Mutschler, Willi

Conference Paper

Identification of DSGE Models - A Comparison of Methods and the Effect of Second Order Approximation

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Time Series Analysis, No. C16-V1

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Mutschler, Willi (2014) : Identification of DSGE Models - A Comparison of Methods and the Effect of Second Order Approximation, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Time Series Analysis, No. C16-V1

This Version is available at:
http://hdl.handle.net/10419/100598

Identification of DSGE models – A comparison of methods and the effect of second order approximation *

Willi Mutschler†

Institute for Econometrics and Economic Statistics,
Center for Quantitative Economics (CQE),
University of Münster, Germany

Version: February 20, 2014
CQE Working paper.
Comments are greatly appreciated.

Abstract

Several methods have been introduced into the literature on identification in DSGE models, however, a comparative study is missing. The contribution of this paper is threefold: First, all methods are derived analytically in the same notation and framework. Second, it is shown how to extend the methods to linear approximations to the second order. It is argued that this can improve overall identification of a DSGE model via imposing additional restrictions on the mean. Third, all methods are applied on DSGE models that are known to have lack of identification. Do the methods come to the same conclusion? Also practical difficulties as well as the difference between using analytical and numerical derivatives are discussed.

JEL Classification: C10, C51, C52, E1

*This paper was presented at the Econometrics and Empirical Economics Seminar in Münster, Germany (2013); at the course on Identification Analysis and Global Sensitivity Analysis for Macroeconomic Models in Ispra, Italy (2013); and at the Fall Macro Midwest Meeting in Minneapolis, USA (2013). The author thanks Nikolay Iskrev, Michel Juillard, Junior Maih, Zhongjun Qu, Marco Ratto, Stephanie Schmitt-Grohé, Ron Smith, Mark Trede, Bernd Wilfing and seminar participants for critical and helpful comments and encouragement.

†Corresponding author. Center for Quantitative Economics, Institut für Ökonometrie und Wirtschaftsstatistik, Westfälische Wilhelms-Universität Münster, Am Stadtgraben 9, 48143 Münster, Tel.: +49-251-83-22914, Fax: +49-251-83-22012, Email: willi.mutschler@uni-muenster.de. The most recent version as well as the Matlab code can be found on www.mutschler.eu/identification.html.

1 Introduction

Many different methods for solving and estimating DSGE models have been developed and used in order to get a detailed description and thorough estimation of dynamic macroeconomic relationships. Recently, the question of identification of DSGE models has proven to be of major importance, especially since identification of a model precedes estimation and inference. Identification problems arise if distinct parameter values do not lead to distinct probability distributions of the data. Even with an infinite sample it is not possible to pin down some parameters, no matter what estimation procedure one uses.[1] In a full-information setting this often evokes a badly shaped likelihood function, which modern Bayesian estimation can conveniently circumvent by using tight priors. However, the comparison of prior and posterior can be misleading, since data need not be informative about parameters (Canova and Sala 2009; Koop, Pesaran, and Smith 2012). Lack of identification leads thus to wrong conclusions from estimation and inference, whereas the source of identification can hugely influence empirical findings (Ríos-Rull et al. 2012).

Identification criteria in the classical literature are stated in terms of rank and order conditions based upon implications of a unique solution to a system of equations.[2] Identification criteria in the DSGE literature, however, are mainly concerned with two mappings: one from the deep parameters to the reduced-form parameters, i.e. the uniqueness of the solution; and one from the solution to observable data, i.e. the uniqueness of the probability distribution. An adaption of classical concepts is not as straightforward, since the reduced-form parameters are non-linear functions of the deep parameters and there might be complicated cross-equation restrictions, which can often only be evaluated numerically (Komunjer and Ng 2011). Moreover, priors tend to mask the problem (Guerron-Quintana, Inoue, and Kilian 2012), whereas calibrating unidentified parameters can lead to wrong conclusions, since other parameters might depend on the calibrated ones (Canova and Sala 2009).

Nevertheless several formal methods have been proposed to check identification in DSGE models via (i) the autocovariogram (Iskrev 2010), (ii) the spectral density (Komunjer and Ng 2011; Qu and Tkachenko 2012), or (iii) Bayesian indicators (Koop, Pesaran, and Smith 2012). Even though all methods seem similar, there has been – to our knowledge – no study of the advantages and drawbacks of implementing the different methods. A comparative approach is worthwhile, so that a researcher gets robust indication before the model is actually taken to data.

Methodically, we will derive all criteria in the same framework and model representation following Schmitt-Grohé and Uribe (2004). It will be shown that the methods heavily depend on the accuracy of computing derivatives and ranks. For a rigorous comparison we will thus derive analytical derivatives of all solution matrices up to second order, moments and spectral density with respect to the deep parameters. While Iskrev (2010) already uses analytical derivatives, Komunjer and Ng (2011) and Tkachenko and Qu (2012) rely on numerical methods. We show how to implement analytical derivatives into the latter criteria such that we are able to discuss the effect of using analytical compared to numerical derivatives for each method and across criteria. For computing the ranks (and null spaces) we use the singular value decomposition across different tolerance levels.

[1] Mutschler (2012) estimates the An and Schorfheide (2007) model on simulated data via GMM, Impulse-Response-Matching, Maximum-Likelihood and Bayesian methods and finds that no estimation method is able to pin down the parameters in the Taylor-rule. Beltran and Draper (2012) have similar findings using different sample sizes.

[2] For a good textbook overview of the early contributions see Fisher (1966) and Hsiao (1983) and the references therein.

Moreover, all methods focus on the linear solution of the DSGE model to the first order. We will relax this assumption and show how to extend the identification criteria given a linear approximation of the policy functions to the second order using the Magnus-Neudecker definition of the Hessian (Gomme and Klein 2011). The representation of the solution yields additional restrictions on the mean that can be used to identify previously unidentified parameters. However, this comes with a computational burden, since the matrices become larger and the quadratic approximation requires appropriate filter techniques.

To make our comparison and exposition illustrative, all methods are applied on three simple models: (i) the neoclassical growth model, (ii) the Kim (2003) model as a well-known non-identified model, and (iii) the An and Schorfheide (2007) as a prototypical DSGE model. We will show that the methods yield similar results and discuss the reasons the criteria differ. Also it will be shown that an approximation to the second order identifies the Kim (2003) model, whereas overall identification of the An and Schorfheide (2007) model increases slightly, and under some mild restrictions, we are able to identify all parameters of the Taylor-rule.

The paper is accompanied by MATLAB code, in which a researcher can choose in a user-friendly interface between the illustrative models, which parameters to identify at which local point, the identification test as well as the specific options, analytical or numerical derivatives, and the order of approximation. Since all procedures are model independent, other models can be easily included and tested as long as they can be represented in the same framework.[3]

2 DSGE framework

2.1 The model

A DSGE model consists of equilibrium conditions and transition equations for state (x_t) and control (y_t) variables, structural shocks and innovations (u_t), which can be cast into a nonlinear first-order system of expectational difference equations f. We will follow the framework of Schmitt-Grohé and Uribe (2004) and introduce a perturbation parameter σ that scales the standard deviation of the exogenous innovations and captures the stochastic nature of the model. If it is set to 0 the model becomes a deterministic one, for $\sigma > 0$ the model is stochastic. The solution of such rational expectation models is characterized by so-called *policy-functions*, g and h, that solve (at least approximately) the system of equations f. The model is further augmented by an observer or measurement equation with D being a known matrix, which selects a subset (d_t) of the control variables (or a linear combination thereof) that are observable. The observer equation can also capture measurement and specification errors v_t. Thus, DSGE models can be interpreted as *state-space models*.

Let E_t be the expectation operator conditional on information available at time t, then

$$0 = E_t f\left(x_{t+1}, y_{t+1}, x_t, y_t|\theta\right), \tag{1}$$

$$x_{t+1} = h(x_t, \sigma|\theta) + \sigma\eta_x(\theta)\varepsilon_{t+1}, \tag{2}$$

$$y_t = g(x_t, \sigma|\theta), \tag{3}$$

$$d_t = Dy_t + \eta_d(\theta)\varepsilon_t \tag{4}$$

is called the general DSGE model with deep parameters θ. The non-stochastic steady-state is

[3]The code is available on http://www.mutschler.eu/identification.html.

given by $\bar{z} := (\bar{x}', \bar{y}')'$

$$f(\bar{x}, \bar{y}, \bar{x}, \bar{y}|\theta) := f(\bar{z}|\theta) = 0, \qquad \bar{x} = h(\bar{x}, 0|\theta), \qquad \bar{y} = g(\bar{x}, 0|\theta), \qquad \bar{d} = D\bar{y}.$$

The goal of the model is to describe fluctuations around the steady-state as well as the response of the model due to stochastic innovations. Hence, the driving force of the model is the vector of exogenous shocks and innovations. In the proposed setup the stochastic variables u_t and v_t are iid with $E(u_t) = E(v_t) = 0$, $E(u_t u_t') = I_{n_u}$ and $E(v_t v_t') = I_{n_v}$. The variances and cross-correlations are captured by premultiplying matrices η_u and η_v respectively, denoting $\Sigma_u := \sigma^2 \eta_u \eta_u'$ and $\Sigma_v := \eta_v \eta_v'$. For convenience, stochastic innovations and measurement errors are stacked into a common vector $\varepsilon_t = (u_t', v_t')'$. η_x and η_d then choose and scale the standard deviation of the shocks and measurement errors accordingly. σ is set to be dependent on the standard deviation of one of the shocks, while scaling all other variances and cross-correlations accordingly. The state and control vectors can be handled very flexible depending on the model. In particular, the state vector can be partitioned into n_{x_1} endogenous predetermined states and n_{x_2} exogenous predetermined states. In the MATLAB code we also append auxiliary equations ($e_{t+1} = E_t(\varepsilon_{t+1}) = 0$) to the model and consider the state vector $x_t = (x_{1,t}', x_{2,t}', e_{t+1}')'$, so that the policy functions also include reactions to current innovations and measurement errors.[4] Also, in the code we add the measurement equations to the model equations such that the control vector becomes $(y_t', d_t')'$ and D simply picks the last n_d entries. See appendix E on how to squeeze the example models into this framework.

The important property for identification is that d_t is weakly stationary and its time series properties are completely characterized by its time invariant unconditional mean and autocovariances. The same result can be motivated by assuming u_t and v_t are i.i.d. Gaussian. Table 1 summarizes the notation and dimensions of the coefficients, variables and mappings. In the following the dependence on the deep parameters θ will be dropped for notational convenience.

2.2 Local approximation

We will follow Gomme and Klein (2011)'s approach to approximate the policy functions using the Magnus and Neudecker (1999) definition of the Hessian. Define the steady state as $\bar{z} := (\bar{x}', \bar{y}')'$, then the Jacobian $\mathcal{D}f(\bar{z})$ and Hessian $\mathcal{H}f(\bar{z})$ of f evaluated at the steady-state are defined as:

$$\begin{aligned}
\mathcal{D}f(\bar{z}) &:= \begin{pmatrix} \frac{\partial f(\bar{z})}{\partial x_{t+1}'} & \frac{\partial f(\bar{z})}{\partial y_{t+1}'} & \frac{\partial f(\bar{z})}{\partial x_t'} & \frac{\partial f(\bar{z})}{\partial y_t'} \end{pmatrix} := \begin{pmatrix} f_1 & f_2 & f_3 & f_4 \end{pmatrix}, \\
\mathcal{H}f(\bar{z}) &:= \mathcal{D}vec((\mathcal{D}f(\bar{z}))').
\end{aligned}$$

This definition simplifies the analytical derivatives, since no tensor notation is needed and basic matrix algebra can be used, see appendix A for further reference.[5] The second-order Taylor

[4] This is insofar practical as the equivalence to other solution algorithms becomes obvious. For instance Sims (2001) linear solution is of the form $z_{t+1} = \mathcal{A}z_t + \mathcal{B}\varepsilon_{t+1}$ with $z_t = (y_t', x_t')'$. The entries of the matrix $\mathcal{B}$ can be extracted from the last columns of g_x and h_x respectively, see also appendix D.

[5] For recent literature in favor of this definition see also Magnus (2010) and Pollock (2013).

