

Dynamic Identification of DSGE Models

Ivana Komunjer* Serena Ng[†]

September 2009

PRELIMINARY: COMMENTS WELCOME

Abstract

This paper studies identification of the parameters of a DSGE model using the entire autocovariance structure of the data. Classical results for identification of structural parameters do not apply because some reduced form parameters in DSGE models cannot be identified. We use the parameter restrictions that two models with identical spectrum must satisfy to obtain the rank and order conditions for identification. Both conditions depend on the parametrization of the model alone, and should be checked before any observations are considered. The results are established in a general set up that allows fewer shocks than endogenous variables. Three examples are considered to illustrate the results.

*University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093. Email: komunjer@ucsd.edu

[†]Columbia University, 420 W. 118 St. MC 3308, New York, NY 10027. Email: serena.ng@columbia.edu

This paper was presented at the 2009 NBER summer institute, as well as the econometrics seminar at UC Davis. We thank Roger Farmer, Frank Schorfheide, Giuseppe Ragusa, and Leon Wegge for helpful comments. Part of this work was done while the first author was visiting Columbia University whose warm hospitality is gratefully acknowledged. The second author acknowledges financial support from the National Science Foundation (SES-0549978).

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models have now reached the level of sophistication to permit analysis of important policy and theoretical macroeconomic issues. Whereas the parameters in these models used to be calibrated, numerical advances in the last two decades have made it possible to estimate models with as many as a hundred parameters. Researchers are, however, aware that not all *deep* or *structural* parameters of DSGE models can be consistently estimated due to identification problems. A DSGE model is identifiable whenever changing the values of the model parameters induce a variation in the reduced form parameter that is not observationally equivalent to its original value. In spite of the recognition of this problem, a procedure has yet to be developed that tells us in a systematic manner how many parameters are identifiable, and if so which ones. This is not a trivial problem because unlike in the classical setup, the *reduced form* parameters of DSGE models are generally not identifiable. The contribution of this paper is to propose rank and order conditions for identification of parameters of interest. These can either be the deep parameters of the optimizing model, or the structural parameters of the log-linearized model. Hereafter, we will let θ denote the parameter vector of interest. Given that these models are dynamic, identification is obtained from the autocovariance structure (spectral density) of the observables. The results are new and not yet seen in the literature.

Since the solutions to log-linearized DSGE models are merely systems of linear simultaneous equations with VARMA representations, it is useful to explain why new identification results are needed. First and foremost, economic theory often considers fewer shocks than observed endogenous variables, a condition known as ‘stochastic singularity’. This singular structure can arise because of adding up constraints as in demand systems, or common shocks as in factor models, of which DSGE is an example. As a consequence, VARMA representations of DSGE models involve matrices that are generally not square. This violates the usual invertibility assumption used in the classical work by Hannan (1971), Zellner and Palm (1974), Hatanaka (1975), and Wallis (1977), for example. Even in models that are nonsingular, identification of VARMA models requires the so-called left-co prime condition to rule out common factors in the autoregressive and moving average polynomials. The identified model is the canonical form obtained from solving the Kronecker indices of the system.¹ These methods are cumbersome and are rarely used in economic analysis. Furthermore, we are not interested in identifying the VARMA parameters per se, but the ones that determine the VARMA parameters. More importantly, canonical forms are a useful (and perhaps the only) starting point if we have no other information about the model. As pointed out in Glover and Willems (1974) in control analysis, when theory imposes a lot of restrictions on a linear dynamical

¹These methods are discussed in Solo (1986), Reinsel (2003), and Lutkepohl (2005).

system, we can identify the parameters of interest without appealing to the canonical form. In particular, in any DSGE model we typically know the number of endogenous state variables and the number of exogenous shocks. Thus, unlike a typical VARMA model, we know the order of the model (also known as the Mcmillian degree) which is useful for identification analysis.

Classical work on the identification of linear simultaneous equations pioneered by Koopmans (1950) exploits the structure of the model, but several features of DSGE models make those results inapplicable. First, DSGE models are dynamic. Whereas in a static case, the parameters of two observationally equivalent structures must be related by a constant matrix, this definition of equivalence is inadequate because two models with different dynamic structures can have identical impulse response functions. Therefore, exclusion restrictions that permit identification of static models may not ensure identification of dynamic models. This problem motivated Rubio-Ramírez, Waggoner, and Zha (2007) to develop results for identification of linear structural VARs. Second, in the traditional work on identification, latent shocks are assumed to be independent and identically distributed (iid) across time. The iid property in particular implies that the distribution of the observables remains constant across time. Identification is then discussed under the assumption that an infinite time series of the data is available, which allows the econometrician to eventually observe the distribution of the DSGE model variables. The iid assumption on the shocks is overly restrictive for time series data. It is common to use the weaker white noise assumption so that identification can be based solely on the autocovariances of the observables and not their entire distribution, which may well vary across time. Finally, DSGE models have no exogenous variables other than the shocks, unlike in Deistler (1976). Latent shocks cannot be used for identification.

There is another important feature of DSGE models that invalidates the classical rank conditions summarized by Fisher (1966) and Rothenberg (1971). In many economic models covered by the classical analysis, the reduced form parameter of the model is known to be identifiable. In such cases, the parameters of the model θ are (locally) identifiable from the distribution of the observables if and only if the mapping from θ to the reduced form parameters is (locally) one to one. The identification problem then becomes simply a question of uniqueness of solutions to systems of equations. This leads to the well known rank conditions for (local) identification (see, e.g., Theorem 6 in Rothenberg, 1971). What makes Rothenberg's (1971) analysis inapplicable here is that the reduced form parameters are generally *not identifiable* in DSGE models as we will see.

For all the reasons above, a new identification theory needs to be developed that allows for stochastically singular models and the possibility that the reduced form parameters are not identifiable. To circumvent the problem of stochastic singularity, it is not uncommon to complete the system by specifying enough shocks so as to match the number of endogenous variables. However,

as argued in Chari, Kehoe, and McGrattan (2009), many shocks in moderate to large new Keynesian models are not structural. We therefore consider a general setup that does not restrict the dimension of the shocks.

Our identification results have implications for frequentist estimation as non-identifiable parameters are not consistently estimable; the results are also useful in a Bayesian context. As Poirier (1998) pointed out, Bayesian analysis is not immune to identification problems. It is well known that local non identification leads to pathological behavior of the posteriors when flat priors are used. This has serious consequences for the posterior computation in practice. Indeed, the local non identification problems lead to reducible Markov chains in which the region of locally unidentified parameter values becomes an absorbing state. This violates the convergence conditions typically imposed for Gibbs samplers. A solution to the problem is to use informative priors (see, e.g., Chao and Phillips, 1998; Kleibergen and van Dijk, 1998). Knowing which parameters are locally unidentifiable would help researchers pinpoint where the specification of such priors is the most important.

Parameters that are not identifiable cannot be consistently estimated. In spite of the importance of this problem, the literature on identification of DSGE models is small. Beyer and Farmer (2007) shows that a determinate solution can be observationally equivalent to one that is not. While indeterminate solutions cannot be ruled out, most solutions are determinate, and identification of determinate solutions is not even completely understood. Canova and Sala (2009) suggests to examine the properties of the impulse responses evaluated at different parameter values, while Consolo, Favero, and Paccagnini (2009) compares the properties of the VAR implied by the DSGE model with those of a factor augmented VAR, also implied by the DSGE model. Both approaches, while useful, have not been shown to be necessary or sufficient for identification. Iskrev (2009) suggests to use all the autocovariances for identification, but does not exploit the parameter restrictions that equivalent systems must satisfy. Moreover, Iskrev's (2009) analysis assumes that the number of shocks equals the number of variables in the system, and that the reduced form parameters of the DSGE model are identifiable from the autocovariances. Our point of departure is that this is not true in general.

The plan of the paper is as follows. In Section 2 we present our setup, state the model assumptions, and study the properties of the observables. In Section 3, we derive operational restrictions that two observationally equivalent reduced form parameters must satisfy. We also make precise the sense in which some components of the DSGE model reduced form parameter are not identifiable in general. Section 4 presents our rank and order conditions for identification. The key feature of the order condition is that it only depends on the dimension of the variables in the DGSE model;

hence it is easy to compute. In Section 5 we then use three examples to illustrate our results. Finally, Section 6 discusses the implications and extensions of this work, and concludes the paper.

2 Setup

In this section we introduce a generic DSGE model and discuss its properties. In particular, we focus on the features of the model that make the classical identification results inapplicable. A careful analysis of the observables at the end of the section prepares the ground for the new identification results that follow.

2.1 Model

Consider a generic nonlinear discrete time DSGE model in which the parameter of interest θ belongs to a parameter set $\Theta \subseteq \mathbb{R}^{n_\theta}$.² Assume that a solution to the log-linearized version of the model exists and is unique see, e.g. Sims (2002). Equivalent representations of the solution can be obtained via algorithms of Anderson and Moore (1985), Uhlig (1999), Sims (2002), Klein (2000), and King and Watson (2002), for example. In this paper, we focus on the following representation. Let K_t be a vector of observed endogenous (state) variables whose values are known at time t , W_t be a vector of observed endogenous (jump) variables, Z_t be a vector of latent exogenous variables, and ϵ_t be a vector of unobserved shocks that satisfy $E_t \epsilon_{t+1} = 0$. These vectors (as well as their leads and lags) are of dimensions n_K , n_W , n_Z and n_Z , respectively. The variables evolve according to the recursive equilibrium law of motion given by:

$$\begin{aligned} K_{t+1} &= P(\theta)K_t + Q(\theta)Z_t \\ W_t &= R(\theta)K_t + S(\theta)Z_t \\ Z_{t+1} &= \Psi(\theta)Z_t + \epsilon_{t+1}. \end{aligned} \tag{1}$$

The matrices $P(\theta)$, $Q(\theta)$, $R(\theta)$, $S(\theta)$ and $\Psi(\theta)$ are of dimensions $(n_K \times n_K)$, $(n_K \times n_Z)$, $(n_W \times n_K)$, $(n_W \times n_Z)$, and $(n_Z \times n_Z)$, respectively. We will refer to the equations in (1) as the (reduced form) solution equations.³ The solution (1) consists of two sets of equations: first are the equations for the endogenous state and jump variables $Y_t \equiv (K_t', W_t')$. The second set of equations specifies the law of motion of the exogenous shock process.

As written, the equations in (1) do not include identities even though variables defined by identities are often of interest. Consider a new set of endogenous jump variables \bar{W}_t defined

²By assumption, parameters known to be only identifiable from the steady state are excluded from θ .

³For details on how (1) can be obtained using alternative solution methods, see Uhlig (1999), for example.

through \bar{n}_W identities:

$$\bar{W}_t = \begin{pmatrix} \bar{A}(\theta) & \bar{B}(\theta) \end{pmatrix} \begin{pmatrix} K_{t+1} \\ W_t \end{pmatrix}.$$

Then letting $\widetilde{W}_t \equiv (W_t', \bar{W}_t')'$, the new solution equations can be written as:

$$\begin{aligned} K_{t+1} &= P(\theta)K_t + Q(\theta)Z_t \\ \widetilde{W}_t &= \widetilde{R}(\theta)K_t + \widetilde{S}(\theta)Z_t \end{aligned}$$

with a new $\tilde{n}_W \times n_K$ matrix \widetilde{R} and a new $\tilde{n}_W \times n_Z$ matrix \widetilde{S} ($\tilde{n}_W \equiv n_W + \bar{n}_W$) defined as:

$$\widetilde{R}(\theta) \equiv \begin{pmatrix} R(\theta) \\ \bar{A}(\theta)P(\theta) + \bar{B}(\theta)R(\theta) \end{pmatrix} \quad \text{and} \quad \widetilde{S}(\theta) \equiv \begin{pmatrix} S(\theta) \\ \bar{A}(\theta)Q(\theta) + \bar{B}(\theta)S(\theta) \end{pmatrix}.$$

The above system of solution equations is thus quite general.

We now state the main model assumptions:

Assumption 1 For every $\theta \in \Theta$ and every $(t, s) \geq 1$, $E(\epsilon_t) = 0$ and $E(\epsilon_t \epsilon_s') = \delta_{t-s} \Sigma(\theta)$ with $\Sigma(\theta)$ positive definite.

Assumption 2 For every $\theta \in \Theta$ and any $z \in \mathbb{C}$, $\det(\text{Id} - \Psi(\theta)z) \det(\text{Id} - P(\theta)z) = 0$ implies $|z| > 1$. Moreover, for every $(\theta, \tilde{\theta}) \in \Theta^2$, $P(\theta)$ and $\Psi(\tilde{\theta})$ have no eigenvalues in common.

Assumption 3 $n_K \leq n_Z \leq n_K + n_W$.

Assumption 4 For every $\theta \in \Theta$, we have: (i) $Q(\theta)$ full rank; (ii) $(Q(\theta)', S(\theta)')$ full rank.

Assumption 5 Let $\Lambda(\theta) \equiv (P(\theta), Q(\theta), R(\theta), S(\theta), \Psi(\theta), \Sigma(\theta))$ denote the reduced form parameter. Then, the mapping $\Lambda : \Theta \rightarrow \mathbb{R}^{n_\Lambda}$ with Θ an open subset of \mathbb{R}^{n_θ} and $n_\Lambda \equiv (n_K + n_W)(n_Z + n_K) + 2n_Z^2$ which to each $\theta \in \Theta$ assigns $\Lambda(\theta)$ is continuously differentiable on Θ .

Assumption 1 states that the shocks $\{\epsilon_t\}$ follow a vector generalization of white noise, which we denote by $\{\epsilon_t\} \sim WN(0, \Sigma(\theta))$. This property of the errors is weaker than that of being iid. DSGE models with rational expectations models assume $\{\epsilon_t\}$ to be a martingale difference sequence, $E_t \epsilon_{t+1} = 0$, whose mean and covariance across time properties are the same as those of a white noise process. The key additional feature of Assumption 1 is to require the contemporaneous shocks covariance matrix $\Sigma(\theta)$ to be constant. Note that $\Sigma(\theta)$ is not required to be of the form $\sigma^2 \text{Id}$ (with $\sigma > 0$). In other words, we allow different components of the shock vector ϵ_t to be correlated as in Curdia and Reis (2009). Since $\Sigma(\theta)$ is nonsingular, the shocks ϵ_t are required to be linearly independent.

The first requirement of Assumption 2 holds when all the eigenvalues of the matrices $\Psi(\theta)$ and $P(\theta)$ lie inside the unit circle, i.e. when $\Psi(\theta)$ and $P(\theta)$ stable. The second restriction rules out situations when the internal propagating mechanism of one structure (as given by $P(\theta_0)$) coincide with the dynamics of the exogenous process in another structure (as given by $\Psi(\theta_1)$). Assumption 3 assumes at least as many shocks as endogenous state variables, but fewer than the total number of endogenous variables. Most DSGE models have this property.

