second, the distance between θ and its observationally equivalent $\bar{\theta}$ strongly depends on the value of ρ_{gz} . In particular, for the benchmark values of other elements in θ considered in this example, θ becomes identical to $\bar{\theta}$ (i.e., the model is globally identified) when ρ_{gz} is around -0.0713 . Third, while the observationally equivalent model parameterizations are indistinguishable on the basis of the probability distribution, they may have distinct economic implications. For instance, θ and $\bar{\theta}$ imply different responses to government spending and productivity shocks, and also give different answers about the importance of these shocks for fluctuations in the observable variables, even though in our particular example these differences are small. Perhaps more importantly, since the values of the discount factors β and $\bar{\beta}$ in the observationally equivalent pairs θ and $\bar{\theta}$ may be very far from each other, it may matter which of them is chosen to evaluate welfare implications of alternative economic policies, hence affecting the model-based policy recommendations.

8. CONCLUDING REMARKS

In this paper, we have developed a framework for analyzing global identification of DSGE models. It is derived by combining the results on the observational equivalence

of the ABCD-representations established earlier in the literature by Komunjer and Ng (2011) with inherent constraints imposed on these representations by the original model structure. We used this framework to design an algorithm that efficiently searches for observationally equivalent deep parameter sets. The main appeal of the algorithm is that the DSGE model needs to be solved only for the parameter value for which the global identification is checked, but not for all candidate alternative parameterizations. This makes our procedure fast to implement and probably also numerically more accurate. We also demonstrated how the algorithm can be used to detect global identification failure in a variant of the model considered by An and Schorfheide (2007) a small-scale DSGE model widely used in the literature on identification.

Our framework can be easily applied to more sophisticated models. We did (though do not report) it for the medium-sized small open economy model described in Justiniano and Preston (2010). The reason why this setup might be of interest is related to some identification problems encountered by these authors while they estimated the model. However, applying our algorithm to this example did not result in finding parameter values that would be observationally equivalent to the baseline parameterization used by Justiniano and Preston (2010). This is consistent with Iskrev (2010) or Komunjer and Ng (2011), who show that identification problems in richer models, like the canonical Smets and Wouters (2007) setup, are less widespread than one might expect given the evidence for small models. It follows that the identification problems plaguing estimation of medium-sized DSGE models might be more of a weak rather than strong nature, that is, they might result from low curvature of the likelihood function for a typical set of macroeconomic time series rather than from intrinsic features of the model structure.

REFERENCES

- Adjemian, S., H. Bastani, F. Karame, M. Juillard, J. Maih, F. Mihoubi, G. Perendia, J. Pfeifer, M. Ratto, and S. Villemot (2011), “Dynare: Reference Manual Version 4.” Dynare Working Papers, CEPREMAP. [1245]
- An, S. and F. Schorfheide (2007), “Bayesian analysis of DSGE models.” *Econometric Reviews*, 26 (2–4), 113–172. [1245, 1254, 1257, 1261]
- Anderson, G. (2008), “Solving linear rational expectations models: A horse race.” *Computational Economics*, 31 (2), 95–113. [1246]
- Anderson, G. and G. Moore (1985), “A linear algebraic procedure for solving linear perfect foresight models.” *Economics Letters*, 17 (3), 247–252. [1246]
- Beyer, A. and R. E. A. Farmer (2007), “On the indeterminacy of determinacy and indeterminacy, comments on “Testing for indeterminacy”.” *American Economic Review*, 97 (1), 524–529. [1243]
- Blanchard, O. J. and C. M. Kahn (1980), “The solution of linear difference models under rational expectations.” *Econometrica*, 48 (5), 1305–1311. [1246]
- Canova, F., F. Ferroni, and C. Matthes (2014), “Choosing the variables to estimate singular DSGE models.” *Journal of Applied Econometrics*, 29 (7), 1099–1117. [1245]