Variable	Dimension	Interpretation
θ	$\Theta \subseteq \mathbb{R}_0^{n_\theta}$	Vector of deep parameters
σ	$\mathbb{R}_0^+$	Perturbation parameter that scales the variance of ε_t
D	$n_d \times n_y$	Selection matrix
$\eta_u(\theta)$	$n_x \times n_u$	Correlation matrix between shocks, depending on θ
$\eta_v(\theta)$	$n_d \times n_v$	Correlation matrix between measurement errors, depending on θ
$\eta_x(\theta)$	$n_x \times (n_u + n_v)$	Correlation matrix between shocks and measurement errors, depending on θ
$\eta_d(\theta)$	$n_d \times (n_u + n_v)$	Correlation matrix between shocks and measurement errors, depending on θ
.....	.....	.....
x_t	$n_x \times 1$	Vector of state variables at time t
y_t	$n_y \times 1$	Vector of control variables at time t
d_t	$n_d \times 1$	Vector of observable variables
$\overline{z} = (\overline{x}', \overline{y}')'$	$(n_x + n_y) \times 1$	Vector of steady-state variables
u_t	$n_u \times 1$	Vector of stochastic innovations at time t, $E(u_t) = 0$, $E(u_t {u_t}') = I_{n_u}$
v_t	$n_v \times 1$	Vector of measurement errors at time t, $E(v_t) = 0$, $E(v_t {v_t}') = I_{n_v}$
$\varepsilon_t = ({u_t}', {v_t}')'$	$(n_u + n_v) \times 1$	Vector of shocks and measurement errors at time t $E(\varepsilon_t) = 0$, $E(\varepsilon_t {\varepsilon_t}') = I_{n_\varepsilon}$
.....	.....	.....
$\mu_x(\theta)$	$n_x \times 1$	Unconditional expectation of state variables, depending on θ
$\mu_y(\theta)$	$n_y \times 1$	Unconditional expectation of control variables, depending on θ
$\mu_d(\theta)$	$n_d \times 1$	Unconditional expectation of observable variables, depending on θ
$\Sigma_x(\theta)$	$n_x \times n_x$	Unconditional covariance matrix of states, depending on θ
$\Sigma_y(\theta)$	$n_y \times n_y$	Unconditional covariance matrix of controls, depending on θ
$\Sigma_d(\theta)$	$n_d \times n_d$	Unconditional covariance matrix of observables, depending on θ
$\Sigma_x(t\|\theta)$	$n_x \times n_x$	Unconditional autocovariogram of states with lag t, depending on θ
$\Sigma_y(t\|\theta)$	$n_y \times n_y$	Unconditional autocovariogram of controls with lag t, depending on θ
$\Sigma_d(t\|\theta)$	$n_d \times n_d$	Unconditional autocovariogram of observables with lag t, depending on θ
.....	.....	.....
$f(x_{t+1}, y_{t+1}, x_t, y_t\|\theta)$	$\mathbb{R}^{2n_x+2n_y} \to \mathbb{R}^{n_x+n_y}$	Mapping of structural equations, depending on θ
$g(x_t, \sigma\|\theta)$	$\left(\mathbb{R}^{n_x} \times \mathbb{R}^+\right) \to \mathbb{R}^{n_y}$	Policy function for control variables, depending on θ
$h(x_t, \sigma\|\theta)$	$\left(\mathbb{R}^{n_x} \times \mathbb{R}^+\right) \to \mathbb{R}^{n_x}$	Policy function for state variables, depending on θ
.....	.....	.....
$f_1(\overline{z}\|\theta)$	$n \times n_x$	Derivative of f w.r.t x_{t+1} evaluated at $\overline{z}$, depending on θ
$f_2(\overline{z}\|\theta)$	$n \times n_y$	Derivative of f w.r.t y_{t+1} evaluated at $\overline{z}$, depending on θ
$f_3(\overline{z}\|\theta)$	$n \times n_x$	Derivative of f w.r.t x_t evaluated at $\overline{z}$, depending on θ
$f_4(\overline{z}\|\theta)$	$n \times n_y$	Derivative of f w.r.t y_t evaluated at $\overline{z}$, depending on θ
$H(\overline{z}\|\theta)$	$n(2n_x + 2n_y) \times (2n_x + 2n_y)$	Magnus-Neudecker Hessian of f evaluated at $\overline{z}$, depending on θ
$g_x(\theta)$	$n_y \times n_x$	First order solution matrix, depending on θ
$h_x(\theta)$	$n_x \times n_x$	First order solution matrix, depending on θ
$g_{xx}(\theta)$	$n_y n_x \times n_x$	Second order solution matrix, depending on θ
$h_{xx}(\theta)$	$n_x^2 \times n_x$	Second order solution matrix, depending on θ
$g_{\sigma\sigma}(\theta)$	$n_y \times 1$	Second order solution matrix, depending on θ
$h_{\sigma\sigma}(\theta)$	$n_x \times 1$	Second order solution matrix, depending on θ
.....	.....	.....
$M(q)$	$n_d + \frac{n_d(n_d+1)}{2} + (q-1)n_d^2 \times n_\theta$	Jacobian w.r.t. θ of first and second moments up to lag q
J	$n_d + \frac{n_d(n_d+1)}{2} + (q-1)n_d^2 \times n_\theta$	Jacobian w.r.t. θ of first moments and solution matrices
.....	.....	.....
$dX = \frac{\partial vec(X)}{\partial \theta}$	$(n_1 + n_2) \times n_\theta$	If X is $n_1 \times n_2$ then dX is the derivative of $vec(X)$ w.r.t θ

Table 1: Notation and dimensions

approximation at the non-stochastic steady-state is then given by

$$\begin{aligned} x_{t+1} &= \bar{x} + h_x(\bar{x},0)\cdot(x_t-\bar{x}) + \sigma\eta_x\varepsilon_{t+1} \\ &\quad + \frac{1}{2}\left[I_{n_x}\otimes(x_t-\bar{x})'\otimes(x_t-\bar{x})'\right]vec(h_{xx}(\bar{x},0)') + \frac{1}{2}\sigma^2 h_{\sigma\sigma}(\bar{x},0), \qquad (5) \\ y_t &= \bar{y} + g_x(\bar{x},0)\cdot(x_t-\bar{x}) \\ &\quad + \frac{1}{2}\left[I_{n_y}\otimes(x_t-\bar{x})'\otimes(x_t-\bar{x})'\right]vec(g_{xx}(\bar{x},0)') + \frac{1}{2}\sigma^2 g_{\sigma\sigma}(\bar{x},0). \qquad (6) \end{aligned}$$

g_x and g_σ are the gradients of g with respect to x_t and σ respectively, g_{xx} and $g_{\sigma\sigma}$ the corresponding Magnus-Neudecker-Hessians, all evaluated at $(\bar{x},0)$. The same notation applies to h_x, h_σ, h_{xx} and $h_{\sigma\sigma}$. Schmitt-Grohé and Uribe (2004) show that all linear terms as well as cross terms in σ, i.e. $g_\sigma, g_{x\sigma}, g_{\sigma x}, h_\sigma, h_{x\sigma}, h_{\sigma x}$, are equal to zero, since the approximation is around $\sigma = 0$. Notice also, that in a linearization of the first order (or log-linearization) the terms with the $1/2$ drop out.

There are several methods and algorithms to calculate the matrices $h_x(\bar{x},0)$ and $g_x(\bar{x},0)$, since these are the coefficients of a first order linearization or log-linearization of the model. We follow Klein (2000) to obtain $h_x(\bar{x},0)$ and $g_x(\bar{x},0)$ using the generalized Schur decomposition.[6] The other matrices can be calculated by inserting (3) and (2) into (1):

$$E_t f\left(h(x_t,\sigma)+\sigma\eta_x\varepsilon_{t+1}, g(h(x_t,\sigma)+\sigma\eta_x\varepsilon_{t+1},\sigma), x_t, g(x_t,\sigma)\right) = 0.$$

This expression is known at the non-stochastic steady-state, thus, all derivatives of f must be 0 when evaluated at the steady state. Differentiating f twice using the chain-rule of Magnus and Neudecker (1999, p. 110), evaluating the Jacobian $\mathcal{D}f = (f_1\ f_2\ f_3\ f_4)$ and Hessian H of f at the non-stochastic steady-state, and setting it to 0 yields (after some algebra):

$$\begin{bmatrix} vec(g_{xx}) \\ vec(h_{xx}) \end{bmatrix} = -Q^{-1}vec(A), \qquad (7)$$

$$\begin{bmatrix} h_{ss} \\ g_{ss} \end{bmatrix} = -B^{-1}C, \qquad (8)$$

with

$$\begin{aligned} Q &= \begin{bmatrix} h_x'\otimes f_2\otimes h_x' + I_{n_x}\otimes f_4\otimes I_{n_x} & I_{n_x}\otimes(f_1\otimes I_{n_x} + f_2 g_x\otimes I_{n_x}) \end{bmatrix}, \\ A &= (I_{n_x+n_y}\otimes M')HM, \\ B &= \begin{bmatrix} f_1 + f_2 g_x & f_2 + f_4 \end{bmatrix}, \\ C &= f_2 trm[(I_{n_y}\otimes(\eta_x\eta_x'))g_{xx}] + trm[(I_{n_x+n_y}\otimes N')HN(\eta_x\eta_x')], \\ M &= \begin{bmatrix} h_x \\ g_x h_x \\ I_{n_x} \\ g_x \end{bmatrix}, \quad N = \begin{bmatrix} I_{n_x} \\ g_x \\ 0_{(n_x+n_y)\times n_x} \end{bmatrix}, \end{aligned}$$

and trm defines the matrix trace of an $nm\times n$ matrix $[Y_1'\ Y_2'\ \dots\ Y_m']'$ as the $m\times 1$ vector

[6] See Anderson (2008) for a comparison of algorithms, which are basically all equivalent and differ only (slightly) in computational burden. Further, all provide and check the Blanchard and Kahn (1980) conditions that are necessary in order to have a stable saddle-path solution, i.e. a unique mapping between state and control variables.

$[tr(Y_1)\ tr(Y_2)\ \dots\ tr(Y_m)]'$. See Gomme and Klein (2011) for the derivation. For our purpose it is sufficient to note that there exist analytical closed-form solutions that we will differentiate with respect to the deep parameters in section 2.4.

2.3 Unconditional Moments

The important assumption for identification is (weak) stationarity of model variables and data. Given the approximated solution (5) and (6) of the DSGE model we can calculate the unconditional moments. Note that we will follow the literature and focus on the concept of second order accurate moments, that is to keep only first and second order terms in the formula to compute moments. Indeed, when computing the variance, the only possible second order term is the product of two first order approximations. But this is the same as when computing the variance from the first order approximation alone, since all other cross-product terms would result in something of order higher than two.

Denoting Σ_x, Σ_y and Σ_d for the variances of x_t, y_t and d_t respectively, we get

$$\begin{aligned}
\Sigma_x &:= E(x_t - \bar{x})(x_t - \bar{x})' = h_x E(x_{t-1} - \bar{x})(x_{t-1} - \bar{x})' h_x' + \underbrace{\sigma^2 \eta_x E(\varepsilon_t \varepsilon_t') \eta_x'}_{=:\Sigma_{\varepsilon_x}} \\
&\Leftrightarrow \Sigma_x = h_x \Sigma_x h_x' + \Sigma_{\varepsilon_x}, \qquad (9) \\
\Sigma_y &:= E(y_t - \bar{y})(y_t - \bar{y})' = g_x E(x_t - \bar{x})(x_t - \bar{x})' g_x' \\
&\Leftrightarrow \Sigma_y = g_x \Sigma_x g_x', \\
\Sigma_d &:= E(d_t - \bar{d})(d_t - \bar{d})' = E(D(y_t - \bar{y}) + \eta_d \varepsilon_t)(D(y_t - \bar{y}) + \eta_d \varepsilon_t)' \\
&\Leftrightarrow \Sigma_d = D \Sigma_y D' + \underbrace{\eta_d E(\varepsilon_t \varepsilon_t') \eta_d'}_{=:\Sigma_{\varepsilon_d}} = D g_x \Sigma_x g_x' D' + \Sigma_{\varepsilon_d}. \qquad (10)
\end{aligned}$$

Using vectorization (or an algorithm for Lyapunov equations) we can solve (9)

$$vec(\Sigma_x) = (I_{n_x^2} - h_x \otimes h_x)^{-1} vec(\Sigma_{\varepsilon_x})$$

and are hence able to calculate the autocovariances for $t \in \mathbb{N} \setminus \{0\}$:

$$\begin{aligned}
\Sigma_x(t) &:= E[(x_t - \bar{x})(x_0 - \bar{x})'] = (h_x)^t \Sigma_x, \\
\Sigma_y(t) &:= E[(y_t - \bar{y})(y_0 - \bar{y})'] = g_x (h_x)^t \Sigma_x g_x', \\
\Sigma_d(t) &:= E[(d_t - \bar{d})(d_0 - \bar{d})'] = D g_x (h_x)^t \Sigma_x g_x' D'. \qquad (11)
\end{aligned}$$

Regarding the mean, notice that $E\left[(x_t - \bar{x})' \otimes (x_t - \bar{x})'\right] = vec(\Sigma_x)'$. Thus, taking expectations of (5), (6) and (4) yields

$$\mu_x := E(x_t) = \bar{x} + (I_{n_x} - h_x)^{-1} \frac{\Lambda_x}{2}, \qquad (12)$$

$$\mu_y := E(y_t) = \bar{y} + g_x(\mu_x - \bar{x}) + \frac{\Lambda_y}{2}, \qquad (13)$$

$$\mu_d := E(d_t) = D \mu_y. \qquad (14)$$

with

$$\Lambda_x = [I_{n_x} \otimes vec(\Sigma_x)'] \, vec(h'_{xx}) + \sigma^2 h_{\sigma\sigma}, \qquad \Lambda_y = \left[I_{n_y} \otimes vec(\Sigma_x)'\right] vec(g'_{xx}) + \sigma^2 g_{\sigma\sigma}.$$

In a linear approximation to the first order $\Lambda_x = \Lambda_y = 0$ and we have the expectation being equal to the steady-state of the model, $\mu_x = \bar{x}, \mu_y = \bar{y}$ and $\mu_d = D\bar{y}$, which is known as the *certainty-equivalence-property*. Going to a higher order approximation we are able to break with this principle.