When $n_K \leq n_Z \leq n_K + n_W$, Assumptions 4(i) and 4(ii) ensure that the $n_K \times n_Z$ matrix $Q(\theta)$ and the $(n_K + n_W) \times n_Z$ matrix $(Q(\theta)', S(\theta)')$ are of rank n_K and n_Z , respectively. When those rank conditions fail, then the dynamics of Z_t are no longer fully transmitted to Y_t ; hence, not all information about the shocks can be recovered from the observables. Note that Assumption 4(ii) allows for the identities to be included in the system (1). If the original system in (1) satisfies Assumptions 3 and 4(ii), then the new system is such that $n_K \leq n_Z \leq n_K + \tilde{n}_W$, and $(Q(\theta)', \tilde{S}(\theta)')$ also has rank n_Z . Indeterminate solutions can be considered by suitably expanding the state vector. However, the rank conditions rule out cases when the solution is unique but an expanded state vector is assumed.

Finally, Assumption 5 puts smoothness requirements on the mapping from the parameter vector of interest θ to system matrices $P(\theta), Q(\theta), R(\theta), S(\theta), \Psi(\theta)$ in (1), and the covariance matrix $\Sigma(\theta)$ in Assumption 1. In what follows, we collect all the components of those matrices in the reduced form parameter

$$\Lambda(\theta) \equiv (P(\theta), Q(\theta), R(\theta), S(\theta), \Psi(\theta), \Sigma(\theta))$$

of dimension $n_\Lambda \equiv (n_K + n_W)(n_Z + n_K) + 2n_Z^2$. Given the reduced form solution (1), the objective of the exercise is to estimate θ (or Λ) and eventually to conduct inference about (functions of) θ from the DSGE model observables. Before that, we need to discuss the properties of $\{Y_t\}$.

2.2 Observables

Under Assumptions 1 and 2 the process for $\{Y_t\}$ whose dimension we denote by $n_Y \equiv n_K + n_W$ has the following VMA(∞) representation:

$$Y_t = \begin{pmatrix} K_t \\ W_t \end{pmatrix} = \sum_{j=0}^{\infty} h(j; \theta) \epsilon_{t-j}. \quad (2)$$

The $(n_K + n_W) \times n_Z$ matrices $h(j; \theta)$ are the Markov parameters of the sequence $\{Y_t\}$, obtained from the transfer function (also called impulse response function):

$$\begin{aligned}
H(z; \theta) &= \begin{pmatrix} H_K(z; \theta) \\ H_W(z; \theta) \end{pmatrix} \equiv \sum_{j=0}^{\infty} h(j; \theta) z^j \\
&= \begin{pmatrix} z[\text{Id} - P(\theta)z]^{-1}Q(\theta)[\text{Id} - \Psi(\theta)z]^{-1} \\ \{R(\theta)z[\text{Id} - P(\theta)z]^{-1}Q(\theta) + S(\theta)\}[\text{Id} - \Psi(\theta)z]^{-1} \end{pmatrix}. \tag{3}
\end{aligned}$$

Under our assumptions, the sequence of endogenous variables $\{Y_t\}$ has two important properties which we now present in the form of two Lemmas.

Lemma 1 (*Covariance Stationarity*) *Let Assumptions 1 and 2 hold. Then for every $\theta \in \Theta$, $\{Y_t\}$ is weakly stationary with $E(Y_t) = 0$ and $E(Y_t Y_s') \equiv \Gamma(s - t; \theta) = \sum_{j=0}^{\infty} h(j; \theta) \Sigma(\theta) h'(j + s - t; \theta)$, for all $(t, s) \geq 1$.*

Lemma 1 is an immediate consequence of (2) and (3). The covariance stationarity result is crucial for subsequent developments as we can exploit the time invariant second moment properties of the endogenous variables $\{Y_t\}$ to identify the parameter of interest θ . Indeed, when the process is weakly stationary, the autocovariances completely summarize the properties of $\{Y_t\}$, and the econometrician eventually observes its entire autocovariance generating function, which for any $z \in \mathbb{C}$ is defined as:

$$\Omega(z; \theta) \equiv \sum_{j=-\infty}^{+\infty} \Gamma(j; \theta) z^{-j}.$$

Note that for every $\theta \in \Theta$, $\Omega(z; \theta)$ is a square $n_Y \times n_Y$ matrix that satisfies $\Omega(z^{-1}; \theta)' = \Omega(z; \theta)$, for every $z \in \mathbb{C}$. The real function $\Omega(\exp(i\omega); \theta)$, defined for any $\omega \in \mathbb{R}$, is called the spectral density (or spectrum) of the observables $\{Y_t\}$. Furthermore,

$$\Omega(\exp(i\omega); \theta) = \Gamma(0; \theta) + 2 \sum_{j=1}^{\infty} \Gamma(j; \theta) \cos(j\omega),$$

which is always positive semi definite. The autocovariances $\Gamma(j; \theta)$ are then obtained from the spectral density $\Omega(\exp(i\omega); \theta)$ via the following inversion formula:

$$\Gamma(j; \theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Omega(\exp(i\omega)) \exp(-i\omega j) d\omega.$$

The autocovariance matrices $\Gamma(j; \theta)$ satisfy $\Gamma(j; \theta) = \Gamma(-j; \theta)'$. With a slight abuse of terminology, we shall hereafter refer to $\Omega(z; \theta)$ as the spectral density as well.

It is important to emphasize that the white noise Assumption 1 does not prevent the observables to have higher moments that change in time, nor does it prevent the unobservables ϵ_t to be conditionally heteroskedastic. Imposing further restrictions on ϵ such as being homoskedastic or iid would imply further properties on the observables that could be used for identification.

From the VMA(∞) representation in (2) it follows that the process $\{Y_t\}$ is causal with $\sum_{j=0}^{\infty} |h(j; \theta)| < \infty$. Under our assumptions, $\{Y_t\}$ is also invertible meaning that $\{\epsilon_t\}$ is spanned by the past and present values of $\{Y_t\}$, or that $\{\epsilon_t\}$ is ‘fundamental’ for $\{Y_t\}$.

Lemma 2 (*Invertibility*) *Let Assumptions 1 through 4 hold. Then for every $\theta \in \Theta$, $\{Y_t\}$ is invertible and we can write $\epsilon_t = \sum_{j=0}^{\infty} g(j; \theta) Y_{t-j}$ for all $t \geq 1$ with $\sum_{j=0}^{\infty} |g(j; \theta)| < \infty$.*

Proof of Lemma 2 is in Appendix A. The invertibility condition is usually stated in terms of the eigenvalues of $H(z; \theta)$, which is sufficient. The necessary condition that the rank of $H(z; \theta)$ is n_Z is, however, enough for us to obtain a VAR(∞) representation of the model (see, e.g., Hannan and Diestler, 1988). Economic theory usually cannot rule out non-fundamentalness, a situation that arises when agents have more information than the econometrician. Giannone and Reichlin (2006) argue that superior information is much less likely to happen when there are more variables than shocks in the system. Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007) suggest a way to check fundamentalness.

We need invertibility of $\{Y_t\}$ to ensure that the spectral density matrix $\Omega(z; \theta)$ has constant rank n_Z almost everywhere in \mathbb{C} . This *minimum phase* (or ‘miniphase’) property will allow us to go from the spectral density matrix to the transfer function that generated it.⁴ Under regularity conditions (see, e.g., Deistler, 2006):

$$\Omega(z; \theta) = H(z; \theta) \Sigma(\theta) H(z^{-1}; \theta)' \quad (4)$$

with the transfer function for $H(z; \theta)$ as defined in (3).

3 Identifiability and Observational Equivalence

In this section, we ask the following question: when do two distinct values of the reduced form parameter of a DSGE model yield the same spectrum? We start our analysis with a definition of observational equivalence that applies to dynamic models. We then derive necessary and sufficient conditions under which two sets of reduced form parameters $\Lambda(\theta_0)$ and $\Lambda(\theta_1)$ of the DSGE model (1) are observationally equivalent.

3.1 Definitions

In static analysis, observational equivalence is typically defined with reference to the distribution (or density) of the observed variables Y_t . In a dynamic setting, this distribution may well vary across

⁴A linear, time invariant system is said to have minimum phase if the system and its inverse are invertible and stable. Under Assumptions 1 through 4, the reduced form solution of our DSGE model has this property.

time. However, in models that are linear, the process $\{Y_t\}$ is known to be covariance stationary. In such cases, observational equivalence is defined with reference to the entire autocovariance structure of the observables. In other words, we need to work with spectral densities of the observables.

It is clear from (3) and (4) that the spectral density $\Omega(z; \theta)$ depends on the deep parameter θ only through the reduced form parameter $\Lambda(\theta)$ and we have $\Omega(z; \theta) = \Omega(z; \Lambda(\theta))$. Each value θ_0 of θ implies the value $\Lambda_0 \equiv \Lambda(\theta_0)$ which completely characterizes the spectrum of $\{Y_t\}$. Now the determination of $\Omega(z; \theta)$ from $\Lambda(\theta)$ is straightforward once a value of θ is given. The converse problem, which is that of identification, is to start with a prescribed spectral density $\Omega(z; \theta)$ and ask which sextuple $\Lambda(\theta)$ is such that (3) and (4) hold.

The question of identification from the spectral density can be stated as follows: having observed the spectral density function $\Omega(z; \theta_0)$, under what conditions is it possible to uncover the value θ_0 that generated it? It is possible that two different values θ_0 and θ_1 of the parameter θ in (1) lead to the reduced form parameters $\Lambda(\theta_0)$ and $\Lambda(\theta_1)$ with the same spectral density. We say that $\Lambda(\theta_0)$ and $\Lambda(\theta_1)$ are observationally equivalent if $\Omega(z; \theta_1) = \Omega(z; \theta_0)$ for all $z \in \mathbb{C}$, or more formally, if for every $z \in \mathbb{C}$:

$$H(z; \theta_0)\Sigma(\theta_0)H(z^{-1}; \theta_0)' = H(z; \theta_1)\Sigma(\theta_1)H(z^{-1}; \theta_1)'. \quad (5)$$

The problem of identifying the DSGE model parameter θ from the spectral density of the observables can be defined as follows.

Definition 1 *The DSGE model (1) is said to be locally identifiable from the spectral density of $\{Y_t\}$ at a point $\theta_0 \in \Theta$ if there exists an open neighborhood of θ_0 such that for every θ_1 in this neighborhood $\Lambda(\theta_0)$ and $\Lambda(\theta_1)$ are observationally equivalent only if $\theta_1 = \theta_0$.*

Put in words, in a neighborhood of θ_0 there are no two distinct reduced form parameters that generated the same spectral density of the observables. As already pointed out, the crucial assumption of Fisher (1966) and Rothenberg (1971)—that $\Lambda(\theta)$ itself is (globally or locally) identifiable—does not hold in DSGE models. To see this, let U be a full rank $n_Z \times n_Z$ matrix, and consider the system:

$$\begin{aligned} K_{t+1} &= P(\theta)K_t + \tilde{Q}(\theta)\tilde{Z}_t \\ W_t &= R(\theta)K_t + \tilde{S}(\theta)\tilde{Z}_t \end{aligned} \quad (6)$$

with $\tilde{Z}_t \equiv U^{-1}Z_t$, and $\tilde{Q}(\theta) \equiv Q(\theta)U$, $\tilde{S}(\theta) = S(\theta)U$, $\tilde{\Psi}(\theta) = U^{-1}\Psi(\theta)U$, $\tilde{\Sigma}(\theta) = U^{-1}\Sigma(\theta)U^{-1}$. We then have that $\tilde{\Lambda}(\theta) = (\tilde{P}(\theta), \tilde{Q}(\theta), \tilde{R}(\theta), \tilde{S}(\theta), \tilde{\Psi}(\theta), \tilde{\Sigma}(\theta))$ is observationally equivalent to $\Lambda(\theta)$, since $\Omega(z; \Lambda(\theta)) = \Omega(z; \tilde{\Lambda}(\theta))$ for every $z \in \mathbb{C}$. Yet, $\Lambda(\theta)$ is different from $\tilde{\Lambda}(\theta)$. The reduced form parameter $\Lambda(\theta)$ is thus not identifiable from the spectral density of the observables. This means

that Rothenberg's (1971) rank condition alone will in general not be sufficient to guarantee the identifiability of θ .

It is known in work by Glover (1973) and Glover and Willems (1974) that how to identify the parameter θ in (1) ultimately depends on whether or not the inputs ϵ_t are observed. If they are observed, then the parameters of the model can be identified from the transfer function in (3) if the mapping from the parameter space to the Markov parameters is locally one-to-one. When the input noises are unobserved, then in general, identification can only be based on the autocovariance generating function since the properties of stationary processes are completely summarized by their second moments.⁵

To derive conditions for identification of θ from the spectrum, we proceed along the lines of work by Glover and Willems (1974) in control theory, but take into account that our state vector is partially observed. First, we look for all reduced form parameter values $\Lambda(\theta_1)$ that are observationally equivalent to $\Lambda(\theta_0)$. For instance, if it is possible to find a value θ_1 such that the transformation in (6) can be written as $\tilde{\Lambda}(\theta_0) = \Lambda(\theta_1)$, then $\Lambda(\theta_0)$ and $\Lambda(\theta_1)$ are observationally equivalent. The transformation in (6) is one example from which observationally equivalent parameters can be obtained; however, we need to find *all* such transformations. Second, once the set of all reduced form parameters $\Lambda(\theta_1)$ that are observationally equivalent to $\Lambda(\theta_0)$ is obtained, we impose conditions that ensure that in this set θ_1 is always equal to θ_0 .

Now from (5), many observationally equivalent transformations $\Lambda(\theta_1)$ with $\Omega(z; \theta_1) = \Omega(z; \theta_0)$ can arise because: (i) each $H(z; \theta)$ can potentially be obtained from a multitude of quintuples $(P(\theta), Q(\theta), R(\theta), S(\theta), \Psi(\theta))$ in (3), and (ii) there can be transfer functions $H(z; \theta)$ and shock covariance matrices $\Sigma(\theta)$ which jointly satisfy (4). In the next two subsections, we derive the necessary and sufficient conditions for both sources of observational equivalence.

3.2 Equivalence of Transfer Functions

Assume for the time being that ϵ_t in (1) is observable. Then, from (2), the econometrician can observe the entire transfer function $H(z; \theta)$. The question which we now ask is as follows: given $H(z; \theta)$ is it possible to uniquely recover the quintuple $(P(\theta), Q(\theta), R(\theta), S(\theta), \Psi(\theta))$ that generated it?