- Canova, F. and L. Sala (2009), “Back to square one: Identification issues in DSGE models.” *Journal of Monetary Economics*, 56 (4), 431–449. [1243]
- Clarida, R., M. Gertler, and J. Galí (1999), “The science of monetary policy: A New Keynesian perspective.” *Journal of Economic Literature*, 37 (4), 1661–1707. [1254]
- Cochrane, J. H. (2011), “Determinacy and identification with Taylor rules.” *Journal of Political Economy*, 119 (3), 565–615. [1243]
- Del Negro, M. and F. Schorfheide (2013), “DSGE model-based forecasting.” In *Handbook of Economic Forecasting*, Vol. 2 (G. Elliott and A. Timmermann, eds.), 57–110, Elsevier. [1243]
- Fernández-Villaverde, J., J. F. Rubio-Ramírez, and M. S. Santos (2006), “Convergence properties of the likelihood of computed dynamic models.” *Econometrica*, 74 (1), 93–119. [1245]
- Fernández-Villaverde, J., J. F. Rubio-Ramírez, T. J. Sargent, and M. W. Watson (2007), “ABCs (and Ds) of understanding VARs.” *American Economic Review*, 97 (3), 1021–1026. [1246]
- Fukac, M., D. F. Waggoner, and T. Zha (2007), “Local and Global Identification of DSGE Models: A Simultaneous-Equation Approach.” Working Paper, Federal Reserve Bank of Atlanta. [1244]
- Herbst, E. P. and F. Schorfheide (2016), *Bayesian Estimation of DSGE Models*. Princeton University Press. [1257]
- Iskrev, N. (2010), “Local identification in DSGE models.” *Journal of Monetary Economics*, 57 (2), 189–202. [1244, 1245, 1261]
- Justiniano, A. and B. Preston (2010), “Can structural small open-economy models account for the influence of foreign disturbances?” *Journal of International Economics*, 81 (1), 61–74. [1261]
- Kim, J. (2003), “Functional equivalence between intertemporal and multisectoral investment adjustment costs.” *Journal of Economic Dynamics and Control*, 27 (4), 533–549. [1243]
- King, R. G. and M. W. Watson (1998), “The solution of singular linear difference systems under rational expectations.” *International Economic Review*, 39 (4), 1015–1026. [1245]
- King, R. G. and M. W. Watson (2002), “System reduction and solution algorithms for singular linear difference systems under rational expectations.” *Computational Economics*, 20 (1–2), 57–86. [1246]
- Klein, P. (2000), “Using the generalized Schur form to solve a multivariate linear rational expectations model.” *Journal of Economic Dynamics and Control*, 24 (10), 1405–1423. [1245, 1246]
- Kocięcki, A. and M. Kolasa (2018), “Supplement to ‘Global identification of linearized DSGE models.’” *Quantitative Economics Supplemental Material*, 86, <https://doi.org/10.3982/QE530>. [1254, 1255]

Komunjer, I. and S. Ng (2011), “Dynamic identification of dynamic stochastic general equilibrium models.” *Econometrica*, 79 (6), 1995–2032. [1244, 1245, 1247, 1248, 1257, 1258, 1259, 1261]

Lubik, T. A. and F. Schorfheide (2003), “Computing sunspot equilibria in linear rational expectations models.” *Journal of Economic Dynamics and Control*, 28 (2), 273–285. [1246]

Morris, S. D. (2014), “Maximum Likelihood Estimation of Globally Unidentified Dynamic Stochastic General Equilibrium Models Using a-priori Information.” Working Paper, Bowdoin College. [1244]

Qu, Z. and D. Tkachenko (2012), “Identification and frequency domain quasi-maximum likelihood estimation of linearized dynamic stochastic general equilibrium models.” *Quantitative Economics*, 3 (1), 95–132. [1244, 1245, 1258]

Qu, Z. and D. Tkachenko (2017), “Global identification in DSGE models allowing for indeterminacy.” *Review of Economic Studies*, 84, 1306–1345. [1244, 1245, 1248, 1250]

Rothenberg, T. J. (1971), “Identification in parametric models.” *Econometrica*, 39 (3), 577–591. [1248]

Schorfheide, F. (2011), “Estimation and Evaluation of DSGE Models: Progress and Challenges.” NBER Working Papers 16781, National Bureau of Economic Research, Inc. [1254, 1255]

Sims, C. A. (2002), “Solving linear rational expectations models.” *Computational Economics*, 20 (1–2), 1–20. [1246]

Smets, F. and R. Wouters (2007), “Shocks and frictions in US business cycles: a Bayesian DSGE approach.” *American Economic Review*, 97 (3), 586–606. [1246, 1254, 1261]

Uhlig, H. (1999), “A toolkit for analyzing nonlinear dynamic stochastic models easily.” In *Computational Methods for the Study of Dynamic Economies* (R. Marimon and A. Scott, eds.), 30–61, Oxford University Press. [1246, 1249]

Co-editor Frank Schorfheide handled this manuscript.

Manuscript received 14 January, 2015; final version accepted 16 February, 2018; available online 22 February, 2018.