2.4 Analytical derivatives of solution matrices and moments

We will derive the analytical derivatives of all solution matrices $(g_x, g_{xx}, g_{\sigma\sigma}, h_x, h_{xx}, h_{\sigma\sigma})$ as well as the moments $(\mu_x, \mu_y, \mu_d, \Sigma_x)$ with respect to the deep parameters θ. Following ideas from Iskrev (2008) and Schmitt-Grohé and Uribe (2012, Supplemental Material, Sec. A.3) we view f as well as the Jacobian of f as a function of θ and of the steady-state vector $\overline{z}(\theta) := (\overline{x}(\theta)', \overline{y}(\theta)')'$, which is also a function of θ. Thus, implicitly we have $f(\overline{z}(\theta), \theta) = 0$. Differentiating yields

$$df := \frac{\partial f(\overline{z}(\theta), \theta)}{\partial \theta'} = \frac{\partial f}{\partial \overline{z}'} \frac{\partial \overline{z}}{\partial \theta'} + \frac{\partial f}{\partial \theta'} = 0 \qquad \Leftrightarrow \qquad \frac{\partial \overline{z}}{\partial \theta'} = -\left[\frac{\partial f}{\partial \overline{z}'}\right]^{-1} \frac{\partial f}{\partial \theta'}.$$

This expression can easily be obtained analytically using MATLAB's symbolic toolbox. The derivative of the Jacobian $\mathcal{D}f(\overline{z}(\theta), \theta)$ with respect to θ is then given by

$$d\mathcal{D}f := \frac{\partial vec(\mathcal{D}f(\overline{z}(\theta), \theta))}{\partial \theta'} = \frac{\partial vec(\mathcal{D}f)}{\partial \overline{z}'} \frac{\partial \overline{z}}{\partial \theta'} + \frac{\partial vec(\mathcal{D}f)}{\partial \theta'}.$$

Note that $d\mathcal{D}f$ can be partitioned into

$$d\mathcal{D}f = \begin{pmatrix} \frac{\partial vec(\partial f(\overline{z})/\partial x'_{t+1})}{\partial \theta'} \\ \frac{\partial vec(\partial f(\overline{z})/\partial y'_{t+1})}{\partial \theta'} \\ \frac{\partial vec(\partial f(\overline{z})/\partial x'_{t})}{\partial \theta'} \\ \frac{\partial vec(\partial f(\overline{z})/\partial y'_{t})}{\partial \theta'} \end{pmatrix} =: \begin{pmatrix} df_1 \\ df_2 \\ df_3 \\ df_4 \end{pmatrix}.$$

This approach can be extended to calculate the analytical derivative of the Magnus-Neudecker-Hessian with respect to θ, since $H := \mathcal{H}f(\overline{z}(\theta), \theta)$:

$$dH := \frac{\partial vec(\mathcal{H}f(\overline{z}(\theta), \theta))}{\partial \theta'} = \frac{\partial vec(\mathcal{H}f)}{\partial \overline{z}'} \frac{\partial \overline{z}}{\partial \theta'} + \frac{\partial vec(\mathcal{H}f)}{\partial \theta'}.$$

Our MATLAB code writes all analytical derivatives using symbolic expressions into script files for further evaluation.

Derivatives of first-order solution matrices

Let $K_{n,q}$ be the commutation[7] matrix of order (n, q) and

$$F = -(h'_x g'_x \otimes I_{n_x+n_y}) df_2 - (h'_x \otimes I_{n_x+n_y}) df_1 - (g'_x \otimes I_{n_x+n_y}) df_4 - df_3,$$

[7] See Magnus and Neudecker (1999, p. 46) for the definition and Magnus and Neudecker (1999, p. 182) for an application regarding derivatives.

then the derivatives of the first-order solution matrices are given by:

$$\begin{aligned}
\begin{bmatrix} dg_x \\ dh_x \end{bmatrix} &= \left[(h_x' \otimes f_2) + (I_{n_x} \otimes f_4) \quad (I_{n_x} \otimes f_2 g_x) + (I_{n_x} \otimes f_1) \right]^{-1} \cdot F, \\
dg_x' &= K_{n_y, n_x} dg_x, \\
dh_x' &= K_{n_x, n_x} dh_x, \\
dh_x^t &= (I_{n_x} \otimes (h_x)^{t-1}) dh_x + (h_x' \otimes I_{n_x}) dh_x^{t-1}, \qquad t \geq 2.
\end{aligned}$$

See Schmitt-Grohé and Uribe (2012, Supplemental Material, Sec. A.3) for the derivation of these results. Note that $dX := \frac{\partial vec(X)}{\partial \theta'}$.

Derivatives of second-order solution matrices

Differentiating (7) with respect to θ requires the analytical derivatives of A and Q^{-1}, whereas differentiating (8) with respect to θ requires the analytical derivatives of B^{-1} and C. See appendix B for the derivation of these objects. Then the analytical derivatives of the second-order solution matrices with respect to θ can be summarized as

$$\begin{aligned}
d \begin{bmatrix} vec(g_{xx}) \\ vec(h_{xx}) \end{bmatrix} &= -Q^{-1} dA - (vec(A)' \otimes I_{n_x^2 (n_x + n_y)}) dQ^{-1} \\
d \begin{bmatrix} h_{ss} \\ g_{ss} \end{bmatrix} &= -(C' \otimes I_n) dB^{-1} - B^{-1} dC
\end{aligned}$$

Note that $dX := \frac{\partial vec(X)}{\partial \theta'}$.

Derivatives of moments

Differentiating (12), (13) and (14) with respect to θ requires the analytical derivatives of Λ_x and Λ_y, whereas differentiating (9) and (11) is straightforward using the vec-operator. See appendix B for the derivation of these objects. The analytical derivatives of the matrices determining the moments are then given by

$$\begin{aligned}
d\mu_x &= d\overline{x} + \left(\frac{\Lambda_x'}{2} \otimes I_{n_x} \right) \left([(I_{n_x} - h_x)']^{-1} \otimes (I_{n_x} - h_x)^{-1} \right) dh_x + (I_{n_x} - h_x)^{-1} \frac{d\Lambda_x}{2}, \\
d\mu_y &= d\overline{y} + \left[(\mu_x - \overline{x})' \otimes I_{n_y} \right] dg_x + (I_{n_x} \otimes g_x)(d\mu_x - d\overline{x}) + \frac{d\Lambda_y}{2}, \\
d\mu_d &= D d\mu_y, \\
d\Sigma_x &= \left[I_{n_x^2} - (h_x \otimes h_x) \right]^{-1} \left[(h_x \Sigma_x \otimes I_{n_x}) dh_x + (I_{n_x} \otimes h_x \Sigma_x) dh_x' + d\Sigma_{\varepsilon_x} \right], \\
d(\Sigma_d(t)) &= (D \otimes D) \left[(I_y \otimes g_x (h_x)^t \Sigma_x) dg_x' + (g_x \otimes g_x (h_x)^t) d\Sigma_x + (g_x \Sigma_x \otimes g_x) d(h_x^t) + (g_x \Sigma_x (h_x')^t \otimes I_y) dg_x \right].
\end{aligned}$$

where we used $d(X^{-1}) = (-(X')^{-1} \otimes X^{-1}) dX$, see Magnus and Neudecker (1999, p. 184).

Once again in a linear approximation to the first order this simplifies to certainty-equivalence, i.e. $d\mu_x = d\overline{x}, d\mu_y = d\overline{y}$ and $d\mu_d = D d\overline{y}$. Thus, only changes of the steady-state affect the mean up to first order. However, we will see that an approximation to the second order yields additional restrictions on the mean, which can tighten identification. In the next section it will be shown how to incorporate these additional restrictions into formal identifiability criteria and tests.

3 Identification criteria based on rank conditions

Suppose that data d_t is generated by the model with parameter vector θ_0. The criteria we will derive stem basically all from Theorem 4 in Rothenberg (1971), which essentially states identifiability conditions based on injectivity of functions. Formally, given an objective function $\mathfrak{f}(\theta)$ a sufficient condition for θ_0 being globally identified is given by

$$\mathfrak{f}(\theta_1) = \mathfrak{f}(\theta_0) \Rightarrow \theta_1 = \theta_0$$

for any $\theta_1 \in \Theta$. If this is only true for values θ in an open neighborhood of θ_0, the identification of θ_0 is local. Since most estimation methods in the econometric DSGE literature – e.g. full-information likelihood methods or limited-information methods like impulse-response matching or GMM – exploit information from the first two moments of data, we will focus on the mean, autocovariances and spectrum of the process. Since population moments are functions of data, the fundamental idea is to check, whether the mapping from θ to the population moments is unique. Then basic mathematical results for systems of equations can be applied. This set of criteria is the most basic and the closest to the ideas of the early work on identification in systems of linear equations, since it is based upon the uniqueness of a solution. Consequently, rank and order conditions are going to be derived, and it is also possible to pinpoint the parameters responsible of potential identification problems.

3.1 Iskrev (2010)'s approach

Suppose the DSGE-model is the data-generating process, then the vector of theoretical first and second moments of data d_t is accordingly given by equations (10), (11) and (14). Collect for $t = 0, 1, \ldots, T-1$ all elements in vectors

$$m(\theta, T) := \begin{pmatrix} vech(\Sigma_d)' & vec(\Sigma_d(1))' & \ldots & vec(\Sigma_d(T-1))' \end{pmatrix}',$$
$$\overline{m}(\theta, T) := \begin{pmatrix} \mu_d' & m(\theta, T)' \end{pmatrix}'.$$

$m(\theta, T)$ is a function of θ that determines the second moments of the data, whereas $\overline{m}(\theta, T)$ determines the first two moments of data. $\overline{m}(\theta, T)$ is of dimension $n_d + \frac{n_d(n_d+1)}{2} + (T-1)n_d^2$. If either ε_t is Gaussian or there are no distributional assumptions about the structural shocks, then the model-implied restrictions on $\overline{m}(\theta, T)$ contain all information that can be used for the estimation of θ. Iskrev (2010)'s test checks whether the derivatives of (the mean and) the predicted autocovariogram of the observables with respect to the vector of identifiable parameters has rank equal to the number of identifiable parameters. Formally, if $m(\theta, q)$ and $\overline{m}(\theta, q)$ are continuously differentiable functions of θ, then θ_0 is locally identifiable if the Jacobians

$$M(q) := \frac{\partial m(\theta_0, q)}{\partial \theta'} \qquad \text{and} \qquad \overline{M}(q) := \frac{\partial \overline{m}(\theta_0, q)}{\partial \theta'}$$

have full column rank at θ_0 for $q \leq T$, i.e. equal to n_θ.This gives immediate rise to a necessary condition: the number of identifiable parameters does not exceed the dimension of $m(\theta, T)$ or $\overline{m}(\theta, T)$. In fact, Iskrev (2010) establishes another order condition. For this, stack all elements of

the mean and the solution matrices that depend on θ into a vector τ:

$$\tau(\theta) = \tau_1(\theta) := \begin{pmatrix} \mu_d' & vec(h_x)' & vec(g_x)' & vech(\Sigma_{\varepsilon_x})' & vech(\Sigma_{\varepsilon_d})' \end{pmatrix}'$$

and consider the factorization

$$\overline{M}(q) = \frac{\partial \overline{m}(\theta, q)}{\partial \tau(\theta)'} \frac{\partial \tau(\theta)}{\partial \theta'}.$$

An immediate corollary implies that a point θ_0 is locally identifiable only if the rank of

$$J := \frac{\partial \tau(\theta_0)}{\partial \theta'}$$

at θ_0 is equal to n_θ. This condition is, however, only necessary, because τ may be unidentifiable.

Incorporating an approximation to the second order Implementing an approximation to the second order is straightforward, since in $\overline{m}_T$ only the expression for the mean μ_d changes. However, for the order condition we have to add the second order solution matrices to $\tau(\theta)$, thus we consider instead the vector

$$\tau(\theta) = \tau_2(\theta) := \begin{pmatrix} \tau_1(\theta)' & vec(h_{xx})' & vec(g_{xx})' & h_{\sigma\sigma}' & g_{\sigma\sigma}' \end{pmatrix}'.$$

Analytical and numerical derivatives For calculating the Jacobians analytically we need the derivatives of all solution matrices and moments, which we already derived in section 2.4. For calculating the Jacobians numerically we use the two-sided finite difference method described in appendix C.

Implementation and interpretation Check if J has full column rank for either τ_1 or τ_2 depending on the order of approximation. If the rank of J is equal to n_θ, then check the ranks of $M(q)$ and $\overline{M}(q)$. Start with the smallest q for which the order condition is satisfied, and then increase the number of moments if the rank condition fails. The reason is that if θ is identifiable, $M(q)$ and $\overline{M}(q)$ are likely to have full rank for q much smaller than T. If J is rank deficient at θ_0, then this point is not identifiable in the model. If $M(q)$ is rank deficient this means that θ_0 cannot be identified from the second moments only given the set of observed variables and the number of observations, whereas a rank deficient $\overline{M}(q)$ indicates nonidentification from the first two moments.

Pinpointing the source of identification is also possible, one has just to evaluate the rows of a rank deficient J, $M(q)$ or $\overline{M}(q)$. This will be a vector of zeros for any $q \leq T$, if θ_j does not affect the first and second moments. Further the columns that are linearly dependent indicate that these parameters are indistinguishable. We calculate the rank and analyze the null space of $J, M(q)$ as well as of $\overline{M}(q)$ using the singular value decomposition to pinpoint problematic parameters.