It is clear from (3) that to any value θ_0 of θ we can associate a unique transfer function $H(z; \theta_0)$. In addition, to each θ_0 corresponds a unique quintuple $(P(\theta_0), Q(\theta_0), R(\theta_0), S(\theta_0), \Psi(\theta_0))$ which we call a *realization* of the transfer function $H(z; \theta_0)$. We now characterize all realizations $(P(\theta_1), Q(\theta_1), R(\theta_1), S(\theta_1), \Psi(\theta_1))$ that are equivalent to $(P(\theta_0), Q(\theta_0), R(\theta_0), S(\theta_0), \Psi(\theta_0))$ in

⁵There are, however, exceptions when identification can be achieved through the Markov parameters even if the input is not observed. See Grewal and Glover (1976) and Glover and Willems (1974).

a sense that they yield the same transfer function $H(z; \theta_1) = H(z; \theta_0)$ everywhere in \mathbb{C} .

Lemma 3 (*Similarity Transform*) *Let Assumptions 1 through 4 hold. Consider the realizations $(P(\theta_0), Q(\theta_0), R(\theta_0), S(\theta_0), \Psi(\theta_0))$ and $(P(\theta_1), Q(\theta_1), R(\theta_1), S(\theta_1), \Psi(\theta_1))$ of the transfer functions $H(z; \theta_0)$ and $H(z; \theta_1)$, respectively, with $(\theta_0, \theta_1) \in \Theta^2$. Then the two realizations are equivalent, i.e. $H(z; \theta_0) = H(z; \theta_1)$ for every $z \in \mathbb{C}$, if and only if $P(\theta_1) = P(\theta_0)$, $R(\theta_1) = R(\theta_0)$, and there exists a full rank $n_Z \times n_Z$ matrix T such that:*

$$Q(\theta_1) = Q(\theta_0)T^{-1}, \quad S(\theta_1) = S(\theta_0)T^{-1} \quad \text{and} \quad \Psi(\theta_1) = T\Psi(\theta_0)T^{-1}.$$

Proof of Lemma 3 is in Appendix B. The proof goes roughly is as follows. We consider the state space representation of our model and show that it is minimal, so that the dimension of the latent state vector is no larger than any other system with the same output spectral density. To show minimality, we need to establish the system is controllable and observable.⁶ Now a well known result in control theory⁷ is that all equivalent minimal realizations must be linearly related through a similarity transformation, which is simply a linear change in the coordinates of the latent state variables. Our state variables are partially observed, but the similarity transforms can still be applied to the system. Solving the similarity transform equations taking into account that our state vector is partially observed, and expressing them in terms of the system matrices $P(\theta), Q(\theta), R(\theta), S(\theta)$ and $\Psi(\theta)$ gives the stated results.

Lemma 3 says the following: assume that it is possible for the econometrician to observe the transfer function $H(z; \theta)$. Then from the transfer function, the econometrician can uniquely recover the system matrices $P(\theta)$ and $R(\theta)$ that generated it. However, the matrix $Q(\theta), S(\theta)$ and $\Psi(\theta)$ can only recovered up to an unknown full rank matrix T . Put differently, Lemma 3 shows that $P(\theta)$ and $R(\theta)$ are exactly identified from the transfer function. However, further restrictions are needed in order to also identify $Q(\theta), S(\theta)$, and $\Psi(\theta)$.

3.3 Equivalence of Spectral Densities

Now the econometrician only observes the autocovariance generating function of the observables $\{Y_t\}$, or equivalently, its spectral density $\Omega(z; \theta)$. Each spectral density is obtained from the transfer function $H(z; \theta)$ and the errors' covariance matrix $\Sigma(\theta)$ according to the equation in (4). The question which we now ask is as follows: given a spectral density $\Omega(z; \theta)$, is it possible to uniquely recover the couple $(H(z; \theta), \Sigma(\theta))$ that generated it.

⁶Controllability means that for any initial state, it is always possible to design an input sequence that puts the system to a desired final state. Observability means that we can always reconstruct the initial state from observing the evolution of the output, provided that we also know the evolution of the input. Formally, these conditions hold when the controllability and observability matrices are full rank.

⁷See, for example, Theorem 8.13 of (Gourieroux and Monfort, 1997).

It is clear that there can be an infinity of transfer functions $H(z; \theta)$ and shock covariance matrices $\Sigma(\theta)$ which jointly satisfy (4). However, invertible systems have the *minimum phase* property, and in minimum phase systems, all observationally equivalent couples $(H(z; \theta_0), \Sigma(\theta_0))$ and $(H(z; \theta_1), \Sigma(\theta_1))$ are related through a particular transformation. A necessary and sufficient condition for two couples $(H(z; \theta_0), \Sigma(\theta_0))$ and $(H(z; \theta_1), \Sigma(\theta_1))$ to be observationally equivalent is as follows:

Lemma 4 (*Minimum Phase*) *Let Assumptions 1 through 4 hold. Consider two couples $(H(z; \theta_0), \Sigma(\theta_0))$ and $(H(z; \theta_1), \Sigma(\theta_1))$ with spectral densities $\Omega(z; \theta_0)$ and $\Omega(z; \theta_1)$, respectively, where $(\theta_0, \theta_1) \in \Theta^2$. Then the two couples are equivalent, i.e. $\Omega(z; \theta_0) = \Omega(z; \theta_1)$ for every $z \in \mathbb{C}$, if and only if there exists a full rank $n_Z \times n_Z$ matrix U such that:*

$$\text{for every } z \in \mathbb{C} \quad H(z; \theta_1) = H(z; \theta_0)U, \quad \text{and} \quad U\Sigma(\theta_1)U' = \Sigma(\theta_0).$$

A proof is given in Appendix B. The result of Lemma 4 is an equivalence statement that contains two implications. The sufficiency part is obvious: indeed, if the two couples $(H(z; \theta_0), \Sigma(\theta_0))$ and $(H(z; \theta_1), \Sigma(\theta_1))$ are related to each other by the above transformation involving a full rank matrix U , then

$$H(z; \theta_0)\Sigma(\theta_0)H(z^{-1}; \theta_0)' = H(z; \theta_0)U\Sigma(\theta_1)U'H(z^{-1}; \theta_0)' = H(z; \theta_1)\Sigma(\theta_1)H(z^{-1}; \theta_1)'$$

so the two spectral density matrices are the same.

The crux of Lemma 4 is the necessity part, which consists in showing that the two spectral densities are the same only if $(H(z; \theta_0), \Sigma(\theta_0))$ and $(H(z; \theta_1), \Sigma(\theta_1))$ are related to each other by the above transformation involving U . The problem of finding the transfer function from a given spectral density is problem known in control theory as ‘spectral factorization’ (see, e.g., Anderson, 1969). Its solution requires that the rank of the spectral density matrix $\Omega(z; \theta)$ of Y_t be constant and known almost everywhere in \mathbb{C} . When $\{Y_t\}$ is causal and invertible, the constant rank property holds provided the shocks ϵ_t have a full rank covariance matrix $\Sigma(\theta)$. In addition, the rank of $\Omega(z; \theta)$ is then shown to be equal to n_Z , the size of the shocks. It is worth pointing out that the result of Lemma 4 does not require the size of the shocks (n_Z) to be the same as that of the observables ($n_K + n_W$). In other words, Lemma 4 applies in systems that are possibly stochastically singular.

A sketch of the proof is as follows. We first show that for any $\theta_0 \in \Theta$, the rank of $H(z; \theta_0)$ is n_Z for all $|z| > 1$. For this, we call $L(\theta)$ the Cholesky triangle of $\Sigma(\theta)$ (i.e. $\Sigma(\theta) = L(\theta)L(\theta)'$) and we consider a $n_Z \times (n_K + n_W)$ matrix $W_d(z) \equiv L(\theta_0)'H(z^{-1}; \theta_0)'$ that is a ‘spectral factor’, meaning that $\Omega(z; \theta_0) = W_d(z^{-1})'W_d(z)$. We next show that $W_d(z)$ has rank n_Z . Results in Anderson (1969) and Newcomb and Anderson (1968) can then be used to show the following: if

$\widetilde{W}_d(z) = L(\theta_1)'H(z^{-1}; \theta_1)'$ is another minimum spectral factor, then we must have $\widetilde{W}_d(z) = VW_d(z)$ where V is an orthogonal $n_Z \times n_Z$ matrix (i.e. $V'V = \text{Id}_{n_Z}$). Letting $U \equiv L(\theta_0)V'L(\theta_1)^{-1}$ then yields the result. We refer to this as a miniphase lemma because the spectral factors of minimum phase systems have a constant rank of n_Z .

Combining Lemma 4 with the similarity results obtained in Lemma 3 leads to our first Proposition.

Proposition 1 *Let Assumptions 1 through 4 hold. Consider two reduced form parameters $\Lambda(\theta_0)$ and $\Lambda(\theta_1)$ with $(\theta_0, \theta_1) \in \Theta^2$. Then $\Lambda(\theta_0)$ and $\Lambda(\theta_1)$ are observationally equivalent if and only if:*

$$P(\theta_1) = P(\theta_0), \quad R(\theta_1) = R(\theta_0)$$

and there exists a full rank $n_Z \times n_Z$ matrix U such that:

$$Q(\theta_1) = Q(\theta_0)U, \quad S(\theta_1) = S(\theta_0)U, \quad \Psi(\theta_1) = U^{-1}\Psi(\theta_0)U, \quad \Sigma(\theta_1) = U^{-1}\Sigma(\theta_0)U^{-1}'.$$

A proof is given in Appendix Appendix B. Proposition 1 establishes that the two transformation matrices T and U defined in Lemma 3 and Lemma 4, respectively, are related to each other through $T = U^{-1}$.

The conditions in Proposition 1 are necessary and sufficient for $\Lambda(\theta_0)$ and $\Lambda(\theta_1)$ to have the same spectrum. Like for the equivalence results derived in Lemma 3 and Lemma 4, the sufficiency part of Proposition 1 is obvious. On the other hand, the necessity part is nontrivial to derive. We now give an interpretation of the necessity statement.

We know from Lemma 4 that $\Lambda(\theta_1)$ and $\Lambda(\theta_0)$ are observationally equivalent if only if the Markov parameters obtained under the two reduced form parameters satisfy:

$$\begin{aligned} Q(\theta_1) &= Q(\theta_0)U & (7) \\ S(\theta_1) &= S(\theta_0)U \\ \sum_{j=0}^k P(\theta_1)^j Q(\theta_1) \Psi(\theta_1)^{k-j} &= \left[\sum_{j=0}^k P(\theta_0)^j Q(\theta_0) \Psi(\theta_0)^{k-j} \right] U \\ S(\theta_1) \Psi(\theta_1)^k + R(\theta_1) \sum_{j=0}^{k-1} P(\theta_1)^j Q(\theta_1) \Psi(\theta_1)^{k-1-j} &= \left[S(\theta_0) \Psi(\theta_0)^k + R(\theta_0) \sum_{j=0}^{k-1} P(\theta_0)^j Q(\theta_0) \Psi(\theta_0)^{k-1-j} \right] U \end{aligned}$$

for every $k \geq 0$, where U is a $n_Z \times n_Z$ full rank matrix.⁸ Proposition 1 provides the conditions under which the above infinite set of equations holds. It provides an operational definition of observational equivalence of two spectral densities in terms of the system matrices $P(\theta), Q(\theta), R(\theta), S(\theta), \Psi(\theta)$

⁸The equations above are obtained by applying Lemma 4 to Markov parameters $h(k; \theta_1)$ and $h(k; \theta_0)$ in (2).

and $\Sigma(\theta)$. Its key implication is that some of the reduced form parameters are not identifiable. If the parameter of interest is the reduced form parameter $\Lambda(\theta)$ itself, then Proposition 1 shows that only the components of $\Lambda(\theta)$ corresponding to system matrices $P(\theta)$ and $R(\theta)$ in (1) are identifiable from the spectrum. These features hold even when there are as many shocks as endogenous variables in the system.

It is important to point out that the restriction to minimal state systems allows us to focus on systems related by similarity transform, while the minimum phase property narrows down equivalent models to those couples $(H(z; \theta), \Sigma(\theta))$ that satisfy Lemma 4. If one (or both) property fails to hold, observational equivalence can only be defined in terms of weaker restrictions on the spectral density (see, e.g. Glover and Willems, 1974).

4 Rank and Order Conditions for Identification

Having pinned down how two reduced form parameters with the same spectrum must be related, we now use those restrictions to derive conditions that are necessary and sufficient for θ to be locally identified. We start by considering the case when identification is possible from the spectrum alone. We then turn to the case when additional a priori restrictions on θ are needed.

4.1 Identifiability from the spectrum

Let $\lambda : \Theta \times \mathbb{R}^{n_Z^2} \rightarrow \mathbb{R}^{n_\Lambda}$ be a mapping which to each deep parameter $\theta \in \Theta$ and $n_Z \times n_Z$ full rank matrix U assigns a transformed reduced parameter vector $\lambda(\theta, U)$ of dimension $n_\Lambda \equiv (n_K + n_W)(n_Z + n_K) + 2n_Z^2$, defined by:⁹

$$\lambda(\theta, U) \equiv \begin{pmatrix} \text{vec}(P(\theta)) \\ \text{vec}(Q(\theta)U) \\ \text{vec}(R(\theta)) \\ \text{vec}(S(\theta)U) \\ \text{vec}(U^{-1}\Psi(\theta)U) \\ \text{vec}(U^{-1}\Sigma(\theta)U^{-1'}) \end{pmatrix} \quad (8)$$

Then θ_0 is identifiable from the spectrum of $\{Y_t\}$ if and only if the system of n_Λ equations given by $\lambda(\theta_0, \text{Id}) = \lambda(\theta, U)$ has a unique solution $\theta = \theta_0$ and $U = \text{Id}$. We have the following result:

Lemma 5 *Let Assumptions 1 through 5 hold. Then the DSGE model (1) is locally identifiable from the spectral density of $\{Y_t\}$ at a point $\theta_0 \in \Theta$ if and only if the mapping λ is locally injective at $\theta = \theta_0$ and $U = \text{Id}$.*

⁹If $A \in \mathbb{R}^{m \times n}$ then $\text{vec}(A)$ is defined to be the nm -vector formed by stacking the columns of A on top of one another, i.e. $\text{vec}(A) \in \mathbb{R}^{nm}$. Note that $\text{vec}(A') = T_{m,n}\text{vec}(A)$ where $T_{m,n}$ is an $nm \times nm$ permutation matrix that satisfies $T_{n,m}T_{m,n} = \text{Id}_{mn}$, $T_{n,m} = T_{m,n}^{-1}$, and $T_{m,n} = T'_{n,m}$. Moreover, if $B \in \mathbb{R}^{p \times q}$ then $A \otimes B$ denotes the Kronecker product (or tensor product) of A and B , i.e. $A \otimes B \in \mathbb{R}^{mp \times nq}$.