3.2 Komunjer and Ng (2011)'s approach

Komunjer and Ng (2011) derive two conditions for identification via equivalent spectral densities depending on the relation between the number of shocks and observables. We will focus on

singular and squared systems ($n_\varepsilon \leq n_d$) and assume fundamental innovations. Moreover, in the commonly used squared case ($n_\varepsilon = n_d$) both conditions coincide.[8]

Consider the linear solution of the minimal DSGE model, i.e. the model whose dynamics are entirely driven by the exogenous states $x_{2,t}$ and the structural innovations and measurement errors $e_t = \begin{pmatrix} \sigma\eta_u u_t \\ \eta_v v_t \end{pmatrix}$ with $E(e_t) = 0, E(e_t e_t') = \begin{pmatrix} \Sigma_u & 0 \\ 0 & \Sigma_v \end{pmatrix} =: \Sigma_e$:

$$x_{2,t} = \overline{x}_2 + \underbrace{h_{x_2}}_{\widetilde{\mathcal{A}}}(x_{2,t-1} - \overline{x}_2) + \underbrace{h_{e_2}}_{\widetilde{\mathcal{B}}} e_t,$$
$$d_t = \overline{d} + \underbrace{g_{d_2}}_{\widetilde{\mathcal{C}}}(x_{2,t-1} - \overline{x}_2) + \underbrace{g_{d_e}}_{\widetilde{\mathcal{D}}} e_t.$$

This minimal system has the smallest possible dimension n_{x_2} of the state vector that is able to capture all dynamics and has the familiar state-space solution. As DSGE models are based upon microfoundations n_{x_2} is for small and medium-sized DSGE models not hard to determine.[9]

Formal conditions for minimality require that for every $\theta \in \Theta$:

(i) Controllability: For any initial state, it is always possible to design an input sequence that puts the system in the desired final state, i.e. the matrix $\left[\widetilde{\mathcal{B}} \quad \widetilde{\mathcal{A}}\widetilde{\mathcal{B}} \quad \dots \quad \widetilde{\mathcal{A}}^{n_{x_2}-1}\widetilde{\mathcal{B}}\right]$ has full row rank,

(ii) Observability: Given the evolution of the input it is always possible to reconstruct the initial state by observing the evolution of the output, i.e. the matrix $\left[\widetilde{\mathcal{C}}' \quad \widetilde{\mathcal{A}}'\widetilde{\mathcal{C}}' \quad \dots \quad \widetilde{\mathcal{A}}^{n_{x_2}-1'}\mathcal{C}'\right]'$ has full column rank.

Since d_t is weakly stationary and e_t is either white noise or iid, the time-series properties of d_t are completely characterized by the time-invariant unconditional mean and autocovariances or, equivalently, there exist a $MA(\infty)$ representation

$$d_t = \bar{d} + \widetilde{\mathcal{D}}e_t + \sum_{j=1}^{\infty} \widetilde{\mathcal{C}}\widetilde{\mathcal{A}}^{j-1}\widetilde{\mathcal{B}}e_{t-j} = \bar{d} + \widetilde{H}_e(L^{-1};\theta)e_t,$$

where L is the lag-operator. For $z \in \mathbb{C}$ the transfer function (z-transform) is

$$\widetilde{H}_e(z;\theta) := \widetilde{\mathcal{D}} + \sum_{j=1}^{\infty} \widetilde{\mathcal{C}}\widetilde{\mathcal{A}}^{j-1}\widetilde{\mathcal{B}}z^{-j} = \widetilde{\mathcal{D}} + \widetilde{\mathcal{C}}\left[I_{n_{x_2}} - \widetilde{\mathcal{A}}z\right]^{-1}\widetilde{\mathcal{B}}.$$

Assuming left-invertibility of the matrix $\widetilde{H}_e$, d_t has the familiar Wold representation with e_t being fundamental. For square models $n_\varepsilon = n_d$, left-invertibility holds if $det(\widetilde{H}_e(z;\theta)) \neq 0$ in $|z| > 1$. When $n_\varepsilon \neq n_d$ left-invertibility requires that $\widetilde{H}_e(z;\theta)$ has full column rank in $|z| > 1$.[10]

Similar to Iskrev's approach we collect all hyperparameters of the state space solution into a vector $\Lambda(\theta) := \left(vec(\widetilde{\mathcal{A}})', vec(\widetilde{\mathcal{B}})', vec(\widetilde{\mathcal{C}})', vec(\widetilde{\mathcal{D}})', vech(\Sigma_e)'\right)'$ with dimension $n_\Lambda = n_{x_2}^2 + n_{x_2}n_\varepsilon +$

[8] For instance, in a model with anticipated shocks one has to use the innovation representation of the model and derive a slightly different rank condition.

[9] For the derivation of this model representation and some practical issues regarding the minimal state vector see Appendix D.

[10] This can be checked by computing the Rosenbrock system matrix $\mathbb{P}(z;\theta) = \begin{pmatrix} I_{n_{x_2}} - \widetilde{\mathcal{A}}z & \widetilde{\mathcal{B}} \\ -\widetilde{\mathcal{C}} & \widetilde{\mathcal{D}} \end{pmatrix}$. Then $rank(\mathbb{P}(z;\theta)) = n_{x_2} + rank(\widetilde{H}_e(z,\theta))$ for every complex z without the set of eigenvalues of $\widetilde{\mathcal{A}}$.

$n_d n_{x_2} + n_d n_\varepsilon + n_\varepsilon(n_\varepsilon+1)/2$. Further assume that $\Lambda : \theta \mapsto \Lambda(\theta)$ is continuously differentiable on Θ. For all $z \in \mathbb{C}$ the spectral density matrix of d_t is defined as

$$\Omega_d(z;\theta) := \Sigma_d + \sum_{j=1}^{\infty} \Sigma_d(j) z^{-j} + \sum_{j=1}^{\infty} \Sigma_d(-j) z^{-j} = \widetilde{H}_e(z;\theta)\Sigma_e(\theta)\widetilde{H}_e(z^{-1};\theta)'.$$

Observational equivalence of θ_0 and θ_1 from the spectral density is hence defined as

$$\forall z \in \mathbb{C} : \widetilde{H}_e(z;\Lambda(\theta_1)) \cdot \Sigma_e(\theta_1) \cdot \widetilde{H}_e(z^{-1};\Lambda(\theta_1))' = \widetilde{H}_e(z;\Lambda(\theta_0)) \cdot \Sigma_e(\theta_0) \cdot \widetilde{H}_e(z^{-1};\Lambda(\theta_0))'$$

implies $\theta_0 = \theta_1$. Equivalent spectral densities arise if

(i) for given $\Sigma_e(\theta)$, each transfer function $\widetilde{H}_e(z;\Lambda(\theta))$ is potentially obtained from a multitude of quadruples $(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}, \widetilde{\mathcal{C}}, \widetilde{\mathcal{D}})$,

(ii) there are many pairs $\widetilde{H}_e(z;\Lambda(\theta))$ and $\Sigma_e(\theta)$ that jointly generate the same spectral density.

Using results from control theory it can be shown that this is equivalent to the existence of a $n_{x_2} \times n_{x_2}$ similarity transformation matrix T and a $n_\varepsilon \times n_\varepsilon$ full column rank matrix $U = L_\varepsilon(\theta_0) V L_\varepsilon(\theta_1)^{-1}$ such that

$$\begin{aligned}
&\widetilde{\mathcal{A}}(\theta_1) = T\widetilde{\mathcal{A}}(\theta_0)T^{-1}, && \widetilde{\mathcal{B}}(\theta_1) = T\widetilde{\mathcal{B}}(\theta_0)U, && \widetilde{\mathcal{C}}(\theta_1) = \widetilde{\mathcal{C}}(\theta_0)T^{-1} \\
&\widetilde{\mathcal{D}}(\theta_1) = \widetilde{\mathcal{D}}(\theta_0)U, && \Sigma_e(\theta_1) = U^{-1}\Sigma_e(\theta_0)U^{-1'}, &&
\end{aligned}$$

with L_e being the Cholesky decomposition of $\Sigma_e(\theta) = L_e L_e'$ and V a constant matrix such that $VV' = I$.[11] Now define a continuously differentiable mapping $\delta : \Theta \times \mathbb{R}^{n_{x_2}^2} \times \mathbb{R}^{n_\varepsilon^2} \to \mathbb{R}^{n_\Lambda}$ as

$$\delta(\theta, T, U) := \begin{pmatrix} vec(T\widetilde{\mathcal{A}}(\theta)T^{-1}) \\ vec(T\widetilde{\mathcal{B}}(\theta)U) \\ vec(\widetilde{\mathcal{C}}(\theta)T^{-1}) \\ vec(\widetilde{\mathcal{D}}(\theta_0)U) \\ vech(U^{-1}\Sigma_e(\theta_0)U^{-1'}) \end{pmatrix}.$$

θ is now locally identifiable from the spectral density of d_t at a point θ_0 if and only if $\delta(\theta, T, U)$ is locally injective at $(\theta_0, I_{n_{x_2}}, I_{n_\varepsilon})$. A sufficient condition is thus, that the matrix of partial derivatives of $\delta(\theta, T, U)$ has full column rank at $(\theta_0, I_{n_{x_2}}, I_{n_\varepsilon})$. Denote this matrix by

$$\begin{aligned}
\Delta(\theta_0) &:= \begin{pmatrix} \frac{\partial\delta(\theta_0, I_{n_{x_2}}, I_{n_\varepsilon})}{\partial\theta'} & \frac{\partial\delta(\theta_0, I_{n_{x_2}}, I_{n_\varepsilon})}{\partial vec(T)'} & \frac{\partial\delta(\theta_0, I_{n_{x_2}}, I_{n_\varepsilon})}{\partial vec(U)'} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial vec(\widetilde{\mathcal{A}})}{\partial\theta'} & \widetilde{\mathcal{A}}' \otimes I_{n_{x_2}} - I_{n_{x_2}} \otimes \widetilde{\mathcal{A}} & 0_{n_{x_2}^2 \times n_\varepsilon^2} \\ \frac{\partial vec(\widetilde{B})}{\partial\theta'} & \widetilde{\mathcal{B}}' \otimes I_{n_{x_2}} & I_{n_\varepsilon} \otimes \widetilde{\mathcal{B}} \\ \frac{\partial vec(\widetilde{\mathcal{C}})}{\partial\theta'} & -I_{n_{x_2}} \otimes \widetilde{\mathcal{C}} & 0_{n_d n_{x_2} \times n_\varepsilon^2} \\ \frac{\partial vec(\widetilde{\mathcal{D}})}{\partial\theta'} & 0_{n_d n_\varepsilon \times n_{x_2}^2} & I_{n_\varepsilon} \otimes \widetilde{\mathcal{D}} \\ \frac{\partial vech(\Sigma_e)}{\partial\theta'} & 0_{(n_\varepsilon(n_\varepsilon+1)/2) \times n_{x_2}^2} & -2\Xi_{n_\varepsilon}[\Sigma_e \otimes I_{n_\varepsilon}] \end{pmatrix} \\
&=: \begin{pmatrix} \Delta_\Lambda(\theta_0) & \Delta_T(\theta_0) & \Delta_U(\theta_0) \end{pmatrix}
\end{aligned}$$

[11] The key insight behind this result is that the spectral density can be factorized and due to left invertibility the matrix V is not a polynomial matrix of unknown degree, but a constant matrix.

with Ξ_{n_ε} being the left-inverse of the $n_\varepsilon^2 + n_\varepsilon(n_\varepsilon+1)/2$ duplication matrix $\mathcal{G}_{n_\varepsilon}$ for $vech(\Sigma_e)$.[12] There are 3 blocks to consider:

1. The $(n_\Delta \times n_\theta)$ block defined by $\Delta_\Lambda(\theta_0)$: The rank of $\Delta_\Lambda(\theta_0)$ must equal n_θ for the mapping $\Lambda(\theta)$ being locally invertible at θ.
2. The $(n_\Delta \times n_{x_2}^2)$ block defined by $\Delta_T(\theta_0)$: The rank of $\Delta_T(\theta_0)$ must equal $n_{x_2}^2$ so that the identity matrix is the only local similarity transformation.
3. The $(n_\Delta \times n_\varepsilon^2)$ block defined by $\Delta_U(\theta_0)$: The rank of $\Delta_U(\theta_0)$ must equal n_ε^2 so that the spectral factorization is locally uniquely determined.

with $n_\Delta := (n_{x_2}+n_d)(n_{x_2}+n_\varepsilon)+n_\varepsilon(n_\varepsilon+1)/2$. This yields the following order and rank condition:

- Order condition (necessary): $n_\theta + n_{x_2}^2 + n_\varepsilon^2 \leq n_\Delta$
- Rank condition (necessary and sufficient): $rank(\Delta(\theta_0)) = n_\theta + n_{x_2}^2 + n_\varepsilon^2$

Incorporating an approximation to the second order Within this framework it is straightforward to check identification via mean restrictions either from the first or second order approximation. The restrictions in equation (14) simply augment the rows of the objective function δ, that is

$$\overline{\delta}(\theta,T,U) := \begin{pmatrix} \mu_d(\theta) \\ \delta(\theta,T,U) \end{pmatrix}, \qquad \overline{\Delta}(\theta_0) := \begin{pmatrix} \frac{\partial \mu_d(\theta_0)}{\partial \theta'} & 0_{n_d \times n_{x_2^2}} & 0_{n_d \times n_\varepsilon^2} \\ \Delta_\Lambda(\theta_0) & \Delta_T(\theta_0) & \Delta_U(\theta_0) \end{pmatrix}.$$

Notice that the rank condition doesn't change, since the restrictions are incorporated as additional equations for solving the same number of unknowns. The order condition, however, requires $n_\theta + n_{x_2}^2 + n_\varepsilon^2 \leq n_d + n_\Delta$.

Analytical and numerical derivatives In section 2.4 we already derived the derivatives of h_x and g_x. The derivatives of $\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}, \widetilde{\mathcal{C}}$ and $\widetilde{\mathcal{D}}$ are easily obtained by choosing the rows of dh_x and dg_x corresponding to the minimal state vector and the auxiliary shock variables. $d\Sigma_e$ consists of $d\Sigma_u$ and $d\Sigma_v$, for which we have closed-form expressions. For calculating the Jacobians numerically we use the two-sided finite difference method described in appendix C to obtain the Jacobians of all minimal solution matrices, the covariance matrix and the mean vector.

Implementation and interpretation Having the derivatives $d\widetilde{\mathcal{A}}, d\widetilde{\mathcal{B}}, d\widetilde{\mathcal{C}}, d\widetilde{\mathcal{D}}, d\Sigma_e$ and $d\mu_d$, we are able to set up Δ as well as $\overline{\Delta}$ and check the order and rank condition. The order condition (dimension of $\overline{\Delta}(\theta_0)$) is a simple function of the number of variables that appear in the DSGE model. The rank, however, is a function of the parameters in the linearized solution to the DSGE model. A rank-deficient $\overline{\Delta}$ means that some parameters are observational equivalent from the spectral density and mean restrictions (depending on the order of approximation), whereas a rank-deficient Δ indicates nonidentification from the spectral density alone. There are three necessary conditions for observational equivalent spectral densities: (1) Similar to Iskrev's J, if the solution matrices are sensitive to changes in parameters, we get a full rank Δ_Λ, (2) full rank of Δ_T means there exist only one quadruple generating the z-Transform for the spectral density, while (3) full rank of Δ_U indicates that there exist a unique pair of z-Transform and dynamic

[12] See Magnus and Neudecker (1999, p. 49) for the definition of the duplication matrix.

structure of the stochastic innovations that generate the spectral density. However, note that individual full column ranks in all three cases are necessary, but not sufficient for identification. Thus we have to check the ranks of Δ and $\overline{\Delta}$ as well.

We calculate the rank and analyze the null space of $\overline{\Delta}(\theta_0), \Delta(\theta_0), \Delta_\Lambda(\theta_0), \Delta_T(\theta_0), \Delta_U(\theta_0)$ using the singular value decomposition for different tolerance levels. Thus we are able to find parameter dependencies and interactions, that are responsible for nonidentification.

3.3 Tkachenko and Qu (2012)'s approach

Assume that d_t is covariance-stationary, thus it has a vector-moving-average representation

$$d_t = \bar{d} + \sum_{j=0}^{\infty} Dg_x h_x^j \sigma \eta_x \varepsilon_{t-j} + \eta_d \varepsilon_t = \bar{d} + H_\varepsilon(L;\theta)\varepsilon_t$$

with $H_\varepsilon(L;\theta) = Dg_x \left(I_{n_x} - h_x L\right)^{-1} \sigma\eta_x + \eta_d$. Using the Fourier transformation for the lag-operator the spectral density matrix Ω_d is given by

$$\Omega_d(\omega,\theta) = \frac{1}{2\pi} H_\varepsilon(e^{-i\omega};\theta) \cdot \underbrace{E(\varepsilon_t \varepsilon_t')}_{=I_{n_\varepsilon}} \cdot H_\varepsilon(e^{-i\omega};\theta)^*, \qquad \omega \in [-\pi;\pi],$$

with * denoting the conjugate transpose of a complex valued matrix. The spectral density matrix has n_d^2 entries and each one is a map from Θ to complex valued functions defined over $[-\pi;\pi]$ in a Banach space.[13] Similar to Rothenberg (1971), who looks at the Hessian of the parametric density function in the Gaussian case, the authors focus on

$$G(\theta_0) = \int_{-\pi}^{\pi} \left(\frac{\partial vec(\Omega_d(\omega;\theta_0)')}{\partial \theta'}\right)' \left(\frac{\partial vec(\Omega_d(\omega;\theta_0))}{\partial \theta'}\right) d\omega$$

and assume that the spectral density ist continuous in ω and continuous and differentiable in θ. The dimension of $G(\theta_0)$ is always $n_\theta \times n_\theta$. Let $\theta_0 \in \Theta$ be a regular point, then it can be shown that θ is locally identifiable at a point θ_0 from the spectrum of d_t if and only if $G(\theta_0)$ is nonsingular, i.e. its rank is equal to n_θ.

Incorporating an approximation to the second order This framework is very flexible and allows for incorporating conditions for identification from a subset of frequencies, partial and conditional identification as well as general constraints: The procedure is to simply add additional restrictions to the G matrix. We include the mean restriction (14) conditional on the order of approximation and consider

$$\bar{G}(\theta_0) = G(\theta_0) + \frac{\partial \mu_d(\theta_0)'}{\partial \theta} \frac{\partial \mu_d(\theta_0)}{\partial \theta'}.$$

θ is locally identifiable at a point θ_0 from the mean and the spectrum of d_t if and only if the rank of $\overline{G}(\theta_0)$ is equal to n_θ.

Analytical and numerical derivatives We will now show how to obtain the derivative of $\Omega_d(\omega;\theta_0)$ w.r.t. θ analytically:

[13] If the spectral density matrix is continuous there is a one-to-one relationship to the autocovariogram $\Sigma_d(j) = \int_{-\pi}^{\pi} e^{ik\omega} \Omega_d(\omega,\theta) d\omega, j = 0, \pm 1, \ldots$.

1. Divide the interval $[-\pi;\pi]$ into N subintervals to obtain $N+1$ frequency indices, ω_s denotes the $s-$th frequency in the partition.
2. Solve the DSGE model using $\theta = \theta_0$.
3. For $s = 1, \dots N+1$:
   (i) Derive the derivative of $H_\varepsilon(e^{-i\omega_s};\theta_0)$ and its conjugate transpose using the expression in appendix B for each ω_s.
   (ii) Calculate for each ω_s

$$d\Omega_d(\omega_s,\theta_0) = \frac{1}{2\pi}\left[(H_\varepsilon^{*'} \otimes I_{n_d})dH_\varepsilon + (I_{n_\varepsilon} \otimes H_\varepsilon)dH_\varepsilon^*\right].$$

4. Approximate $G(\theta_0)$ using

$$G(\theta_0) \approx \frac{2\pi}{N+1}\sum_{s=1}^{N+1} d\Omega_d(\omega_s,\theta_0)' d\Omega_d(\omega_s,\theta_0).$$

Regarding numerical derivatives we use the two-sided central difference method described in appendix C to compute for each ω_s the non-vectorized derivative $\frac{\partial\Omega_d(\omega_s;\theta_0)}{\partial\theta_j}$ and stack these into a big matrix. The typical element of G is then given by

$$G_{jk}(\theta) = \int_{-\pi}^{\pi} tr\left\{\frac{\partial\Omega_d(\omega;\theta)}{\partial\theta_j}\frac{\partial\Omega_d(\omega;\theta)}{\partial\theta_k}\right\} d\omega$$

which can be approximated by

$$G_{jk}(\theta_0) \approx \frac{2\pi}{N+1}\sum_{s=1}^{N+1} tr\left\{\frac{\partial\Omega_d(\omega_s;\theta_0)}{\partial\theta_j}\frac{\partial\Omega_d(\omega_s;\theta_0)}{\partial\theta_k}\right\}, \qquad j,k = 1,\dots,n_\theta$$

Note, that even in the analytical case we still have to divide the interval $[-\pi;\pi]$ into sufficient subintervals N to numerically approximate the integral.

Implementation and interpretation Since G as well as $\overline{G}$ are real, symmetric and positive semidefinite, the rank can be easily computed using the singular value decomposition and counting non-zero singular values. A full rank G indicates identification via the spectrum, whereas a full rank $\overline{G}$ also adds the mean restrictions to the spectrum. If either is rank deficient, we check identification for all possible subsets of the parameter vector separately. That is, we start with one-element subsets, i.e. we check identification for each parameter separately. Then we continue with all possible two-element subsets and check the rank for each subset individually. We do this for all j-subsets, $j = 1,\dots,n_\theta - 1$. This procedure is computationally very easy, since for the one-element subsets we only have to check whether the diagonal elements of G and $\overline{G}$ are nonzero, whereas for the higher subsets we can pick the relevant elements of the G and $\overline{G}$ matrix respectively, and calculate its rank using the singular value decomposition. If for any j-subset the rank is deficient, we conclude that the parameters in the subset are not identifiable and exclude them from the higher ordered subsets. Thus, we are able to pinpoint the source of nonidentification. A further benefit of considering the frequency domain is the possibility of constructing nonidentification curves that can show how large the neighborhood of the nonidentified parameter values is. This can be insightful for finding regions of nonidentification.

3.4 Discussion of methods based on rank conditions

All presented methods exploit the dynamic structure of the solution of a DSGE model to define mappings and establish conditions for local injectivity of the mappings. For all procedures we are able to derive necessary as well as sufficient conditions for identification, and to pinpoint the source of nonidentification, i.e. problematic parameters. Including the mean restrictions from a second order approximation is also straightforward.

The point of departure, however, is different: Iskrev's approach is based in the time domain, whereas Qu and Tkachenko derive conditions in the frequency domain. It is known that given some regularity conditions there is a one-to-one mapping between the time and frequency domain, thus the criteria should yield the same results. Komunjer and Ng's approach can be considered to be in between both approaches, since they establish conditions without actually computing autocovariances or the spectral density. However, their assumptions are the strictest as they rely on minimality and left-invertibility. Deriving and checking the minimality of the model can be tedious in practice.[14] Iskrev's and Qu and Tkachenko's method is very general as we only assume the existence of the VARMA representation of the DSGE model. However, we have to actually compute the derivative of the autocorrelogram and of the spectrum, which may leave scope for numerical errors and imprecision. In particular choosing the lag order T as well as number of subintervals N for the frequencies may change results, since strictly speaking the criteria are only valid for $T, N \to \infty$. In practice, however, this is not a question of heavily sensitive results[15], but rather one of speed: the higher T or N, the more time the calculations need. Komunjer and Ng's approach is hence the fastest, since we only have to evaluate the solution matrices and their derivatives (which we also have to do for the other criteria). In this line of thought note that all methods depend heavily on the solution matrices and suffer from possible numerical instability of the solution algorithm. However, since we used the same framework and algorithm across methods, we are able to neglect this effect in the comparative approach in section 5.

The different interpretations of Iskrev's and Komunjer and Ng's criteria can also be used as diagnostics for model building. For instance J as well as Δ_Λ check the mapping from the structural parameters to the solution parameters. The evaluation yields parameters that do not influence the reduced-form solution and may be thus obsolete. A researcher is hence able to reparameterize the model prior to estimation. Moreover, given a known shock a rank deficient $\Delta_{\Lambda T}$ indicates that two structures (e.g. two different policies) might cause the same impulse response of the model, so we have to be careful interpreting the importance of shocks. In contrast given a rank deficient $\Delta_{\Lambda U}$ we cannot be sure, whether it is the size of the shock or similar propagating mechanisms, that yield the same dynamic structure of the model. Qu and Tkachenko's test does not give such diagnostics, however, their approach can be used for a quasi maximum likelihood estimation in the frequency domain. Moreover, it is possible to get insight into the size of the local neighborhood of the unidentified parameters.

Lastly, all procedures check only local identification.[16] Thus, one has to make sure that this procedure is valid for a sufficient range of parameters. However, given prior beliefs and upper and

[14] In a model with non-fundamental innovations ($n_d < n_\varepsilon$) we also have to work with the innovation representation of the model.

[15] In most practical cases T in between 10 and 100 will be sufficient, since the higher the lag the less informative the identification restrictions. Further we experienced with different values for N and find that the results hardly change. Thus an N in the order of 10000 is sufficient as well.

[16] See Komunjer (2012) for issues regarding global identification. She uses results from unconditional moment restriction models (properness and homeomorphism) to establish identification conditions in the fashion of GMM identification conditions.

lower values this can be granted. A possible identification analysis would be to generate many random draws θ_i from Θ that are theoretically plausible (lower or upper bounds, prior distributions, conditions for existence and uniqueness should be satisfied) and check the aforementioned criteria.

4 Identification criteria based on Bayesian methods (NOT YET IN THE CODE, discussion incomplete)

4.1 Bayesian comparison indicator

It is well known that informative marginal priors are sufficient to get well-defined posteriors even for non-identifiable parameters. The usual approach in the DSGE literature is to compare the prior and posterior distribution of a parameter. If they differ, there is "learning", i.e. data seems to be informative about a parameter. However, from an identification point of view, this can be misleading as was shown by Kadane (1974) and Poirier (1998).