A proof of Lemma 5 is in Appendix C. The mapping in 8 defines a system of $n_\Lambda = (n_K + n_W)(n_Z + n_K) + 2n_Z^2$ equations in $(n_\theta + n_Z^2)$ unknowns. We now look for necessary and sufficient conditions under which a (locally) unique solution to this system is $\theta = \theta_0$ and $U = \text{Id}$. It is a well known result (Fisher, 1966) that if the rank of the matrix of partial derivatives of $\lambda(\theta, U)$ with respect to the components of θ and U remains constant in a neighborhood of (θ_0, Id) , then a necessary and sufficient condition for λ to be locally injective is that:

$$\text{rank} \left(\frac{\partial \lambda(\theta_0, \text{Id})}{\partial \theta} \quad \frac{\partial \lambda(\theta_0, \text{Id})}{\partial \text{vec} U} \right) = n_\theta + n_Z^2. \quad (9)$$

See also Glover and Willems (1974) for a similar result. Necessary and sufficient conditions for identifiability of θ from the spectrum are given in the following Proposition:

Proposition 2 *Let Assumptions 1 through 5 hold. If the rank of $\Delta(\theta)$ remains constant in a neighborhood of θ_0 where*

$$\Delta(\theta) \equiv \left(\frac{\partial \lambda(\theta, \text{Id})}{\partial \theta} \quad \frac{\partial \lambda(\theta, \text{Id})}{\partial \text{vec} U} \right) = \begin{pmatrix} \frac{\partial \text{vec} P(\theta)}{\partial \theta} & 0_{n_K^2 \times n_Z^2} \\ \frac{\partial \text{vec} Q(\theta)}{\partial \theta} & \text{Id}_{n_Z} \otimes Q(\theta) \\ \frac{\partial \text{vec} R(\theta)}{\partial \theta} & 0_{n_K n_W \times n_Z^2} \\ \frac{\partial \text{vec} S(\theta)}{\partial \theta} & \text{Id}_{n_Z} \otimes S(\theta) \\ \frac{\partial \text{vec} \Psi(\theta)}{\partial \theta} & \text{Id}_{n_Z} \otimes \Psi(\theta) - \Psi(\theta)' \otimes \text{Id}_{n_Z} \\ \frac{\partial \text{vec} \Sigma(\theta)}{\partial \theta} & -(\text{Id}_{n_Z^2} + T_{n_Z, n_Z})(\Sigma(\theta) \otimes \text{Id}_{n_Z}) \end{pmatrix},$$

then a necessary and sufficient rank condition for the DSGE model (1) to be locally identified from the spectrum of $\{Y_t\}$ at a point θ_0 in Θ is: $\text{rank} \Delta(\theta_0) = n_\theta + n_Z^2$. Moreover, a necessary order condition is:

$$n_\theta \leq (n_K + n_W)(n_Z + n_K) + \frac{n_Z(n_Z + 1)}{2}.$$

The proof is in Appendix C. Proposition 2 contains two results. First is a rank condition which is both necessary and sufficient for the DSGE model (1) to be identified. This condition is entirely new and not yet seen in the literature on identification. It extends the rank conditions of Glover (1973) and Glover and Willems (1974) that are necessary and sufficient for θ to be identifiable from the transfer function. Glover's (1973) and Glover and Willems's (1974) results require the shocks ϵ_t to be observed. When they are unobserved, then our rank condition applies. The second result of Proposition 2 is an order restriction on n_θ without which θ cannot be identified from the spectrum. The intuition behind is simple: the number of linearly independent equations defined by the mapping $\lambda(\theta, U)$ in (8) needs to be at least as large as the number of unknowns $n_\theta + n_Z^2$. Since the last n_Z^2 equations correspond to the matrix $U^{-1}\Sigma(\theta)U^{-1'}$ which is always symmetric, they only contain $\frac{n_Z(n_Z+1)}{2}$ independent components. Hence the order restriction follows.

Whether or not the DSGE model is identifiable from the spectrum, crucially depends on the dynamics of Z_t in (1). This is because in both the rank and order conditions, the dimension n_θ of θ depends on the assumptions placed on the system matrix Ψ and the covariance matrix Σ . In particular, when in the DSGE model (1) the shocks are mutually uncorrelated univariate AR(1) processes, Ψ and Σ are diagonal matrices, which contributes to $2n_Z$ unknowns in n_θ , instead of $n_Z^2 + n_Z(n_Z + 1)/2$ unknowns when the shocks are VAR(1) and those matrices are unrestricted (we recall that Σ is always required to be symmetric). Hence, it may well be the case that the unrestricted DSGE model with VAR(1) shocks fails to be identified, while the opposite holds under the AR(1) specification.

What makes Proposition 2 important is the result that the deep parameter θ can be identified from the spectrum even though the reduced form parameter $\Lambda(\theta)$ is itself not identifiable. An immediate consequence of the rank condition in Proposition 2 is the following simple *necessary* condition for identification:

$$\text{rank} \left(\frac{\partial \text{vec}(\Lambda(\theta_0))}{\partial \theta} \right) = n_\theta. \quad (10)$$

While necessary, the above condition by itself is not sufficient to guarantee identifiability. Not surprisingly, our rank condition in Proposition 2 is stronger than (10). This is because the reduced form parameter $\Lambda(\theta)$ is not identifiable so a stronger requirement is needed to identify θ . Intuitively, the identifiability of the deep parameter obtains whenever the deviations in the reduced form parameter $\Lambda(\theta)$ induced by a change in θ are incompatible with observationally equivalent transformations in Proposition 1.

Canova and Sala (2009) and Iskrev (2007, 2008) suggest to check the rank of $\frac{\partial \text{vec}(\Lambda(\theta_0))}{\partial \theta}$, just as suggested in (10). However, their recommendation comes as a consequence of considering $\frac{\partial \log L}{\partial \theta} = \frac{\partial \log L}{\partial \text{vec} \Lambda} \frac{\partial \text{vec} \Lambda}{\partial \theta}$, and assuming that $\frac{\partial \log L}{\partial \text{vec} \Lambda}$ is full rank. But the reduced form parameter $\Lambda(\theta)$ is in fact not identifiable as Proposition 1 shows, and we make no reference to the likelihood L . The rank condition on $\frac{\partial \text{vec}(\Lambda(\theta_0))}{\partial \theta}$ alone is still necessary for identification; it is however not sufficient. This result has a classical flavor, even though we work with assumptions that would not be valid in a classical setup.

Finally, it is worth emphasizing that the results of Proposition 2 do not involve the spectrum itself. This makes our approach fundamentally different from approaches based on the information matrix (Iskrev, 2007). Whether or not the DSGE model (1) is identifiable depends on the parametrization alone and our rank and order conditions should be checked before any data observations are considered.

4.2 Identifiability Under A Priori Restrictions

When the order condition in Proposition 2 fails, then it is not possible to identify θ_0 from the spectrum alone. We need to impose a set of a priori restrictions on some of the components of θ . Let θ_0 satisfy a set of a priori restrictions:

$$\varphi(\theta_0) = 0 \quad (11)$$

where $\varphi : \Theta \rightarrow \mathbb{R}^{n_\varphi}$ is assumed to be continuous.

The constraints imposed by Proposition 1 are already defined by the mapping λ in (8). We now look for conditions under which the mapping $\chi : \Theta \times \mathbb{R}^{n_Z^2} \rightarrow \mathbb{R}^{n_\varphi + n_\Lambda}$ which to each deep parameter θ and full rank $n_Z \times n_Z$ matrix U assigns $\chi(\theta, U)$ defined by:

$$\chi(\theta, U) \equiv \begin{pmatrix} \varphi(\theta) \\ \lambda(\theta, U) \end{pmatrix}$$

is locally injective at $\theta = \theta_0$ and $U = \text{Id}$. As before, this condition is necessary and sufficient for θ to be locally identifiable from the spectrum and the a priori restrictions (11). If the rank of the matrix of partial derivatives of $\chi(\theta, U)$ remains constant in a neighborhood of (θ_0, Id) , then a necessary and sufficient condition for χ to be locally injective is that:

$$\text{rank} \left(\frac{\partial \chi(\theta_0, \text{Id})}{\partial \theta} \quad \frac{\partial \chi(\theta_0, \text{Id})}{\partial \text{vec} U} \right) = n_\theta + n_Z^2. \quad (12)$$

We then have the following result:

Proposition 3 *Let Assumptions 1 through 5 hold and assume that φ is continuously differentiable on Θ . If the rank of $\Delta_\varphi(\theta)$ remains constant in a neighborhood of θ_0 where*

$$\Delta_\varphi(\theta) \equiv \left(\frac{\partial \chi(\theta, \text{Id})}{\partial \theta} \quad \frac{\partial \chi(\theta, \text{Id})}{\partial \text{vec} U} \right) = \begin{pmatrix} \frac{\partial \varphi(\theta)}{\partial \theta} & 0_{n_\varphi \times n_Z^2} \\ \frac{\partial \text{vec} P(\theta)}{\partial \theta} & 0_{n_K^2 \times n_Z^2} \\ \frac{\partial \text{vec} Q(\theta)}{\partial \theta} & \text{Id}_{n_Z} \otimes Q(\theta) \\ \frac{\partial \text{vec} R(\theta)}{\partial \theta} & 0_{n_K n_W \times n_Z^2} \\ \frac{\partial \text{vec} S(\theta)}{\partial \theta} & \text{Id}_{n_Z} \otimes S(\theta) \\ \frac{\partial \text{vec} \Psi(\theta)}{\partial \theta} & \text{Id}_{n_Z} \otimes \Psi(\theta) - \Psi(\theta)' \otimes \text{Id}_{n_Z} \\ \frac{\partial \text{vec} \Sigma(\theta)}{\partial \theta} & -(\text{Id}_{n_Z^2} + T_{n_Z, n_Z})(\Sigma(\theta) \otimes \text{Id}_{n_Z}) \end{pmatrix},$$

then a necessary and sufficient rank condition for the DSGE model (1) satisfying the a priori restrictions (11) to be locally identified from the spectrum of $\{Y_t\}$ at θ_0 is: $\text{rank} \Delta_\varphi(\theta_0) = n_\theta + n_Z^2$. Moreover, a necessary order condition for identification is:

$$n_\varphi \geq n_\theta - \left[(n_K + n_W)(n_Z + n_K) + \frac{n_Z(n_Z + 1)}{2} \right].$$

A proof of Proposition 3 is in Appendix C. When the rank condition in Proposition 3 holds, we say that θ is *conditionally identified* at θ_0 , where the conditioning information is given by the constraints $\varphi(\theta) = 0$. Proposition 3 shows that in this case that is likely true for many models, the full rank condition on $\Delta(\theta_0)$ is not the appropriate identification condition as we also need to take into account the additional a priori restrictions. Typically, those restrictions would be fixing the values of some components of θ . Proposition 3 shows how many of these restrictions are necessary for identification, and is very useful in empirical work.

The fact that identification conditions for θ differ depending on whether or not $n_\theta \leq (n_K + n_W)(n_Z + n_K) + n_Z(n_Z + 1)/2$ is entirely new and not seen in the existing literature on identification. Aforementioned work by Canova and Sala (2009) and Iskrev (2007, 2008) does not include a priori restrictions on the deep parameter. The rank and order conditions for identification are simple and can be checked even prior to solving the model. This is unlike the problem of weak identification, which is a finite sample problem.

It is worth emphasizing that our identification conditions do not depend on any population moments of the data. This is what distinguishes our approach from either likelihood or moment based identification methods, which typically require full rank Fisher information matrix (see, e.g., Iskrev, 2007), or a full rank Hessian of the GMM objective function. Unlike $\Delta(\theta)$ and $\Delta_\varphi(\theta)$, these matrices depend on the sample moments of the data. Using our approach, the researchers can establish whether or not their DSGE models are identified prior to collecting any data, and thus abstracting from issues relating to measurement error. Since the solution parameter $\Lambda(\theta)$ is a function of the system matrices obtained in popular numerical solutions to DSGE models (which we will show), the rank conditions are easy to verify.

5 Examples

This section consists of three examples. Example 1 consists of one equation. Example 2 is a simple growth model with fewer shocks than variables in the system. This example is also used to illustrate how output from different solution algorithms can be used in evaluating the rank and order conditions. Example 3 is a model with three shocks and three endogenous variables. This example helps to understand the usefulness of the P, Q, R, S representation in identification.

5.1 Example 1

An and Schorfheide (2007) considered two structural models with the same reduced form:

$$y_t = \psi y_{t-1} + \eta_t, \quad \eta_t \sim \text{iid}(0, 1) \tag{13}$$

Model M_1 has $\psi = \rho$, and is given by:

$$y_t = \frac{1}{\alpha} E_t y_{t+1} + u_t, \quad u_t = \rho u_{t-1} + \epsilon_t, \quad \epsilon_t \sim \text{iid}(0, (1 - \rho/\alpha)^2)$$

with deep parameter $\theta \equiv (\alpha, \rho)$. Model M_2 has $\psi = \frac{1}{2}(\alpha - \sqrt{\alpha^2 - 4\phi\alpha})$ and is given by:

$$y_t = \frac{1}{\alpha} E_t y_{t+1} + \epsilon_t, \quad \epsilon_t \sim \text{iid}\left(0, \left[\frac{\alpha + \sqrt{\alpha^2 - 4\phi\alpha}}{2\alpha}\right]^2\right)$$

with deep parameter $\theta \equiv (\alpha, \phi)$.

The reduced form (13) is a special case of (1) obtained by setting $Z_t = \eta_t$, $K_t = y_t$, and leaving W_t is empty so $n_Z = 1$, $n_K = 1$, $n_W = 0$. The reduced form parameters are $P(\theta) = \psi$, $Q(\theta) = 1$, $\Psi(\theta) = 0$, and $\Sigma(\theta) = 1$. They satisfy Assumptions 1 through 4 whenever $0 < |\psi| < 1$. Now note that $n_\Lambda = (n_W + n_K)(n_K + n_Z) + 2n_Z^2 = 4$, $n_Z = 1$, $n_\theta = 2$, so for both models we have:

$$n_\theta = 2 < (n_W + n_K)(n_K + n_Z) + \frac{n_Z(n_Z + 1)}{2} = 3.$$

In other words, both models satisfy our order condition in Proposition 2. We now examine the rank condition. For this, note that the 4×3 matrix $\Delta(\theta)$ defined in Proposition 2 calculated for each of the two models is:

$$M_1 : \Delta(\theta) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad \text{and} \quad M_2 : \Delta(\theta) = \begin{pmatrix} \frac{1}{2} \left(1 - \frac{\alpha - 2\phi}{\sqrt{\alpha^2 - 4\phi\alpha}}\right) & \frac{\alpha}{\sqrt{\alpha^2 - 4\phi\alpha}} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Hence, for both models the rank of $\Delta(\theta) = 2 < n_\theta + n_Z^2 = 3$, so θ is not identifiable. This is not surprising as in model M_1 the component α of the deep parameter can never be identified from (13), while in M_2 the two components α and ϕ cannot be identified separately.