Suppose the deep parameters can be divided into two subsets $\theta = (\theta_1', \theta_2')'$ with $\theta_1 \in \Theta_1$ being non-identifiable and $\theta_2 \in \Theta_2$ identifiable. The permissible values for the parameters are each possibly conditionally dependent: $\Theta_1(\theta_2) := \{\theta_1 | (\theta_1', \theta_2')' \in \Theta\}$ and $\Theta_2(\theta_1) := \{\theta_2 | (\theta_1', \theta_2')' \in \Theta\}$. Given Bayes' rule the marginal posterior density for the unidentified $\theta_1 \in \Theta_1$ is given by

$$p(\theta_1|d) = \frac{p(\theta_1) f(d|\theta_1)}{p(d)}$$

If data $d = d_0, \ldots, d_T$ is marginally uninformative for θ_1, this implies that $p(\theta_1|d) = p(\theta_1)$ or equivalently $f(d|\theta_1) = p(d)$. Since the likelihood $f(d|\theta_1, \theta_2) = f(d|\theta_2)$ does not depend on θ_1 we have

$$\begin{aligned} f(d|\theta_1) &= \int_{\Theta_2(\theta_1)} f(d|\theta_1, \theta_2) p(\theta_2|\theta_1) d\theta_2 = \int_{\Theta_2(\theta_1)} f(d|\theta_2) p(\theta_2|\theta_1) d\theta_2 \\ p(d) &= \int_{\Theta_2(\theta_1)} f(d|\theta_2) p(\theta_2) d\theta_2 = \int_{\Theta_1} \int_{\Theta_2(\theta_1)} f(d|\theta_2) p(\theta_2, \theta_1) d\theta_1 d\theta_2. \end{aligned}$$

For this to be equal, two conditions have to be met: First, the priors have to be independent, i.e. $p(\theta_2, \theta_1) = p(\theta_2)p(\theta_1)$ and secondly, the parameter space has to be a product space, i.e. $\Theta = \Theta_1 \times \Theta_2$.[17] Otherwise there is dependence of $p(\theta_1|d)$ on θ_1 through either the conditional density $p(\theta_2|\theta_1)$ or the support $\Theta_1(\theta_2)$. In other words, if these conditions are not satisfied, then "data based learning about θ_2 can 'spill over' onto the unidentified θ_1" (Koop, Pesaran, and Smith 2012, p. 14). Thus, one has to be careful judging identification from "apparent learning", i.e. the posterior $p(\theta_1|d)$ being different than the prior $p(\theta_1)$.

[17] Poirier (1998) defines this as θ_1 and θ_2 being variation free.

However, if data d is conditionally uninformative for θ_1, i.e. $p(\theta_1|\theta_2, d) = p(\theta_1|\theta_2)$ we get

$$p(\theta_1|\theta_2,d) = \frac{p(\theta_1,\theta_2|d)}{p(\theta_2|d)} = \frac{p(\theta_1|\theta_2,d)p(\theta_2|d)}{p(\theta_2|d)} = \frac{p(\theta_1|\theta_2,d)\left(p(\theta_2)f(d|\theta_2)\right)/p(d)}{\left(\int\limits_{\Theta_1(\theta_2)} p(\theta_2)f(d|\theta_2,\theta_1)p(\theta_1|\theta_2)d\theta_1\right)/p(d)}$$
$$= \frac{p(\theta_2)p(\theta_1|\theta_2)f(d|\theta_2)}{p(\theta_2)f(d|\theta_2)} = p(\theta_1|\theta_2)$$

for some $\theta_2 \in \Theta_2$ and $\theta_1 \in \Theta_1(\theta_2)$. Now, when looking at conditional informativeness, the aforementioned dependencies do not occur, since neither the density nor the support are required to be independent. This is the idea of proposition 2 in Poirier (1998) which states, that data is always conditionally uninformative for a subset of nonidentified parameters given a subset of identified parameters. Thus the posterior for $\theta_1 \in \Theta_1$ should be rewritten with respect to the conditional prior beliefs given the identified $\theta_2 \in \Theta_2(\theta_1)$:

$$\begin{aligned} p(\theta_1|d) &= \int\limits_{\Theta_2(\theta_1)} p(\theta_1,\theta_2|d)d\theta_2 = \int\limits_{\Theta_2(\theta_1)} p(\theta_1|\theta_2,d)p(\theta_2|d)d\theta_2 \\ &= \int\limits_{\Theta_2(\theta_1)} p(\theta_1|\theta_2)p(\theta_2|d)d\theta_2 \\ &= E_{\theta_2|y}\left[p(\theta_1|\theta_2)\right]. \end{aligned}$$

Following Koop, Pesaran, and Smith (2012) we apply this idea to DSGE models and compare the properties of $p(\theta_1|d)$ and $E_{\theta_2|y}\left[p(\theta_1|\theta_2)\right]$ in addition to $p(\theta_1|d)$ and $p(\theta_1)$.

4.2 Bayesian learning rate indicator

Naturally identification should become better as more data becomes available. With an infinite sample the role of the prior vanishes and Bayesian asymptotics are identical to the asymptotic distribution theory for maximum likelihood. However, this requires θ to be identified. Thus, if the convergence does not occur, θ is not or only weakly identified. Koop, Pesaran, and Smith (2012) use this idea to derive an indicator, that is focused on the rate at which learning, interpreted as increasing posterior precision, occurs.

Again, stack all non-constant elements of the solution matrices that depend on θ into a $k \times 1$ vector-valued function $\tau(\theta)$ and denote the Jacobian with $J(\theta) = \partial\tau(\theta)/\partial\theta'$.[18] The log-likelihood function $l_T(\tau) := l(\tau, d)$ is only dependent on the solution parameters and data of size T. Further assume that the average Hessian of the log-likelihood

$$Q_T(\tau) = \frac{-1}{T}\frac{\partial^2 l_T(\tau)}{\partial\tau\partial\tau'}$$

is a positive definite matrix for all admissible values of τ. Standard results from maximum likelihood asymptotics yield that the maximum likelihood estimator $\widehat{\tau}_T$ converges to the true value $\tau_0 = \tau(\theta_0)$ such that

$$\sqrt{T}(\widehat{\tau}_T - \tau_0) \xrightarrow{d} N(0, Q^{-1})$$

[18] In a linear approximation to the first order $\tau = \tau_1$ and to the second order $\tau = \tau_2$, see section 3.1.

with $Q = \underset{T\to\infty}{plim}\ Q_T(\tau_0)$. Given a multivariate normal density with mean vector θ^{pr} and precision H^{pr} as the prior distribution

$$p(\theta) = (2\pi)^{-n_\theta/2}|H^{pr}|^{1/2}exp\{-\frac{1}{2}(\theta-\theta^{pr})'H^{pr}(\theta-\theta^{pr})\},$$

it can be shown that the posterior distribution of θ is approximately normal with mean θ_T^{po} and precision matrix H_T^{po} such that

$$\begin{aligned}
\theta_T^{po} &= (TS_T + H^{pr})^{-1}(TS_T\widehat{\theta}_T + H^{pr}\theta^{pr}) \\
H_T^{po} &= \begin{pmatrix} H_{11,T}^{po} & H_{12,T}^{po} \\ H_{21,T}^{po} & H_{22,T}^{po} \end{pmatrix} = T\begin{pmatrix} S_{11,T} & S_{12,T} \\ S_{21,T} & S_{22,T} \end{pmatrix} + \begin{pmatrix} H_{11}^{pr} & H_{12}^{pr} \\ H_{21}^{pr} & H_{22}^{pr} \end{pmatrix} = TS_T + H^{pr} \\
S_T &= \begin{pmatrix} S_{11,T} & S_{12,T} \\ S_{21,T} & S_{22,T} \end{pmatrix} = J(\theta)'Q_T(\tau)J(\theta),
\end{aligned}$$

assuming for simplicity $k = n_\theta$ and partitioning θ into a non-identified part θ_1 and an identified part θ_2. Using the fact that the marginal distributions of the multivariate normal distribution are also normal it can be shown, that the posterior precision for the unidentified θ_1 is given by

$$H_{11,T}^{po} = (TS_{11,T} + H_{11}^{pr}) - (TS_{12,T} + H_{12}^{pr})(TS_{22,T} + H_{22}^{pr})^{-1}(TS_{21,T} + H_{21}^{pr})$$

Now two cases can be analyzed: the identified and unidentified case.

Identified case In the identified case J has full rank and thus S_T is bounded in T and a full rank matrix for all T. It can be shown that apart from terms of order T^{-2} this implies

$$\begin{aligned}
T^{-1}H_{11,T}^{po} = &(S_{11,T} - S_{12,T}S_{22,T}^{-1}S_{21,T}) \\
&+ T^{-1}(H_{11}^{pr} - H_{12}^{pr}S_{22,T}^{-1}S_{21,T} - S_{12,T}S_{22,T}^{-1}H_{21}^{pr} + S_{12,T}S_{22,T}^{-1}H_{22}^{pr}S_{22,T}^{-1}S_{21,T}) \\
&+ O(T^{-2})
\end{aligned}$$

As can be seen, the dependence of the average posterior precision on the priors diminishes at rate T^{-1}. Taking the limit yields

$$\underset{T\to\infty}{lim}\,(T^{-1}H_{11,T}^{po}) = \underset{T\to\infty}{lim}\,(S_{11,T} - S_{12,T}S_{22,T}^{-1}S_{21,T}) > 0,$$

since S_T is bounded in T and of full rank. Thus, if $H_{11,T}^{po}$ improves at rate T, $T^{-1}H_{11,T}^{po}$ converges to a constant number in the identified case.

Unidentified or weakly identified case Weak identification is defined as in Stock, Wright, and Yogo (2002), as the derivatives of the solution parameters with respect to the weak identified parameters can for finite T be different than zero, but asymptotically not: $J_{11,T} = T^{-1/2}\Delta_{11,T}$ and $J_{12,T} = T^{-1/2}\Delta_{12,T}$. These matrices are bound in T and the unidentified case is simply given by $\Delta_{11,T} = \Delta_{12,T} = 0$. Further since θ_2 is identified, $J_{22,T}$ has full rank for all T:

$$J_T = \begin{pmatrix} J_{11,T} & J_{12,T} \\ J_{21,T} & J_{22,T} \end{pmatrix} = \begin{pmatrix} T^{-1/2}\Delta_{11,T} & T^{-1/2}\Delta_{12,T} \\ J_{21,T} & J_{22,T} \end{pmatrix}$$

Looking at the Cholesky decomposition of $Q_T = U_T' U_T$, with U_T being lower triangular, the following auxiliary matrix is used:

$$\begin{aligned} P_T &= \begin{pmatrix} P_{11,T} & P_{12,T} \\ P_{21,T} & P22,T \end{pmatrix} = \begin{pmatrix} U_{11,T} & 0 \\ U_{21,T} & U22,T \end{pmatrix} \begin{pmatrix} J_{11,T} & J_{22,T} \\ J_{21,T} & J22,T \end{pmatrix} \\ &= \begin{pmatrix} U_{11,T}J_{11,T} & U_{11,T}J_{12,T} \\ U_{21,T}J_{11,T} + U_{22,T}J_{21,T} & U_{21,T}J_{12,T} + U_{22,T}J_{22,T} \end{pmatrix} \\ &= U_T J_T \end{aligned}$$

Under weak identification, this yields

$$\begin{aligned} P_{11,T} &= T^{-1/2} U_{11,T} \Delta_{11,T} \\ P_{12,T} &= T^{-1/2} U_{11,T} \Delta_{12,T} \\ P_{21,T} &= T^{-1/2} U_{21,T} \Delta_{11,T} + U_{22,T} J_{21,T} \\ P_{22,T} &= T^{-1/2} U_{21,T} \Delta_{12,T} + U_{22,T} J_{22,T}. \end{aligned}$$

Now it can be shown that the posterior precision of the unidentified parameters is apart from terms of order T^{-1} given by

$$\begin{aligned} H_{11,T^{po}} = (U_{11,T}\Delta_{11,T} - U_{12,T}\Delta_{12,T}P_{22,T}^{-1}P_{21,T})'(U_{11,T}\Delta_{11,T} - U_{12,T}\Delta_{12,T}P_{22,T}^{-1}P_{21,T}) \\ + H_{11}^{pr} - P_{21,T}'P_{22,T}^{'-1} - H_{12}^{pr}P_{22,T}^{-1}P_{21,T} + P_{21,T}'P_{22,T}^{'-1}H_{22}^{pr}P_{22,T}^{-1}P_{21,T} + O(T^{-1}) \end{aligned}$$

In contrast to the identified case the posterior precision does not rise with T and is bounded by the priors and data given the P, U and Δ matrices. Consequently, the average precision $T^{-1}H_{11,T}^{po}$ tends to zero in the limit.

Bayesian learning rate indicator Given these two results one can use the average precision as an indicator for weak identification given a subset of identified parameters:[19]

$$\lim_{T\to\infty} T^{-1}\overline{H}_{11,T} \begin{cases} 0 & (H_{11,T}^{po} \text{ improves at rate slower than T}) \quad \text{if } \theta_1 \text{ unidentified} \\ \text{a number} & (H_{11,T}^{po} \text{ improves at rate T}) \quad \text{if } \theta_1 \text{ identified} \end{cases}$$

Discussion of Bayesian methods (incomplete)

Both indicators are very easy to implement, since it requires only a few additional steps during an ordinary Bayesian estimation, i.e. the ability to calculate the posterior expectation of the conditional prior and to simulate data of increasing sizes. The second indicator is also useful for detecting weak identification. Caglar, Chadha, and Shibayama (2012) already show how to use the Bayesian learning rate indicator for the workhorse medium-sized DSGE model of Smets and Wouters (2007). However, these procedures require a subset of parameters that are definitely identifiable. In practice this means first finding this subset by using one or several of the aforementioned criteria.

A minor drawback regarding the Bayesian learning rate indicator is also noteworthy: assuming a Gaussian prior is chosen for analytical simplicity and in order to get a closed-form solution.

[19] See Caglar, Chadha, and Shibayama (2012, p. 6).

However, this assumption can be relaxed with the cost of not getting an analytical expression for the posterior precision. Since the focus lies on the rate at which the precision gets updated with the sample size, this is only a minor shortcoming.[20]

Whereas, incorporating an approximation to the second order is very straightforward, as the procedure relies only on the ability to simulate the posterior and data. For higher order approximations there are several things to be careful about. First computing the posterior requires the particle filter or efficient importance sampling. Second, when simulating data one has to use pruning in the higher order approximation.[21] These approaches thus rely heavily on the numerical techniques, algorithms and MCMC methods. Different approaches may lead to different results.