We now search for a priori restrictions that would allow for θ to be identified in each of the models. For model M_1 , any scalar restriction (11) such that $\partial\varphi/\partial\alpha \neq 0$ will be such that the corresponding 5×3 matrix $\Delta_\varphi(\theta)$ becomes of rank 3. Hence, θ will be identifiable under any such restriction. A simple example is the restriction $\alpha = \bar{\alpha}$, i.e. $\varphi(\theta) = \alpha - \bar{\alpha}$. Of course, letting $\varphi(\theta) = \rho - \bar{\rho}$ would fail to satisfy the rank condition. For model M_2 , the restrictions that will satisfy the rank condition must be such that:

$$\det \begin{pmatrix} \frac{\partial\varphi(\theta)}{\partial\alpha} & \frac{\partial\varphi(\theta)}{\partial\phi} \\ \frac{1}{2} \left(1 - \frac{\alpha - 2\phi}{\sqrt{\alpha^2 - 4\phi\alpha}}\right) & \frac{\alpha}{\sqrt{\alpha^2 - 4\phi\alpha}} \end{pmatrix} \neq 0.$$

For instance, setting $\alpha = \bar{\alpha}$ is sufficient to identify θ at any value $\theta_0 = (\alpha_0, \phi_0)$ such that $\alpha_0 \neq 0$. Similarly, setting $\phi = \bar{\phi}$ is also sufficient for identification provided $\phi_0 \neq 0$.

5.2 Example 2

Consider a one-sector stochastic growth model with inelastic labor supply, allowing for the possibility of costly adjustment of capital. The planner's problem is to choose consumption (\mathcal{C}_t) and capital (\mathcal{K}_t) to maximize her expected utility: $E_t \sum_{s=0}^{\infty} \beta^{t+s} \left(\frac{\mathcal{C}_{t+s}}{1-\nu} \right)$ subject to the given production technology and a feasibility constraint: $\mathcal{Q}_t = Z_t \mathcal{K}_t^\alpha (1 - \Phi_t)$ and $\mathcal{Q}_t = \mathcal{C}_t + \mathcal{K}_{t+1} - (1 - \delta)\mathcal{K}_t$. The exogenous technology process is specified as $\mathcal{Z}_t = \mathcal{Z}_0 \exp(z_t)$ where $z_t = \psi z_{t-1} + \epsilon_t$ and $\epsilon_t \sim WN(0, \sigma^2)$ with \mathcal{Z}_0 given. The adjustment cost function is: $\Phi_t(\mathcal{K}_{t+1}, \mathcal{K}_t) = \frac{\phi}{2} \left[\frac{\mathcal{K}_{t+1}}{\mathcal{K}_t} - 1 \right]^2$.

The solution to the log-linearized model takes the form:

$$\begin{aligned} k_{t+1} &= \lambda_{kk} k_t + \lambda_{kz} z_t \\ c_t &= \lambda_{ck} k_t + \lambda_{cz} z_t \end{aligned}$$

with $z_t = \psi z_{t-1} + \epsilon_t$ and $\epsilon_t \sim WN(0, \sigma^2)$, and where $(\lambda_{kk}, \lambda_{kz}, \lambda_{ck}, \lambda_{cz})$ are defined in Appendix D. Here, $P(\theta) = \lambda_{kk}$, $Q(\theta) = \lambda_{kz}$, $R(\theta) = \lambda_{ck}$, $S(\theta) = \lambda_{cz}$, $\Psi(\theta) = \psi$ and $\Sigma(\theta) = \sigma^2$, and the variables in (1) are $Z_t \equiv z_t$, $K_t \equiv k_t$, $W_t \equiv c_t$, with $n_Z = 1$, $n_K = 1$, and $n_W = 1$. The parameter of interest is $\theta \equiv (\alpha, \beta, \delta, \phi, \nu, \psi, \sigma^2)$. The reduced form parameter is $\Lambda(\theta) \equiv (\lambda_{ck}, \lambda_{cz}, \lambda_{kk}, \lambda_{kz}, \psi, \sigma^2)$. The parameter space Θ is chosen so that $\Lambda(\theta)$ satisfies Assumptions 1 through 4.

We consider three versions of the stochastic growth model, all with

$$\theta_0 = (\alpha_0, \beta_0, \delta_0, \phi_0, \nu_0, \psi_0, \sigma_0^2) = (0.36, 0.95, 0.025, \phi, \nu, 0.85, 0.04).$$

The three models differ in the restrictions imposed on ν and ϕ . As this model is sufficiently simple, the method of undetermined coefficients can be used to obtain numerical values for λ for any given θ . Appendix D shows how to use the algorithms of Sims (2002), Klein (2000), and Dynare to compute the P, Q, R, S representation.

Model I: With $\nu = 1, \phi = 0$, we have a log utility and capital is costless to adjust. In this version of the model, the dimension of the deep parameter θ is $n_\theta = 5$. Since $(n_K + n_W)(n_Z + n_K) + n_Z(n_Z + 1)/2 = 5$, the order condition of Proposition 2 holds. We now check the rank condition. The matrix $\Delta(\theta)$ evaluated at $\theta_0 = (0.36, 0.95, 0.025, 0.85, 0.04)$ has rank $6 = n_\theta + n_Z^2$ which satisfies the rank condition in Proposition 2. Hence, the value θ_0 of the deep parameter θ in Model I is identifiable from the spectrum of $\{(k_t, c_t)'\}$.

Model II: With $\nu > 0$ unrestricted and $\phi = 0$, we have a power utility and capital adjustment remains costless. Now, $n_\theta = 6$ and the order condition of Proposition 2 fails to hold. Since $n_\theta - (n_K + n_W)(n_Z + n_K) + n_Z(n_Z + 1)/2 = 1$, we need to introduce at least one restriction.

Consider fixing a particular component of θ to its true value, such as for example fixing β to 0.95. This corresponds to letting $\varphi(\theta) = \beta - 0.95$ in (11). We now need to check that the rank of the matrix $\Delta_\varphi(\theta)$ defined in Proposition 3 equals $n_\theta + n_Z^2 = 7$. When evaluated at $\theta_0 = (0.36, 0.95, 0.025, 2, 0.85, 0.04)$, the rank of $\Delta_\varphi(\theta)$ equals 6 which fails to satisfy the rank condition. Table 1 summarizes the rank results obtained by successively fixing one component θ_i of θ .

θ_i	α	β	δ	ν	ψ	σ
	7	6	7	7	6	7

Table 1: Rank of $\Delta_\varphi(\theta_0)$ in Model II ($\theta_0 = (0.36, 0.95, 0.025, 2, 0.85, 0.04)$)

As seen from Table 1 fixing any component of θ other than β or ψ is sufficient for θ_0 to be conditionally identifiable from the spectrum of $\{(k_t, c_t)'\}$.

Model III: With both $\nu > 0$ and $\phi > 0$ unrestricted, we extend Model II to allow for costly adjustment of capital. Now, $n_\theta = 7$ and $n_\theta - (n_K + n_W)(n_Z + n_K) + n_Z(n_Z + 1)/2 = 2$, so at least two restrictions are necessary to satisfy the order condition in Proposition 2. Similar to above, we consider all restrictions obtained by fixing the values of two distinct components θ_i and θ_j of θ . Table 2 reports the ranks of $\Delta_\varphi(\theta)$ evaluated at $\theta_0 = (0.36, 0.95, 0.025, 2, 0.03, 0.85, 0.04)$. Recall that for θ to be conditionally identified under $\varphi(\theta) = 0$ we now need $\text{rank } \Delta_\varphi(\theta_0) = n_\theta + n_Z^2 = 8$.

$\theta_i \backslash \theta_j$	β	δ	ϕ	ν	ψ	σ
α	8	8	8	8	7	8
β	-	8	8	8	7	8
δ		-	8	8	7	8
ϕ			-	8	7	8
ν				-	7	8
ψ					-	7

Table 2: Rank of $\Delta_\varphi(\theta_0)$ in Model III ($\theta_0 = (0.36, 0.95, 0.025, 2, 0.03, 0.85, 0.04)$)

As seen from Table 2 the two restrictions on α and β , for example, are sufficient to identify θ . The rank condition is, however, not trivially satisfied for all parameter restrictions. For example, restricting ψ and any other component of θ gives a reduced rank of $\Delta_\varphi(\theta)$. Note that the results of Table 2 suggest that when the constraints are put on β and ψ , i.e., when $\varphi(\theta) = (\beta - 0.95, \psi)'$, the point $\theta_1 = (0.36, 0.95, 0.025, 2, 0, 0.85, 0.04)$ is not a regular point of $\Delta_\varphi(\theta)$, i.e. the rank of $\Delta_\varphi(\theta_1)$ differs from the rank of $\Delta_\varphi(\theta)$ when θ is in a neighborhood of θ_1 . Indeed, we have that

$\text{rank } \Delta_\varphi(\theta_1) = 7$ while for any small departure of ψ from $\psi_1 = 0$ we get $\text{rank } \Delta_\varphi(\theta) = 8$.

5.3 Example 3

An and Schorfheide (2007) consider a model whose log-linearized solution is given by:

$$\begin{aligned}
y_t &= E_t y_{t+1} + g_t - E_t g_{t+1} - \frac{1}{\tau} (r_t - E_t \pi_{t+1} - E_t z_{t+1}) \\
\pi_t &= \beta E_t \pi_{t+1} + \kappa (y_t - g_t) \\
c_t &= y_t - g_t \\
r_t &= \rho_r r_{t-1} + (1 - \rho_r) \psi_1 \pi_t + (1 - \rho_r) \psi_2 (y_t - g_t) + \epsilon_{rt} \\
g_t &= \rho_g g_{t-1} + \epsilon_{gt} \\
z_t &= \rho_z z_{t-1} + \epsilon_{zt}
\end{aligned}$$

with $\epsilon_{rt} \sim WN(0, \sigma_r^2)$, $\epsilon_{gt} \sim WN(0, \sigma_g^2)$, and $\epsilon_{zt} \sim WN(0, \sigma_z^2)$ mutually uncorrelated. In the above model, $\kappa = \tau(1 - \nu)/(\nu \bar{\pi}^2 \phi)$, and $\bar{\pi}$ is steady state inflation rate. The parameter vector of interest is $\theta = (\tau, \beta, \kappa, \psi_1, \psi_2, \rho_r, \rho_g, \rho_z, \sigma_r^2, \sigma_g^2, \sigma_z^2)$ which is of dimension $n_\theta = 11$. The definition of θ in An and Schorfheide (2007) reflects the fact that ν , $\bar{\pi}$, and ϕ are not separately identified.

This model has three exogenous shocks $\epsilon_t \equiv (\epsilon_{rt}, \epsilon_{gt}, \epsilon_{zt})'$ ($n_Z = 3$), one observed state variable $K_t \equiv r_{t-1}$ ($n_K = 1$), two latent state processes z_t and g_t , and three additional endogenous variables $W_t \equiv (y_t, \pi_t, c_t)'$ ($n_W = 3$). Letting $Z_t \equiv (\epsilon_{rt}, g_t, z_t)'$, the P, Q, R, S representation of An and Schorfheide's (2007) model obtained by DYNARE (2009), Sims (2002) or Klein (2000) algorithm takes the form:

$$\begin{aligned}
K_{t+1} &= \lambda_{rr} K_t + (\lambda_{r\epsilon_r} \quad 0 \quad \lambda_{r\epsilon_z}) Z_t \\
W_t &= \begin{pmatrix} \lambda_{yr} \\ \lambda_{\pi r} \\ \lambda_{cr} \end{pmatrix} K_t + \begin{pmatrix} \lambda_{y\epsilon_g} & 1 & \lambda_{y\epsilon_z} \\ \lambda_{\pi\epsilon_g} & 0 & \lambda_{\pi\epsilon_z} \\ \lambda_{c\epsilon_g} & 0 & \lambda_{c\epsilon_z} \end{pmatrix} Z_t \\
Z_{t+1} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & \rho_g & 0 \\ 0 & 0 & \rho_z \end{pmatrix} Z_t + \epsilon_{t+1},
\end{aligned}$$

in which the λ coefficients are nonlinear functions of θ . Unlike in our Example 2, the analytic expression of different λ 's is unknown. Their numerical values are however easily obtained by using one of the algorithms above. The choice of the parameter space Θ is such that the above system satisfies our Assumptions 1 through 4.

In this example we have $n_\Lambda = (n_K + n_W)(n_Z + n_K) + 2n_Z^2 = 34$. Since $n_\theta = 11 < (n_K + n_W)(n_Z + n_K) + n_Z(n_Z + 1)/2 = 31$ the order condition in Proposition 2 is satisfied. We now check

the rank condition. For this, we choose θ_0 as in Table 3 of An and Schorfheide (2007),

$$\theta_0 = (2, 0.9975, 0.33, 1.5, 0.125, 0.75, 0.95, 0.9, 4 \times 10^{-6}, 36 \times 10^{-6}, 9 \times 10^{-6}).$$

Then, $\text{rank } \Delta(\theta_0) = 20 = n_\theta + n_Z^2$, so our rank condition in Proposition 2 is satisfied. Hence, θ_0 is identifiable from the spectrum of $\{(r_{t-1}, y_t, \pi_t, c_t)'\}$.

As a simple check of our procedure, if we let the parameter of interest be $\theta = (\tau, \beta, \nu, \phi, \bar{\pi}, \psi_1, \psi_2, \rho_r, \rho_g, \rho_z, \sigma_r^2, \sigma_g^2, \sigma_z^2)$ then $n_\theta = 13$ which still satisfies the order condition. However, we now have that $\text{rank } \Delta(\theta_0) = 21 < n_\theta + n_Z^2 = 22$ which shows that θ_0 is not identifiable. This is not surprising as we have already noted that the components ν , ϕ , and $\bar{\pi}$ are not separately identifiable.

6 Discussion

In the preceding sections, we obtained rank and order conditions for identification of θ from the spectral density $\Omega(z; \theta)$ of the observables by assuming that the shocks are fundamental, and that the model has a minimal state representation. Both properties were shown to hold under Assumptions 1 through 4. There are however important examples of DSGE models in which those assumptions fail. For example, when there are sunspot shocks of the type discussed in Beyer and Farmer (2007). To accommodate this possibility, the state vector may be larger than necessary, which violates the minimal state property need to obtain Lemma 3.

In systems in which some of our assumptions fail, the identification arguments developed in the previous sections no longer apply. In particular, we can no longer rely on the results of Proposition 1 to characterize all observationally equivalent values of the reduced form parameter $\Lambda(\theta)$ in (1). Instead, we need to look directly at the equivalence relation (5):

$$H(z; \theta_1) \Sigma(\theta_1) H(z^{-1}; \theta_1)' = H(z; \theta_0) \Sigma(\theta_0) H(z^{-1}; \theta_0)',$$

for all z in \mathbb{C} , and find the necessary and sufficient conditions under which the above holds only when $\theta_1 = \theta_0$. Without Proposition 1, all we can use for identification are the spectra or the autocovariance generating functions. This is essentially the approach taken in Iskrev (2009), who asked whether θ can be identified from the autocovariance matrices $\Gamma(j; \theta)$ of the observables. In this case, one looks at the necessary and sufficient conditions under which the infinite system of equalities:

$$\Gamma(j; \theta_1) = \Gamma(j; \theta_0)$$

obtained for every j ($-\infty < j < +\infty$) can only hold if $\theta_1 = \theta_0$. Whether it is based on the spectrum or on the autocovariance matrices, this identification approach remains valid even when the assumptions of Lemmas 3 and 4 do not hold. However, for the problem considered here, it

amounts to studying local injectivity properties of an infinite system of nonlinear equations which is both theoretically challenging and computationally difficult.