5 Application (Bayesian indicators not yet)

We will now compare the different methods on three illustrative models: (i) the neoclassical growth model, (ii) the Kim (2003) model and (iii) the An and Schorfheide (2007) model. The first is a well-identified model, whereas the other two models are known to have identification failures. Please refer to appendix E for the details about each model and how to represent them in the proposed framework of chapter 2. We welcome the interested reader to use our user-friendly Matlab code to try different subsets of parameters, local points, analytical or numerical derivatives, order of approximations and other options to confirm results of this section. In the code we also implemented the models of Christiano, Eichenbaum, and Evans (2005), Garcia-Cicco, Pancrazi, and Uribe (2010) and Smets and Wouters (2007).

5.1 The neoclassical growth model

The neoclassical growth model is a simple and well identified model, which we will only use as a benchmark case. Table 2 summarizes the results of the tests for different tolerance levels. As expected, all tests yield that the full parameter vector is identifiable at the given local point and given the observables across tolerance levels.[22] The second order approximation as well as using analytical or numerical derivatives does not change the ranks of all tests.

However, when we change the size of the standard deviation of the shock on technology from $\sigma_a = 1$ to $\sigma_a = 0.01$ we get a different picture, see table 3 for the corresponding ranks and results. The problematic parameter for Iskrev's J at the largest tolerance level is σ_a, whereas for the second moments the parameters $(\alpha, \rho, \beta, \delta, \gamma, \sigma_a)$ are together not distinguishable (for $1e-5$). Tkachenko and Qu's test results in nonidentifiable parameters for tolerance levels down to $1e-09$. However, the source of nonidentification is not as straightforward as in the other tests. Pinpointing the problematic parameters for each tolerance level yields several combinations of unidentifiable subsets. For instance at tolerance level $1e-6$ we have three subsets that yield observational equivalent spectral densities: $\gamma, (\alpha, \rho, \beta, \delta)$ and $(\alpha, \beta, \delta, \sigma_a)$. Komunjer & Ng's $\Delta_\Lambda, \Delta_{\Lambda T}$ and $\Delta_{\Lambda U}$ hint towards σ_a, but also towards (β, γ). Their test is numerically quite robust for this example. This result is of course not surprising, since the technological shock is the only driving force of the dynamics of the neoclassical growth model. Having a very small impulse makes it tough

[20] For identification issues regarding different priors see also Beltran and Draper (2012) and Onatski and Williams (2010).

[21] See Kim et al. (2008) for details regarding pruning.

[22] Values less than tol (e.g 1e-4=0.0001) are regarded as zero. The rank is calculated using the singular-value-decomposition and calculating the nonzero elements on the diagonal.

Table 2: Neoclassical model, $\sigma_a = 1$

	Iskrev			*Komunjer/Ng*					*Qu/Tkachenko*	
tol	J	M	$\overline{M}$	Δ_Λ	$\Delta_{\Lambda T}$	$\Delta_{\Lambda U}$	Δ	$\overline{\Delta}$	G	$\overline{G}$
1e-01	**7**	6	**7**	**7**	10	**11**	14	14	6	6
1e-02	**7**	**7**	**7**	**7**	**11**	**11**	14	**15**	6	**7**
1e-03	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-04	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-05	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-06	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-07	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-08	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-09	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-10	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-11	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-12	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-13	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-14	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-15	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-16	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-17	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-18	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-19	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-20	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
Require	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**

Ranks of identification tests for different tolerance levels tol, bold indicates full rank.
Lags in autocovariogram $T = 100$, subintervalls $N = 10000$, numerical differentiation step 10^{-7}.

Table 3: Neoclassical model, $\sigma_a = 0.01$

	Iskrev			*Komunjer/Ng*					*Qu/Tkachenko*	
tol	J	M	$\overline{M}$	Δ_Λ	$\Delta_{\Lambda T}$	$\Delta_{\Lambda U}$	Δ	$\overline{\Delta}$	G	$\overline{G}$
1e-01	6	2	4	6	9	9	12	13	2	4
1e-02	**7**	3	5	**7**	**11**	**11**	14	**15**	2	4
1e-03	**7**	3	5	**7**	**11**	**11**	**15**	**15**	2	4
1e-04	**7**	5	6	**7**	**11**	**11**	**15**	**15**	3	5
1e-05	**7**	6	**7**	**7**	**11**	**11**	**15**	**15**	3	5
1e-06	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	3	5
1e-07	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	4	5
1e-08	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	4	6
1e-09	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	6	**7**
1e-10	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	6	**7**
1e-11	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-12	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-13	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-14	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-15	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-16	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-17	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-18	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-19	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
1e-20	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**
Require	**7**	**7**	**7**	**7**	**11**	**11**	**15**	**15**	**7**	**7**

Ranks of identification tests for different tolerance levels tol, bold indicates full rank.
Lags in autocovariogram $T = 100$, subintervalls $N = 10000$, numerical differentiation step 10^{-7}.

to distinguish the other parameters, especially when actually calculating the autocovariance or spectral density. Komunjer & Ng's test is in this example most reliable, since it neither computes autocovariances nor the spectral density, but rather focuses on implications given observational equivalence.

Again, the second order approximation does not change the ranks at all. We also experienced with four different numerical differentiation steps $(1e-3, 1e-7, 1e-11, 1e-15)$ for calculating the derivatives numerically and find that the results stay the same expect for the smallest step size $1e-15$, for which results are very similar to table 2.

In summary, we conclude that the tests yield similar results and our comparative approach is useful to understand the model dynamics and dependencies. Also we are able to get robust insight into possible sources of nonidentification. Further, we show that in the neoclassical model a higher order approximation yields no additional restrictions on the mean that can be used to identify the model.

5.2 The Kim (2003) model

This model extends the neoclassical growth model to include investment adjustment costs twofold: First intertemporal adjustment costs, which involve a nonlinear substitution between capital and investment in capital accumulation, are introduced into the capital accumulation equation govern by a parameter ϕ. Second multisectoral costs, which are captured by a nonlinear transformation between consumption and investment, enter the budget constraint given a parameter θ. In the original paper Kim (2003) log-linearizes the model and shows that there is observational equivalence between these two specifications: "[W]hen a model already has a free parameter for intertemporal adjustment costs, adding another parameter for multisectoral adjustment costs does not enrich the model dynamics" (Kim 2003, p. 534). Table 4 confirms this analytical result throughout all identification tests. Analyzing the nullspace of $J, \Delta_\Lambda, \overline{M}$ and $\overline{G}$ yields unanimously the result that the combination (θ, ϕ) is observational equivalent given a first-order approximation. This holds for tolerance levels as low as $1e-14$ for analytical derivatives. Regarding numerical derivatives we see a dependence on the differentiation step: the smaller the step, the more likely we erroneously conclude an identified model. This is not surprising, since the numerical error of the DSGE solution algorithm will be relative large compared to a very small step size, and we cannot compute the rank precisely. We thus conclude that using analytical derivatives yields numerically better results across all criteria.

Note that identification via the second moments only yields some higher rank deficiencies: checking $\overline{M}$ and Δ additionally yields $(a_0, \theta, \phi, \sigma_a)$ as an unidentified set, whereas checking G yields (a_0, σ_a) and (θ, ϕ) as two unidentifiable subsets. Thus, the mean has already some information to strengthen identification of the model, in this case identifying a_0. Considering an approximation to the second order yields even more restrictions on the mean, as can be seen in table 5. The criteria checking identification via the second moments only stay the same, whereas including the mean restrictions now gives full rank for all tests starting at tolerance levels as large as 1e-4 or even 1e-3, and even for numerical derivatives.

We thus conclude that an approximation to the second order yields additional restrictions on the mean to identify θ and ϕ separately. All tests indicate that θ and ϕ are no longer observational equivalent and the model can be identified using the first and second moments of data. This result – as far as we know – is new to the literature.

Table 4: Identification analysis of the Kim (2003) model, 1st-order approximation

	Iskrev			*Komunjer/Ng*					*Qu/Tkachenko*	
Tol	J	M	$\overline{M}$	Δ_Λ	$\Delta_{\Lambda T}$	$\Delta_{\Lambda U}$	Δ	$\overline{\Delta}$	G	$\overline{G}$
1e-01	6 (6,6,6)	6 (6,6,6)	7 (7,7,7)	5 (5,5,5)	9 (9,9,9)	9 (9,9,9)	12 (12,12,12)	12 (12,12,12)	5 (5,5,5)	5 (5,5,5)
1e-02	7 (7,7,7)	6 (6,6,6)	7 (7,7,7)	6 (6,6,6)	10 (10,10,10)	10 (10,10,10)	13 (13,13,13)	14 (14,14,14)	6 (6,6,6)	7 (7,7,7)
1e-03	7 (7,7,7)	6 (6,6,7)	7 (7,7,**8**)	7 (7,7,7)	11 (11,11,11)	11 (11,11,11)	14 (14,14,14)	15 (15,15,15)	6 (6,6,6)	7 (7,7,7)
1e-04	7 (7,7,**8**)	6 (6,6,7)	7 (7,7,**8**)	7 (7,7,**8**)	11 (11,11,**12**)	11 (11,11,**12**)	14 (14,14,15)	15 (15,15,**16**)	6 (6,6,6)	7 (7,7,7)
1e-05	7 (7,7,**8**)	6 (6,6,**8**)	7 (7,7,**8**)	7 (7,7,**8**)	11 (11,11,**12**)	11 (11,11,**12**)	14 (14,14,15)	15 (15,15,**16**)	6 (6,6,6)	7 (7,7,7)
1e-06	7 (7,7,**8**)	6 (6,6,**8**)	7 (7,7,**8**)	7 (7,7,**8**)	11 (11,11,**12**)	11 (11,11,**12**)	14 (14,14,**16**)	15 (15,15,**16**)	6 (6,6,7)	7 (7,7,**8**)
1e-07	7 (7,**8**,**8**)	6 (6,7,**8**)	7 (7,**8**,**8**)	7 (7,**8**,**8**)	11 (11,**12**,**12**)	11 (11,11,**12**)	14 (14,14,**16**)	15 (15,15,**16**)	6 (6,6,7)	7 (7,7,**8**)
1e-08	7 (7,**8**,**8**)	6 (6,7,**8**)	7 (7,**8**,**8**)	7 (7,**8**,**8**)	11 (11,**12**,**12**)	11 (11,**12**,**12**)	14 (14,15,**16**)	15 (15,**16**,**16**)	6 (6,6,7)	7 (7,7,**8**)
1e-09	7 (7,**8**,**8**)	6 (7,7,**8**)	7 (**8**,**8**,**8**)	7 (7,**8**,**8**)	11 (11,**12**,**12**)	11 (11,**12**,**12**)	14 (14,15,**16**)	15 (15,**16**,**16**)	6 (6,6,**8**)	7 (7,7,**8**)
1e-10	7 (**8**,**8**,**8**)	6 (7,**8**,**8**)	7 (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	14 (15,15,**16**)	15 (**16**,**16**,**16**)	6 (6,6,**8**)	7 (7,7,**8**)
1e-11	7 (**8**,**8**,**8**)	6 (7,**8**,**8**)	7 (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	14 (15,**16**,**16**)	15 (**16**,**16**,**16**)	6 (6,6,**8**)	7 (7,7,**8**)
1e-12	7 (**8**,**8**,**8**)	6 (7,**8**,**8**)	7 (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	14 (15,**16**,**16**)	15 (**16**,**16**,**16**)	6 (6,6,**8**)	7 (7,7,**8**)
1e-13	7 (**8**,**8**,**8**)	6 (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	14 (15,**16**,**16**)	15 (**16**,**16**,**16**)	7 (7,7,**8**)	7 (7,7,**8**)
1e-14	7 (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	14 (**16**,**16**,**16**)	15 (**16**,**16**,**16**)	7 (7,**8**,**8**)	7 (7,**8**,**8**)
1e-15	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	15 (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	7 (7,**8**,**8**)
1e-16	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**12** (**12**,**12**,**12**)	**12** (**12**,**12**,**12**)	**16** (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)
1e-17	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**12** (**12**,**12**,**12**)	**12** (**12**,**12**,**12**)	**16** (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)
1e-18	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**12** (**12**,**12**,**12**)	**12** (**12**,**12**,**12**)	**16** (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)
1e-19	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**12** (**12**,**12**,**12**)	**12** (**12**,**12**,**12**)	**16** (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)
1e-20	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**12** (**12**,**12**,**12**)	**12** (**12**,**12**,**12**)	**16** (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)
Require	**8**	**8**	**8**	**8**	**12**	**12**	**16**	**16**	**8**	**8**

Ranks of identification tests with analytical derivatives for different tolerance levels tol, lags in autocovariogram $T = 100$, subintervalls $N = 10000$. Bold indicates full rank.

In parenthesis are the corresponding ranks computed with numerical derivatives given differentiation steps 1e-3, 1e-7 and 1e-11, respectively.