The starting point of our analysis is to fully exploit restrictions that need to exist between observationally equivalent parameters of the DSGE model (1). Unsurprisingly, observational equivalence in a dynamic setting involves different restrictions on the model parameters than those in static models. We provide an operational definition of spectral equivalence and show that some reduced form parameters of DSGE models are generally not identifiable. We then use these restrictions to develop necessary and sufficient conditions for identification. Our results have implications for estimation, both using full information method, and particularly for two step estimation that requires a first step estimation of the reduced form model. This issue is analyzed in our companion paper Komunjer and Ng (2009).

Appendix A Proofs of Section 2

Proof of Lemma 2. Recall that for every $t \geq 1$ we have:

$$Y_t = H(L; \theta)\epsilon_t$$

where the transfer function $H(z; \theta)$ is an $(n_K + n_W) \times n_Z$ matrix of polynomials defined in (3). To show invertibility we use the following result: If for every $\theta \in \Theta$, $H(z^{-1}; \theta)$ is of rank n_Z in $|z| > 1$, then $H(z^{-1}; \theta)$ has a left inverse, i.e. there exists a matrix $G(z^{-1}; \theta)$ of dimension $n_Z \times (n_K + n_W)$ such that $G(z^{-1}; \theta)H(z^{-1}; \theta) = \text{Id}$. This implies that we can write $G(L; \theta)Y_t = \epsilon_t$ so the process is invertible. We now show that for every $\theta \in \Theta$, $\text{rank } H(z^{-1}; \theta) = n_Z$ for every $z \in \mathbb{C}$ such that $|z| > 1$. From (3) we have that:

$$\begin{aligned} \text{rank } H(z^{-1}; \theta) &= \text{rank} \left(\begin{array}{c} z^{-1}[\text{Id} - P(\theta)z^{-1}]^{-1}Q(\theta)[\text{Id} - \Psi(\theta)z^{-1}]^{-1} \\ \{R(\theta)z^{-1}[\text{Id} - P(\theta)z^{-1}]^{-1}Q(\theta) + S(\theta)\}[\text{Id} - \Psi(\theta)z^{-1}]^{-1} \end{array} \right) \\ &= \text{rank} \left[\begin{pmatrix} z[\text{Id} - P(\theta)z^{-1}] & 0 \\ -R(\theta) & \text{Id} \end{pmatrix}^{-1} \begin{pmatrix} Q(\theta) \\ S(\theta) \end{pmatrix} [\text{Id} - \Psi(\theta)z^{-1}]^{-1} \right] \\ &= \text{rank} \begin{pmatrix} Q(\theta) \\ S(\theta) \end{pmatrix} = n_Z \quad \text{for every } |z| > 1 \end{aligned}$$

where the second to last equality follows by Assumption 2, while the last equality follows by combining Assumptions 3 and 4(ii). \square

Appendix B Proofs of Section 3

In the proofs of the results stated in Section 3, we define the state vector

$$S_t \equiv \begin{pmatrix} Z_t \\ K_t \end{pmatrix}$$

and we use the following state space representation of the DSGE model (1):

$$\begin{aligned} S_{t+1} &= A(\theta)S_t + B(\theta)\epsilon_{t+1} \\ Y_t &= C(\theta)S_t \end{aligned} \tag{B.1}$$

where $Y_t = (K_t' \ W_t)'$ as before, while the matrices $A(\theta)$, $B(\theta)$, $C(\theta)$ are given by:

$$A(\theta) \equiv \begin{pmatrix} \Psi(\theta) & 0 \\ Q(\theta) & P(\theta) \end{pmatrix}, \quad B(\theta) \equiv \begin{pmatrix} \text{Id} \\ 0 \end{pmatrix}, \quad C(\theta) \equiv \begin{pmatrix} 0 & \text{Id} \\ S(\theta) & R(\theta) \end{pmatrix} \tag{B.2}$$

Recall that for every $t \geq 1$ we have $Y_t = H(L; \theta)\epsilon_t$. The transfer function $H(z; \theta)$ can be related to the matrices $(A(\theta), B(\theta), C(\theta))$ in (B.2) as:

$$H(z; \theta) = C(\theta) [\text{Id} - zA(\theta)]^{-1} B(\theta) \tag{B.3}$$

In what follows, we call $(A(\theta), B(\theta), C(\theta))$ a *realization* of the transfer function $H(z; \theta)$. For simplicity, we drop θ from the notations below. We first give several definitions that have been adapted from control theory (see, e.g., Antsaklis and Michel, 1997):

Definition 2 (Controllability) Let \mathcal{C} be the $(n_Z + n_K) \times (n_Z(n_Z + n_K))$ matrix defined as:

$$\mathcal{C} \equiv (B \quad AB \quad \dots \quad A^{n_Z+n_K-1}B)$$

A realization (A, B, C) is controllable if and only if $\text{rank} \mathcal{C} = n_Z + n_K$.

Definition 3 (Observability) Let \mathcal{O} be the $(n_K + n_W)(n_Z + n_K) \times (n_Z + n_K)$ matrix defined as:

$$\mathcal{O} \equiv \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n_Z+n_K-1} \end{pmatrix}$$

A realization (A, B, C) is observable if and only if $\text{rank} \mathcal{O} = n_Z + n_K$.

Definition 4 (Minimality) A realization (A, B, C) is minimal (irreducible, of least order) if and only if it is both controllable and observable.

Note that the matrix \mathcal{C} is called the controllability matrix, while \mathcal{O} is referred to as the observability matrix.

We first show the following Lemma whose result will be useful in our subsequent proofs.

Lemma 6 Let Assumptions 1 through 4 hold. Then, for every $\theta \in \Theta$, the realization $(A(\theta), B(\theta), C(\theta))$ in (B.2) is minimal.

Proof of Lemma 6. The proof of Lemma 6 is done in two steps.

STEP 1. We first show controllability. For this, note that for any $1 \leq k \leq n_Z + n_K - 1$, we have:

$$A^k B = \begin{pmatrix} \Psi^k \\ P^{k-1}Q + P^{k-2}Q\Psi + \dots + PQ\Psi^{k-2} + Q\Psi^{k-1} \end{pmatrix}$$

so it holds that for every $1 \leq k \leq n_Z + n_K - 1$:

$$A^k B - A^{k-1} B \Psi = \begin{pmatrix} 0 \\ P^{k-1}Q \end{pmatrix} \tag{B.4}$$

Now consider the following $(n_Z(n_Z + n_K)) \times (n_Z(n_Z + n_K))$ matrix \mathcal{J} :

$$\mathcal{J} \equiv \begin{pmatrix} \text{Id} & -\Psi & & (0) \\ & \ddots & \ddots & \\ & & \ddots & -\Psi \\ (0) & & & \text{Id} \end{pmatrix}$$

Since \mathcal{J} is nonsingular, we have that:

$$\text{rank}(\mathcal{C}\mathcal{J}) = \text{rank}(\mathcal{C})$$

Now using the property in (B.4), we have that:

$$\mathcal{C}\mathcal{J} = \begin{pmatrix} \text{Id} & 0 & 0 & \dots & 0 \\ 0 & Q & PQ & \dots & P^{n_Z+n_K-2}Q. \end{pmatrix}$$

Under Assumption 4(i) we have that $\text{rank}(Q) = n_K$. Then, it holds that:

$$\text{rank} \begin{pmatrix} \text{Id} & 0 \\ 0 & Q \end{pmatrix} = n_Z + n_K$$

The above implies that $\text{rank}\mathcal{C} = n_Z + n_K$, so (A, B, C) is controllable.

STEP 2. To show observability, first note that for any $1 \leq k \leq n_Z + n_K - 1$, we have:

$$CA^k = \begin{pmatrix} P^{k-1}Q + P^{k-2}Q\Psi + \dots + PQ\Psi^{k-2} + Q\Psi^{k-1} & P^k \\ RP^{k-1}Q + RP^{k-2}Q\Psi + \dots + RPQ\Psi^{k-2} + RQ\Psi^{k-1} + S\Psi^k & RP^k \end{pmatrix}$$

Now consider the $(n_K + n_W)(n_Z + n_K) \times (n_K + n_W)(n_Z + n_K)$ matrix \mathcal{Q} defined as:

$$\mathcal{Q} \equiv \begin{pmatrix} \text{Id} & 0 & 0 & 0 & & (0) \\ 0 & \text{Id} & 0 & 0 & & \\ -P & 0 & \text{Id} & 0 & & \\ 0 & 0 & -R & \text{Id} & & \\ & \ddots & \ddots & \ddots & \ddots & \\ & & -P & 0 & \text{Id} & 0 \\ (0) & & 0 & 0 & -R & \text{Id} \end{pmatrix}$$

Note that $\det\mathcal{Q} = 1$ so \mathcal{Q} is of full rank and we have:

$$\text{rank}(\mathcal{Q}\mathcal{O}) = \text{rank}\mathcal{O}$$

Now, straightforward calculations yield:

$$\mathcal{Q}\mathcal{O} = \begin{pmatrix} 0 & \text{Id} \\ S & R \\ Q & 0 \\ S\Psi & 0 \\ \vdots & \vdots \\ Q\Psi^{n_Z+n_K-2} & 0 \\ S\Psi^{n_Z+n_K-1} & 0 \end{pmatrix}$$

Under Assumption 4(ii) it follows that

$$\text{rank} \begin{pmatrix} 0 & \text{Id} \\ S & R \\ Q & 0 \end{pmatrix} = n_Z + n_K$$

which shows that $\text{rank}(\mathcal{O}) = \text{rank}(\mathcal{Q}\mathcal{O}) = n_Z + n_K$, and (A, B, C) is observable. \square

Proof of Lemma 3. According to Lemma 6, in our DSGE model (1), under Assumptions 1 to 4, every realization $(A(\theta), B(\theta), C(\theta))$ of the transfer function $H(z; \theta)$ in (B.3) is minimal. We can then use the following standard result from control theory:

Theorem 3.10 (Antsaklis and Michel, 1997). *Let $(A(\theta_0), B(\theta_0), C(\theta_0))$ and $(A(\theta_1), B(\theta_1), C(\theta_1))$ be realizations of $H(z; \theta_0)$. If $(A(\theta_0), B(\theta_0), C(\theta_0))$ is a minimal realization, then $(A(\theta_1), B(\theta_1), C(\theta_1))$ is also a minimal realization if and only if the two realizations are equivalent, i.e., if and only if there exists a nonsingular $(n_K + n_W) \times (n_K + n_W)$ matrix \mathcal{T} such that $A(\theta_1) = \mathcal{T}A(\theta_0)\mathcal{T}^{-1}$, $B(\theta_1) = \mathcal{T}B(\theta_0)$, and $C(\theta_1) = C(\theta_0)\mathcal{T}^{-1}$.*

Now our model can be put in an (A, B, C) form like (B.2), and by Lemma 6, all realizations are minimal. By the above theorem, they then must be related by a similarity transformation. Given that K_t is observed, our \mathcal{T} matrix must be of the form:

$$\mathcal{T} = \begin{pmatrix} \text{Id} & 0 \\ 0 & T \end{pmatrix}$$

with T full rank. Plugging in the definitions of A, B, C gives the result as desired. \square

Proof of Lemma 4. We first recall several useful results from control theory Anderson, Hitz, and Diem (1974). In the spectral factorization problem, one is given a $n \times n$ matrix $\phi(z)$ of real rational functions of a complex variable z , with $\phi(z^{-1})' = \phi(z)$ and $\phi(\exp(i\omega))$ positive semi-definite for all real ω . One is required to find a $W(z)$ such that $W(z^{-1})'W(z) = \phi(z)$, with $W(z)$ real and rational, of entries analytic in $|z| \geq 1$, and possibly with $W(z)$ of constant rank in $|z| > 1$. A spectral factor $W_d(z)$ of dimension $r \times n$ is termed *minimum phase* if $W_d(z^{-1})'W_d(z) = \phi(z)$ and $W_d(z)$ has rank r everywhere in $|z| > 1$ where r is equal to the rank almost everywhere of $\phi(z)$, i.e. $\text{rank}\phi(z) = r$ for almost every $z \in \mathbb{C}$. As shown in Anderson (1969) or Anderson, Hitz, and Diem (1974), for example, minimum phase spectral factors are unique to within left multiplication by an arbitrary real constant orthogonal $r \times r$ matrix V , i.e. any other minimum phase factor $\widetilde{W}_d(z)$ is of the form:

$$\widetilde{W}_d(z) = VW_d(z) \quad \text{where} \quad V'V = \text{Id}$$

Now, fix $\theta_0 \in \Theta$ and consider a reduced form parameter $\Lambda(\theta_0) = (P(\theta_0), Q(\theta_0), R(\theta_0), S(\theta_0), \Psi(\theta_0), \Sigma(\theta_0))$ that generates the spectral density $\Omega(z; \theta_0)$ of the observables $\{Y_t\}$. Under Assumption 1, the matrix $\Sigma(\theta_0)$ is real, symmetric and positive definite and $L(\theta_0)$ is the lower triangular matrix obtained by the Cholesky decomposition of $\Sigma(\theta_0)$, i.e. $\Sigma(\theta_0) = L(\theta_0)L(\theta_0)'$. Note that since $\Sigma(\theta_0)$ has real entries, $L(\theta_0)$ can be assumed to have real entries as well. In addition, $L(\theta_0)$ is positive

definite. Now consider the following $n_Z \times (n_K + n_W)$ matrix:

$$W_d(z) \equiv L(\theta_0)'H(z^{-1}; \theta_0)'. \quad (\text{B.5})$$

We have

$$\begin{aligned} W_d(z^{-1})'W_d(z) &= H(z; \theta_0)L(\theta_0)L(\theta_0)'H(z^{-1}; \theta_0)' \\ &= H(z; \theta_0)\Sigma(\theta_0)H(z^{-1}; \theta_0)' \\ &\equiv \Omega(z; \theta_0) \end{aligned}$$

for every $z \in \mathbb{C}$. Note that by (3), $\Omega(z; \theta_0)$ is a matrix of real rational functions of a complex variable z , with $\Omega(z^{-1}; \theta_0)' = \Omega(z; \theta_0)$ and $\Omega(\exp(i\omega); \theta_0)$ positive semi-definite for all real ω .

We now show that $\text{rank } W_d(z) = n_Z$ for every $z \in \mathbb{C}$ such that $|z| > 1$. From the proof of Lemma 2 we know that (3) $\text{rank } H(z^{-1}; \theta_0) = n_Z$ for every $|z| > 1$. Using (B.5) and the fact that $L(\theta_0)$ is positive definite, then shows that $\text{rank } W_d(z) = n_Z$ for every $|z| > 1$. It remains to be shown that $\text{rank } \Omega(z; \theta_0) = n_Z$ for almost every $z \in \mathbb{C}$. For this, recall again from the proof of Lemma 2 that:

$$H(z^{-1}; \theta_0) = \begin{pmatrix} z[\text{Id} - P(\theta_0)z^{-1}] & 0 \\ -R(\theta_0) & \text{Id} \end{pmatrix}^{-1} \begin{pmatrix} Q(\theta_0) \\ S(\theta_0) \end{pmatrix} [\text{Id} - \Psi(\theta_0)z^{-1}]^{-1}$$

which is rank deficient only for those $z \in \mathbb{C}$ that are such that $\det[z\text{Id} - P(\theta_0)]\det[z\text{Id} - \Psi(\theta_0)] = 0$, i.e. only at the eigenvalues of the matrices $P(\theta_0)$ and $\Psi(\theta_0)$. Given that there is at most $n_Z + n_K$ different eigenvalues, $H(z^{-1}; \theta_0)$ is of full rank almost everywhere in \mathbb{C} . Since by Assumption 1, $\Sigma(\theta_0)$ is full rank, it follows that $\Omega(z; \theta_0)$ is of rank n_Z almost everywhere in \mathbb{C} .

We can now apply the result in Anderson, Hitz, and Diem (1974) to show that if $\widetilde{W}_d(z) \equiv L(\theta_1)'H(z^{-1}; \theta_1)'$ is another minimum phase factor, then necessarily we have:

$$L(\theta_1)'H(z^{-1}; \theta_1)' = \widetilde{W}_d(z) = VW_d(z) = VL(\theta_0)'H(z^{-1}; \theta_0)'$$

where V is an orthogonal $n_Z \times n_Z$ matrix, $V'V = \text{Id}$. Now let:

$$U \equiv L(\theta_0)V'L(\theta_1)^{-1}$$

be a full rank $n_Z \times n_Z$ matrix. It then follows that $H(z; \theta_1) = H(z; \theta_0)U$ for every $z \in \mathbb{C}$ and $UL(\theta_1)V = L(\theta_0)$, so $UL(\theta_1)VV'L(\theta_1)'U^{-1} = L(\theta_0)L(\theta_0)'$, i.e. $U\Sigma(\theta_1)U' = \Sigma(\theta_0)$. \square

Proof of Proposition 1. The proof is obtained by combining the results of Lemma 3 and Lemma 4. From Lemma 4 we know that $\Lambda(\theta_0)$ and $\Lambda(\theta_1)$ are observationally equivalent if and only if, for every $z \in \mathbb{C}$ we have $H(z; \theta_1) = H(z; \theta_0)U$ with U nonsingular. Expressing this equality in terms

of the (A, B, C) matrices in (B.3) and using the similarity transformation result of Lemma 3 then gives the following necessary and sufficient condition for observational equivalence of $\Lambda(\theta_0)$ and $\Lambda(\theta_1)$:

$$A(\theta_1) = \mathcal{T}A(\theta_0)\mathcal{T}^{-1}, \quad B(\theta_1) = \mathcal{T}B(\theta_0)U, \quad C(\theta_1) = C(\theta_0)\mathcal{T}^{-1}, \quad (\text{B.6})$$

where U is a nonsingular $n_Z \times n_Z$ matrix, and \mathcal{T} is a nonsingular $(n_K + n_W) \times (n_K + n_W)$ matrix. Let

$$\mathcal{T} \equiv \begin{pmatrix} T_1 & T_2 \\ T_3 & T_4 \end{pmatrix}$$

be a full rank $(n_K + n_Z) \times (n_K + n_Z)$ matrix, with sub-matrices (T_1, T_2, T_3, T_4) of dimensions $n_K \times n_K$, $n_K \times n_Z$, $n_Z \times n_K$, and $n_Z \times n_Z$, respectively. Then, writing the above equations for $(B(\theta_1), A(\theta_1), C(\theta_1))$ gives the following:

$$0 = T_2 U \quad (\text{B.7})$$

$$\text{Id} = T_4 U \quad (\text{B.8})$$

$$P(\theta_1)T_1 + Q(\theta_1)T_3 = T_1P(\theta_0) \quad (\text{B.9})$$

$$P(\theta_1)T_2 + Q(\theta_1)T_4 = T_1Q(\theta_0) + T_2\Psi(\theta_0) \quad (\text{B.10})$$

$$\Psi(\theta_1)T_3 = T_3P(\theta_0) \quad (\text{B.11})$$

$$\Psi(\theta_1)T_4 = T_3Q(\theta_0) + T_4\Psi(\theta_0) \quad (\text{B.12})$$

$$R(\theta_1)T_1 + S(\theta_1)T_3 = R(\theta_0) \quad (\text{B.13})$$

$$R(\theta_1)T_2 + S(\theta_1)T_4 = S(\theta_0) \quad (\text{B.14})$$

$$T_1 = \text{Id} \quad (\text{B.15})$$

$$T_2 = 0 \quad (\text{B.16})$$

Now, we can rewrite Equation (B.11) as:

$$(\text{Id}_{n_K} \otimes \Psi(\theta_1)) \text{vec}(T_3) = (P(\theta_0)' \otimes \text{Id}_{n_Z}) \text{vec}(T_3)$$

where $\text{vec}(T_3)$ denotes the vectorization of the matrix T_3 formed by stacking the columns of T_3 into a single column vector. It follows that:

$$[(\text{Id}_{n_K} \otimes \Psi(\theta_1)) - (P(\theta_0)' \otimes \text{Id}_{n_Z})'] \text{vec}(T_3) = 0$$

which has a unique solution $T_3 = 0$ if and only if $\Psi(\theta_1)$ and $P(\theta_0)$ have no eigenvalues in common (Assumption 2). Then, combining the above with (B.15), (B.16), and (B.8) gives that the matrix \mathcal{T} is necessarily of the form:

$$\mathcal{T} = \begin{pmatrix} \text{Id} & 0 \\ 0 & U^{-1} \end{pmatrix}$$

Then, from (B.9) and (B.10) we get:

$$P(\theta_1) = P(\theta_0) \quad \text{and} \quad Q(\theta_1) = Q(\theta_0)U \quad (\text{B.17})$$

Results for $R(\theta_1)$, $S(\theta_1)$, and $\Psi(\theta_1)$ are obtained from Equations (B.13), (B.14), and (B.12), respectively. Using the fact that in addition $\Sigma(\theta_1) = U^{-1}\Sigma(\theta_0)U^{-1'}$ from Lemma 4 then gives the desired result. \square

Appendix C Proofs of Section 4

Proof of Lemma 5 The proof is similar to that of Theorem 1 in Glover and Willems (1974). The necessity is straightforward. We now prove sufficiency by proving the contrapositive: suppose that θ_0 is not locally identifiable. Then there exists an infinite sequence of deep parameter vectors $\{\theta_1, \dots, \theta_k, \dots\}$ (of dimension n_θ) approaching θ_0 such that:

$$\Omega(z; \theta_k) = \Omega(z; \theta_0) \quad \text{for all } z \in \mathbb{C}.$$

From Proposition 1 we know that the above holds if and only if there exists an infinite sequence of full rank $n_Z \times n_Z$ matrices $\{T_1, \dots, T_k, \dots\}$ such that:

$$P(\theta_k) = P(\theta_0), \quad Q(\theta_k)T_k = Q(\theta_0), \quad R(\theta_k) = R(\theta_0), \quad S(\theta_k)T_k = S(\theta_0), \quad (\text{C.1})$$

$$T_k^{-1}\Psi(\theta_k)T_k = \Psi(\theta_0) \quad \text{and} \quad T_k^{-1}\Sigma(\theta_k)T_k^{-1'} = \Sigma(\theta_0). \quad (\text{C.2})$$

In other words, using the notation in (8) we have that:

$$\lambda(\theta_k, T_k) = \lambda(\theta_0, \text{Id})$$

In order to show that the mapping λ is not locally injective, it suffices to show that the sequence $\{T_1, \dots, T_k, \dots\}$ approaches the identity matrix Id . For this, note that from (C.1) we have:

$$\begin{pmatrix} Q(\theta_k) \\ S(\theta_k) \end{pmatrix} T_k = \begin{pmatrix} Q(\theta_0) \\ S(\theta_0) \end{pmatrix}$$

Now, from Assumption 4(ii) we know that the $(n_K + n_W) \times n_Z$ matrix $(Q(\theta_k)', S(\theta_k)')'$ is of rank n_Z ; hence, it is left invertible and we have

$$T_k = \left[(Q(\theta_k)' \quad S(\theta_k)') \begin{pmatrix} Q(\theta_k) \\ S(\theta_k) \end{pmatrix} \right]^{-1} (Q(\theta_k)' \quad S(\theta_k)') \begin{pmatrix} Q(\theta_0) \\ S(\theta_0) \end{pmatrix}$$

By continuity of $Q(\theta)$ and $S(\theta)$ (Assumption 5), it then follows that the sequence $\{T_1, \dots, T_k, \dots\}$ approaches the identity matrix Id . Hence, λ is not injective in the neighborhood of $(\theta_0, \text{Id}_{n_Z})$. \square

Proof of Proposition 2 The proof of Proposition 2 is done in two steps.

STEP 1. We start by establishing the necessary and sufficient (rank) condition. To compute the matrix of partial derivatives of λ , we use the product rule: if $F(X)$ and $G(X)$ are, respectively, $m \times p$ - and $p \times q$ -variate and differentiable functions with respect to X , then

$$\frac{\partial \text{vec}(F(X)G(X))}{\partial \text{vec}X} = (G(X)' \otimes \text{Id}_m) \frac{\partial \text{vec}(F(X))}{\partial \text{vec}X} + (\text{Id}_q \otimes F'(U)) \frac{\partial \text{vec}(G(X))}{\partial \text{vec}X}$$

For example, if X is an $n_Z \times n_Z$ matrix ($X \in \mathbb{R}^{n_Z \times n_Z}$), then applying the above rule to $F(X) = X^{-1}$ (respectively $F(X) = X'$) and $G(X) = X$ we get the following useful results:

$$\frac{\partial \text{vec}(X^{-1})}{\partial \text{vec}X} = -(X^{-1'} \otimes X^{-1}) \quad \text{and} \quad \frac{\partial \text{vec}(X^{-1'})}{\partial \text{vec}X} = -T_{n_Z, n_Z}(X^{-1'} \otimes X^{-1})$$

where T_{n_Z, n_Z} is the $n_Z^2 \times n_Z^2$ permutation matrix that transforms $\text{vec}X$ into $\text{vec}X'$, i.e. $T_{n_Z, n_Z} \text{vec}X = \text{vec}X'$. Note that T_{n_Z, n_Z} is an orthogonal matrix: $T_{n_Z, n_Z} = T_{n_Z, n_Z}^{-1}$ and $T_{n_Z, n_Z} = T_{n_Z, n_Z}'$. In addition, note that we have $\text{rank}(\text{Id}_{n_Z^2} + T_{n_Z, n_Z}) = \frac{n_Z(n_Z+1)}{2}$.

STEP 1A. We first compute the partial derivatives of $\lambda(\theta, U)$ with respect to the components of θ . Using the product rule, it immediately follows that:

$$\frac{\partial \lambda(\theta, U)}{\partial \theta} = \begin{pmatrix} \frac{\partial P(\theta)}{\partial \theta} \\ (U' \otimes \text{Id}_{n_Z}) \frac{\partial Q(\theta)}{\partial \theta} \\ \frac{\partial R(\theta)}{\partial \theta} \\ (U' \otimes \text{Id}_{n_Z}) \frac{\partial S(\theta)}{\partial \theta} \\ (U' \otimes U^{-1}) \frac{\partial \Psi(\theta)}{\partial \theta} \\ (U^{-1} \otimes U^{-1}) \frac{\partial \Sigma(\theta)}{\partial \theta} \end{pmatrix} \quad (\text{C.3})$$

STEP 1B. We now compute the derivatives of $\lambda(\theta, U)$ with respect to U . For $\text{vec}(Q(\theta)U)$ and $\text{vec}(S(\theta)U)$ applying the product rule immediately gives

$$\frac{\partial \text{vec}(Q(\theta)U)}{\partial \text{vec}U} = (\text{Id}_{n_Z} \otimes Q(\theta)) \quad (\text{C.4})$$

$$\frac{\partial \text{vec}(S(\theta)U)}{\partial \text{vec}U} = (\text{Id}_{n_Z} \otimes S(\theta)) \quad (\text{C.5})$$

For $\text{vec}(U^{-1}\Psi(\theta)U)$, we let $F(U) = U^{-1}$ and $G(U) = \Psi(\theta)U$. Then,

$$\begin{aligned} \frac{\partial \text{vec}(U^{-1}\Psi(\theta)U)}{\partial \text{vec}U} &= -((\Psi(\theta)U)' \otimes \text{Id}_{n_Z})(U^{-1'} \otimes U^{-1}) + (\text{Id}_{n_Z} \otimes U^{-1})(\text{Id}_{n_Z} \otimes \Psi(\theta)) \\ &= -((U'\Psi(\theta)'U^{-1'}) \otimes U^{-1}) + (\text{Id}_{n_Z} \otimes (U^{-1}\Psi(\theta))) \\ &= (U' \otimes U^{-1})((\text{Id}_{n_Z} \otimes \Psi(\theta)) - (\Psi(\theta)' \otimes \text{Id}_{n_Z}))(U^{-1'}) \otimes \text{Id}_{n_Z} \end{aligned} \quad (\text{C.6})$$

Note that if the rank of $\Delta(\theta)$ remains constant in a neighborhood of θ_0 , then the rank of the partial derivatives of λ remains constant in a neighborhood of (θ_0, Id) .

STEP 2. We now establish the necessary (order) condition. Counting the number of linearly independent rows in the matrix $\Delta(\theta)$ in (C.9), it is easy to show that:

$$\text{rank}\Delta(\theta) \leq (n_W + n_K)(n_Z + n_K) + n_Z^2 + \frac{n_Z(n_Z + 1)}{2}$$

where we have used the fact that $\Sigma(\theta)$ is always symmetric so:

$$\text{rank} \left(\frac{\partial \text{vec}\Sigma(\theta)}{\partial \theta} \quad -(\text{Id}_{n_Z^2} + T_{n_Z, n_Z})(\Sigma(\theta) \otimes \text{Id}_{n_Z}) \right) \leq \frac{n_Z(n_Z + 1)}{2}.$$

Hence, a necessary condition for (9) to hold is that:

$$n_\theta + n_Z^2 \leq (n_W + n_K)(n_Z + n_K) + n_Z^2 + \frac{n_Z(n_Z + 1)}{2} \quad \text{i.e.} \quad n_\theta \leq (n_K + n_W)(n_Z + n_K) + \frac{n_Z(n_Z + 1)}{2}$$

□

Proof of Proposition 3 To begin, note that the matrix of partial derivatives of χ can be written as:

$$\left(\frac{\partial \chi(\theta, U)}{\partial \theta} \quad \frac{\partial \chi(\theta, U)}{\partial \text{vec}U} \right) = \begin{pmatrix} \frac{\partial \varphi(\theta)}{\partial \theta} & 0_{n_\varphi \times n_Z^2} \\ \frac{\partial \lambda(\theta, U)}{\partial \theta} & \frac{\partial \lambda(\theta, U)}{\partial \text{vec}U} \end{pmatrix}$$

Using the same reasoning as in the proof of Proposition 2, the above can further be written as:

$$\left(\frac{\partial \chi(\theta, U)}{\partial \theta} \quad \frac{\partial \chi(\theta, U)}{\partial \text{vec}U} \right) = \begin{pmatrix} \text{Id}_{n_\varphi} & 0_{n_\varphi \times n_\Lambda} \\ 0_{n_\Lambda \times n_\varphi} & M(U) \end{pmatrix} \Delta_\varphi(\theta) N(U)$$

where the $n_\Lambda \times n_\Lambda$ matrix $M(U)$ and $(n_\theta + n_Z^2) \times (n_\theta + n_Z^2)$ matrix $N(U)$ are as in the proof of Proposition 2, and where we have let:

$$\begin{aligned} \Delta_\varphi(\theta) &\equiv \left(\frac{\partial \chi(\theta, \text{Id})}{\partial \theta} \quad \frac{\partial \chi(\theta, \text{Id})}{\partial \text{vec}U} \right) = \begin{pmatrix} \frac{\partial \varphi(\theta)}{\partial \theta} & 0_{n_\varphi \times n_Z^2} \\ & \Delta(\theta) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial \varphi(\theta)}{\partial \theta} & 0_{n_\varphi \times n_Z^2} \\ \frac{\partial \text{vec}P(\theta)}{\partial \theta} & 0_{n_K^2 \times n_Z^2} \\ \frac{\partial \text{vec}Q(\theta)}{\partial \theta} & \text{Id}_{n_Z} \otimes Q(\theta) \\ \frac{\partial \text{vec}R(\theta)}{\partial \theta} & 0_{n_K n_W \times n_Z^2} \\ \frac{\partial \text{vec}S(\theta)}{\partial \theta} & \text{Id}_{n_Z} \otimes S(\theta) \\ \frac{\partial \text{vec}\Psi(\theta)}{\partial \theta} & \text{Id}_{n_Z} \otimes \Psi(\theta) - \Psi(\theta)' \otimes \text{Id}_{n_Z} \\ \frac{\partial \text{vec}\Sigma(\theta)}{\partial \theta} & -(\text{Id}_{n_Z^2} + T_{n_Z, n_Z})(\Sigma(\theta) \otimes \text{Id}_{n_Z}) \end{pmatrix} \end{aligned}$$

It then follows that:

$$\text{rank} \left(\frac{\partial \chi(\theta, U)}{\partial \theta} \quad \frac{\partial \chi(\theta, U)}{\partial \text{vec}U} \right) = \text{rank}\Delta_\varphi(\theta)$$

from which the rank condition follows. For the order condition, note that we again have:

$$\text{rank} \left(\frac{\partial \chi(\theta, U)}{\partial \theta} \quad \frac{\partial \chi(\theta, U)}{\partial \text{vec} U} \right) \leq n_\varphi + (n_K + n_W)(n_Z + n_K) + n_Z^2 + \frac{n_Z(n_Z + 1)}{2}$$

so a necessary condition for identifiability of θ under the a priori restrictions (11) is:

$$n_\theta + n_Z^2 \leq n_\varphi + (n_K + n_W)(n_Z + n_K) + n_Z^2 + \frac{n_Z(n_Z + 1)}{2} \quad \text{i.e.} \quad n_\varphi \geq n_\theta - \left[(n_K + n_W)(n_Z + n_K) + \frac{n_Z(n_Z + 1)}{2} \right]$$

□

Appendix D Details of Example 2 in Section 5

Let Q^*, C^*, K^* be steady state values of output, consumption, and capital, and let $R^* = (1 - \delta) + \alpha \frac{Q^*}{K^*}$.

The log-linearized model is

$$\begin{aligned} q_t &= \alpha k_t + z_t \\ q_t &= (1 - \delta \frac{K^*}{Q^*}) c_t + \frac{K^*}{Q^*} k_{t+1} - (1 - \delta) \frac{K^*}{Q^*} k_t \\ 0 &= E_t \left[\nu (c_{t+1} - c_t) - \phi \beta \frac{Q^*}{K^*} k_{t+2} + (\alpha(1 - \alpha) + \beta \phi (2 - \delta)) \beta \frac{Q^*}{K^*} k_{t+1} \right. \\ &\quad \left. - \phi \beta (1 - \delta) \frac{Q^*}{K^*} k_t - \alpha \beta \frac{Q^*}{K^*} z_{t+1} \right] \\ z_t &= \psi z_{t-1} + \sigma \epsilon_t \end{aligned}$$

The structural model can be rewritten as

$$\begin{aligned} q_t &= \gamma_1 k_t + z_t \\ q_t &= \gamma_2 c_t + \gamma_3 k_{t+1} + (1 - \gamma_2 - \gamma_3) k_t \\ &= E_t \left[\gamma_4 (c_{t+1} - c_t) + \gamma_5 k_{t+2} + \left(\frac{\gamma_1 (1 - \gamma_1)}{\gamma_1 + \gamma_2 + \gamma_3 - 1} + \frac{(1 - \gamma_2 - 2\gamma_3)\gamma_5}{\gamma_3} \right) k_{t+1} \right. \\ &\quad \left. - (1 - \gamma_2 - \gamma_3) \frac{\gamma_5}{\gamma_3} k_t - \frac{\gamma_1}{\gamma_1 + \gamma_2 + \gamma_3 - 1} z_{t+1} \right] \\ z_t &= \gamma_5 z_{t-1} + \gamma_6 \epsilon_t \end{aligned}$$

where the *structural parameters* are

$$\gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6) \equiv \left(\alpha, \quad 1 - \delta \frac{K^*}{Q^*}, \quad \frac{K^*}{Q^*}, \quad \nu, \quad -\phi \beta \frac{Q^*}{K^*}, \quad \psi, \quad \sigma \right).$$

For this model, λ_{kk} is given by:

$$\lambda_{kk} = \frac{-B + \sqrt{B^2 - 4AC}}{2A}$$

where

$$A \equiv -\nu \frac{K^*}{C^*} - \phi\beta \frac{Y^*}{K^*}, \quad B \equiv \nu \left(1 + \frac{1}{\beta}\right) \frac{K^*}{C^*} + [\alpha(1 - \alpha) + \phi(2 - \delta)]\beta \frac{Y^*}{K^*}, \quad C \equiv -\frac{\nu}{\beta} \frac{K^*}{C^*} - \phi\beta(1 - \delta) \frac{Y^*}{K^*}.$$

And λ_{kz} , λ_{ck} , λ_{cz} equal:

$$\begin{aligned} \lambda_{kz} &= \frac{(1 - \psi)\nu \frac{Y^*}{C^*} + \alpha\beta\psi \frac{Y^*}{K^*}}{\nu \left[\frac{1}{\beta} - \lambda_{kk} + 1 - \psi\right] \frac{K^*}{C^*} + \beta[\alpha(1 - \alpha) - \phi(2 - \delta) - \phi\lambda_{kk} - \psi\phi] \frac{Y^*}{K^*}} \\ \lambda_{ck} &= \frac{K^*}{C^*} \left[\frac{1}{\beta} - \lambda_{kk}\right] \\ \lambda_{cz} &= \frac{K^*}{C^*} \left[\frac{Y^*}{K^*} - \lambda_{kz}\right]. \end{aligned}$$

References

- AN, S., AND F. SCHORFHEIDE (2007): “Bayesian Analysis of DSGE Models,” *Econometric Reviews*, 26: 2-4, 113–172.
- ANDERSON, B. D. O. (1969): “The Inverse Problem of Stationary Covariance Generation,” *Journal of Statistical Physics*, 1, 133–147.
- ANDERSON, B. D. O., K. L. HITZ, AND N. D. DIEM (1974): “Recursive Algorithm for Spectral Factorization,” *IEEE Transactions on Circuits and Systems*, 21, 742–750.
- ANDERSON, G., AND G. MOORE (1985): “A linear algebraic procedure for solving linear perfect foresight models,” *Economic Letters*, 17(3), 247–252.
- ANTSAKLIS, P., AND A. MICHEL (1997): *Linear Systems*. McGraw-Hill, New York.
- BEYER, A., AND R. FARMER (2007): “Indeterminacy of Determinacy and Indeterminacy,” *American Economic Review*, 97:1, 524–529.
- CANOVA, F., AND L. SALA (2009): “Back to Square One: Identification Issues in DSGE Models,” *Journal of Monetary Economics*, 56, 431–449.
- CHAO, J., AND P. PHILLIPS (1998): “Bayesian Posterior Distributions in Limited Information Analysis of the Simultaneous Equation Model Using the Jeffreys Prior,” *Journal of Econometrics*, 87, 4986.
- CHARI, V. V., P. KEHOE, AND E. MCGRATTAN (2009): “New Keynesian Models: Not Yet Useful for Policy Analysis,” *American economic Journal: Macroeconomics*, 1:1, 242–66.
- CONSOLO, A., C. FAVERO, AND A. PACCAGNINI (2009): “On the Statistical Identification of DSGE Models,” *Journal of Econometrics*, 150, 99–115.
- CURDIA, V., AND R. REIS (2009): “Correlated Disturbances and U.S. Business Cycles,” unpublished manuscript.
- DEISTLER, M. (1976): “The Identifiability of Linear Econometric Models with Autocorrelated Errors,” *International Economic Review*, 17, 26–46.
- DEISTLER, M. (2006): “Stationary Processes and Linear Systems,” in *Harmonic Analysis and Rational Approximation*, ed. by J.-D. Fournier, J. Grimm, J. Leblond, and J. R. Partington, pp. 159–179. Springer, Berlin.

- DYNARE (2009): “Version 4,” .
- FERNANDEZ-VILLAYERDE, F., J. RUBIO-RAMIREZ, T. SARGENT, AND M. WATSON (2007): “A,B,Cs and (D)’s for Understanding VARs,” *American Economic Review*, 97:3, 1021–1026.
- FISHER, F. (1966): *The Identification Problem in Econometrics*. McGraw-Hill.
- GIANNONE, D., AND L. REICHLIN (2006): “Does Information Help Recover Structural Shocks From Past Observations,” *Journal of the European Economic Association*, 4:2-3, 455–465.
- GLOVER, K. (1973): “Structural Aspects of System Identification,” Unpublished Dissertation, M.I.T.
- GLOVER, K., AND J. WILLEMS (1974): “Parameterizations of Linear Dynamical Systems: Canonical Forms and Identifiability,” *IEEE Transactions on Automatic Control*, 19:6, 640–646.
- GOURIEROUX, C., AND A. MONFORT (1997): *Time Series and Dynamic Models*. Cambridge University Press.
- GREWAL, K., AND K. GLOVER (1976): “Identifiability of Linear and Nonlinear Dynamic Systems,” *IEEE Transactions on Automatic Control*, 21:6, 833–837.
- HANNAN, E. (1971): “The Identification Problem for Multiple Equation Systems with Moving Average Errors,” *Econometrica*, 39:5, 751–765.
- HANNAN, E., AND M. DIESTLER (1988): *The Statistical Theory of Linear Systems*. Wiley, New York.
- HATANAKA, M. (1975): “On the Global Identification of the Dynamic Simultaneous Equations Model with Stationary Disturbances,” *International Economic Review*, 16:3, 545–553.
- ISKREV, N. (2007): “Evaluating the Information Matrix in Linearized DSGE models,” *Economics Letters*, 99:3, 607–610.
- (2008): “How Much Do We Learn from Estimation of DSGE models? A Case Study of Identification Issues in a New Keynesian Business Cycle Model,” University of Michigan.
- (2009): “Local Identification in DSGE Models,” Banco de Portugal.
- KING, R. G., AND M. W. WATSON (2002): “System Reduction and Solution Algorithms for Singular Linear Difference Systems under Rational Expectations,” *Computational Economics*, 20, pp. 57–86.

- KLEIBERGEN, F., AND H. K. VAN DIJK (1998): “Bayesian Simultaneous Equations Analysis Using Reduced Rank Structures,” *Econometric Theory*, 14, 701743.
- KLEIN, P. (2000): “Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model,” *Journal of Economic Dynamics and Control*, 4:2, 257–271.
- KOMUNJER, I., AND S. NG (2009): “Issues in Limited Information Identification of DSGE models,” unpublished manuscript, Columbia University.
- KOOPMANS, T. C. (ed.) (1950): *Statistical Inference in Dynamic Economic Models*, vol. 10 of *Cowles Commission Monograph*. John Wiley & Sons, Inc.
- LUTKEPOHL, H. (2005): *New Introduction to Multiple Time Series Analysis*. Springer Verlag, Berlin.
- NEWCOMB, R., AND B. ANDERSON (1968): “On the Generation of All Spectral Factors,” *IEEE Transactions on Information Theory*, 14:3, 512–514.
- POIRIER, D. (1998): “Revisign Beliefs in Nonidentified Models,” *Econometric Theory*, 14, 183–209.
- REINSEL, G. (2003): *Elements of Multivariate Time Series Analysis*. Springer, New York, 2nd edn.
- ROTHENBERG, T. (1971): “Identification in Parametric Models,” *Econometrica*, 39:3, 577–591.
- RUBIO-RAMÍREZ, J. F., D. F. WAGGONER, AND T. ZHA (2007): “Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference,” *Review of Economics and Studies*, Forthcoming.
- SIMS, C. (2002): “Solving Linear Rational Expectations Models,” *Computational Economics*, pp. 1–20.
- SOLO, V. (1986): “Topics in Advanced Time Series,” in *Lectures in Probability and Statistics*, ed. by G. del Pino, and R. Rebolledo, vol. 1215. Springer-Verlag.
- UHLIG, H. (1999): “A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily,” in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon, and A. Scott, pp. 30–61. Oxford University Press.
- WALLIS, K. (1977): “Multiple Time Series Analysis and the Final Form of Econometric Models,” *Econometrica*, 45:6, 1481–1497.
- ZELLNER, A., AND F. PALM (1974): “Time Series Analysis and Simultaneous Equation Econometric Models,” *Journal of Econometrics*, 2, 17–54.