Table 5: Identification analysis of the Kim (2003) model, 2^{nd}-order approximation

	Iskrev			*Komunjer/Ng*					*Qu/Tkachenko*	
Tol	J	M	$\overline{M}$	Δ_Λ	$\Delta_{\Lambda T}$	$\Delta_{\Lambda U}$	Δ	$\overline{\Delta}$	G	$\overline{G}$
1e-01	**8** (**8**,**8**,**8**)	6 (6,6,6)	6 (6,6,6)	5 (5,5,5)	9 (9,9,9)	9 (9,9,9)	12 (12,12,12)	13 (13,13,13)	5 (5,5,5)	5 (5,5,5)
1e-02	**8** (**8**,**8**,**8**)	6 (6,6,6)	7 (7,7,7)	6 (6,6,6)	10 (10,10,10)	10 (10,10,10)	13 (13,13,13)	14 (14,14,14)	6 (6,6,6)	7 (7,7,7)
1e-03	**8** (**8**,**8**,**8**)	6 (6,6,7)	**8** (**8**,**8**,**8**)	7 (7,7,7)	11 (11,11,11)	11 (11,11,11)	14 (14,14,14)	15 (15,15,15)	6 (6,6,6)	7 (7,7,7)
1e-04	**8** (**8**,**8**,**8**)	6 (6,6,7)	**8** (**8**,**8**,**8**)	7 (7,7,**8**)	11 (11,11,**12**)	11 (11,11,**12**)	14 (14,14,15)	**16** (**16**,**16**,**16**)	6 (6,6,6)	7 (7,7,7)
1e-05	**8** (**8**,**8**,**8**)	6 (6,6,**8**)	**8** (**8**,**8**,**8**)	7 (7,7,**8**)	11 (11,11,**12**)	11 (11,11,**12**)	14 (14,14,15)	**16** (**16**,**16**,**16**)	6 (6,6,6)	**8** (**8**,**8**,**8**)
1e-06	**8** (**8**,**8**,**8**)	6 (6,6,**8**)	**8** (**8**,**8**,**8**)	7 (7,7,**8**)	11 (11,11,**12**)	11 (11,11,**12**)	14 (14,14,**16**)	**16** (**16**,**16**,**16**)	6 (6,6,7)	**8** (**8**,**8**,**8**)
1e-07	**8** (**8**,**8**,**8**)	6 (6,7,**8**)	**8** (**8**,**8**,**8**)	7 (7,**8**,**8**)	11 (11,**12**,**12**)	11 (11,11,**12**)	14 (14,14,**16**)	**16** (**16**,**16**,**16**)	6 (6,6,7)	**8** (**8**,**8**,**8**)
1e-08	**8** (**8**,**8**,**8**)	6 (6,7,**8**)	**8** (**8**,**8**,**8**)	7 (7,**8**,**8**)	11 (11,**12**,**12**)	11 (11,**12**,**12**)	14 (14,15,**16**)	**16** (**16**,**16**,**16**)	6 (6,6,7)	**8** (**8**,**8**,**8**)
1e-09	**8** (**8**,**8**,**8**)	6 (7,7,**8**)	**8** (**8**,**8**,**8**)	7 (7,**8**,**8**)	11 (11,**12**,**12**)	11 (11,**12**,**12**)	14 (14,15,**16**)	**16** (**16**,**16**,**16**)	6 (6,6,**8**)	**8** (**8**,**8**,**8**)
1e-10	**8** (**8**,**8**,**8**)	6 (7,**8**,**8**)	**8** (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	14 (15,15,**16**)	**16** (**16**,**16**,**16**)	6 (6,6,**8**)	**8** (**8**,**8**,**8**)
1e-11	**8** (**8**,**8**,**8**)	6 (7,**8**,**8**)	**8** (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	14 (15,**16**,**16**)	**16** (**16**,**16**,**16**)	6 (6,6,**8**)	**8** (**8**,**8**,**8**)
1e-12	**8** (**8**,**8**,**8**)	6 (7,**8**,**8**)	**8** (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	14 (15,**16**,**16**)	**16** (**16**,**16**,**16**)	6 (6,6,**8**)	**8** (**8**,**8**,**8**)
1e-13	**8** (**8**,**8**,**8**)	6 (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	14 (15,**16**,**16**)	**16** (**16**,**16**,**16**)	7 (7,7,**8**)	**8** (**8**,**8**,**8**)
1e-14	**8** (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	14 (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	7 (7,**8**,**8**)	**8** (**8**,**8**,**8**)
1e-15	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	7 (**8**,**8**,**8**)	11 (**12**,**12**,**12**)	11 (**12**,**12**,**12**)	15 (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)
1e-16	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**12** (**12**,**12**,**12**)	**12** (**12**,**12**,**12**)	**16** (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)
1e-17	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**12** (**12**,**12**,**12**)	**12** (**12**,**12**,**12**)	**16** (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)
1e-18	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**12** (**12**,**12**,**12**)	**12** (**12**,**12**,**12**)	**16** (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)
1e-19	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**12** (**12**,**12**,**12**)	**12** (**12**,**12**,**12**)	**16** (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)
1e-20	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)	**12** (**12**,**12**,**12**)	**12** (**12**,**12**,**12**)	**16** (**16**,**16**,**16**)	**16** (**16**,**16**,**16**)	**8** (**8**,**8**,**8**)	**8** (**8**,**8**,**8**)
Require	**8**	**8**	**8**	**8**	**12**	**12**	**16**	**16**	**8**	**8**

Ranks of identification tests with analytical derivatives for different tolerance levels tol, lags in autocovariogram $T = 100$, subintervalls $N = 10000$. Bold indicates full rank.

In parenthesis are the corresponding ranks computed with numerical derivatives given differentiation steps 1e-3, 1e-7 and 1e-11, respectively.

5.3 The An and Schorfheide (2007) model

This model is a prototypical DSGE model often cited in the literature. The authors already show that the parameter ν and the steady-state ratio c/y do not influence the log-linearized solution. Also they indicate that the parameters entering the Taylor-rule are at best only weakly identified. Komunjer and Ng (2011), Ratto and Iskrev (2011) and Qu and Tkachenko (2012) confirm this and show that indeed the parameters $\psi_1, \psi_2, \rho_R, \sigma_R$ are not separably distinguishable using the methods described in section 3. However, they all use a slightly different version of the model (log-linearized model, simplified measurement equations, different parametrization), hence, comparing the details of their results is only partly possible. Therefore, we will focus on the original model in its non-linearized form and check the identification criteria across different tolerance levels.

Table 6 shows results for the linear approximation to the first order for analytical and numerical derivatives with different differentiation steps. First, turning to analytical derivatives, all criteria yield correctly rank-deficiency by three, however, there are slight differences for which tolerance levels. The thresholds are 1e-15 for Iskrev's method, 1e-11 for Komunjer and Ng's criteria and 1e-12 for Qu and Tkachenko's test considering identification through first and second moments. The problematic parameter sets are shown in columns two, four and six of table 8 for the first-order approximation. We can robustly confirm that from an identification point of view we have to fix three parameters for the log-linearized model: ν, c/y and one parameter out of the set $(\psi_1, \psi_2, \rho_R, \sigma_R)$. For very small, i.e. strict tolerance levels Iskrev's and Qu and Tkachenko's method fail to detect the Taylor-rule coefficients, whereas Komunjer and Ng's approach even points wrongly to an identified model. On the other hand for very large tolerance levels (e.g. 1e-3=0.001) we are only on the verge of detecting σ_R, since $\sigma_R = 0.002$ is of the same magnitude as the tolerance level. Thus, when selecting a tolerance level for the rank calculation, one has to be aware of the magnitude of the local point as well as not to set it too strict. This issue becomes even more severe, when we calculate the derivatives numerically. Now we have a trade-off between setting the differentiation step too large (e.g. 1e-3), and thus possibly calculating the derivatives imprecisely, or too small (e.g. 1e-11), such that the numerical error from the solution algorithm possibly outweighs the differentiation error. Both result in false rank calculation, as can be seen in table 6, where we fail to detect lack of identification of the Taylor-rule coefficients even for mild tolerance levels across all approaches. Using a feasible trade-off for the numerical differentiation step (e.g. 1e-7), the threshold of correctly determining rank-deficiency by three are now 1e-7 for Iskrev's method, 1e-8 for Komunjer and Ng's criteria and 1e-12 for Qu and Tkachenko's test. So compared to the analytical case, we have to loosen our tolerance level for the rank calculations or otherwise we get wrong results. Therefore we are strongly in favor of analytical procedures and advise using them whenever feasible; the results are unanimously more reliable even for stricter tolerance levels.

Further J and Δ_Λ can be used for diagnostic issues of the model. J is rank-deficient by two, i.e. ν and c/y do not enter neither the mean nor the first-order solution matrices, whereas Δ_Λ is short by 4, i.e. $\nu, c/y, \pi^{(A)}$ and $\gamma^{(Q)}$ do not enter the solution matrices. Thus, we conclude that $\pi^{(A)}$ and $\gamma^{(Q)}$ are only identified via the mean using the first-order approximation, which is in accordance to the measurement equations.

Now, considering an approximation to the second order, the picture changes slightly, see table 7 for the results. Now J is full rank: the mean, and the first- and second-order solution matrices seem to have information for identifying all parameters. Note that a full-rank J is only necessary,

but not sufficient for identification, as is obvious when looking at the ranks of $\overline{M}$, $\overline{\Delta}$ and $\overline{G}$. The influence of different tolerance levels as well as the effect of analytical vs. numerical derivatives are apparently very similar to the first-order approximation, since in the relevant cases we yield rank deficiencies by three. However, columns three, five and seven of table 8 yield some alluring insight about the problematic parameters. For Iskrev's and Komunjer & Ng's criteria, it's the combination of $(\nu, c/y, \psi_1, \psi_2, \rho_R, \sigma_R, \pi^{(A)})$ that jointly yield three problematic sets, whereas Qu & Tkachenko's test points towards $(\nu, \pi^{(A)})$, $(c/y, \pi^{(A)})$ and $(\nu, c/y)$. The appearance of $\pi^{(A)}$ in the problematic sets is very appealing, since fixing $\pi^{(A)}, c/y$ and ν we are indeed able to identify the Taylor-rule coefficients. There is information in the mean given by the long-run inflation that spills over to the Taylor-rule coefficients.

In summary, we conclude that our comparative approach is insightful for this prototypical small DSGE model and we are able to get robust insight into possible sources of nonidentification.[23] All tests yield similar results, however, the choice of tolerance level and numerical differentiation step is a delicate one. Using analytical derivatives as well as comparing the output of the different approaches seems to us the best-practice to gauge identification of the DSGE model. Further, we show that in the An and Schorfheide (2007) model a higher order approximation yields additional restrictions on the mean that can be used to identify the coefficients entering the Taylor rule. This feature of the nonlinear model – as far as we know – is new to the literature.

6 Conclusion [Incomplete]

The goal of this paper is to add a comprehensive study of identification methods to the literature on identification in DSGE models and to emphasize its importance prior to estimation. The contribution of this paper is threefold: First, we analyze and derive all methods for identification of DSGE models theoretically in a consistent notation and framework. We are thus able to see similarities, like the dependence on the solution matrices, and differences like computing autocovariances, spectral densities or other matrices. Also practical and numerical difficulties applying the methods are discussed, and how these can lead to different conclusions. In particular, we derive analytical derivatives for all methods and showe that using these make all methods more robust compared to the use of numerical derivatives, since these heavily dependent on the numerical differentiation step. We argue in favor of using analytical derivatives, whenever feasible, due to its robustness and greater speed than relying on numerical procedures. Second, it is shown how to analytically extend the methods to linear approximations to the second order. It is argued that this can help increasing overall identification of a DSGE model via imposing additional restrictions. In this way we are able to identify the Kim (2003) model from the first two moments of data using a second order approximation as well as the coefficients of the Taylor-rule in the An and Schorfheide (2007) model. Third, we applied all methods on DSGE models that are known to have lack of identification. Most of the times the methods come to the same conclusion, however, the issue of numerical errors due to nonlinearities and very large matrices can make these methods tedious in practice and may lead to unreliable or contradictory conclusions. The example models show that by evaluating different criteria we also gain inside into the dynamic structure of the DSGE model. We argue that in order to thoroughly analyze identification, one has to be aware of the drawbacks of the different methods and check whether different methods

[23] In the Matlab code we also analyze bigger models, i.e. models by Garcia-Cicco, Pancrazi, and Uribe (2010), Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007).

Table 6: Identification analysis of the An and Schorfheide (2007) model, 1st-order approximation

	Iskrev			*Komunjer/Ng*					*Qu/Tkachenko*	
Tol	J	M	$\overline{M}$	Δ_Λ	$\Delta_{\Lambda T}$	$\Delta_{\Lambda U}$	Δ	$\overline{\Delta}$	G	$\overline{G}$
1e-01	10 (10,10,10)	9 (9,9,9)	12 (12,12,12)	8 (8,8,8)	17 (17,17,17)	16 (16,16,16)	23 (23,23,23)	26 (26,26,26)	7 (7,7,7)	9 (9,9,9)
1e-02	11 (11,11,11)	9 (9,9,9)	12 (12,12,12)	9 (9,9,9)	18 (18,18,18)	17 (17,17,17)	25 (25,25,25)	28 (28,28,28)	9 (9,9,9)	12 (12,12,12)
1e-03	13 (13,13,13)	10 (10,10,11)	12 (12,12,13)	11 (11,11,11)	20 (20,20,20)	19 (19,19,19)	27 (27,27,27)	30 (30,30,30)	9 (9,9,9)	12 (12,12,12)
1e-04	13 (13,13,13)	10 (10,10,11)	12 (12,12,13)	11 (11,11,11)	20 (20,20,20)	19 (19,19,20)	28 (28,28,28)	30 (30,30,30)	9 (9,9,9)	12 (12,12,12)
1e-05	13 (13,13,13)	10 (11,10,11)	12 (13,12,13)	11 (11,11,11)	20 (20,20,20)	19 (19,19,20)	28 (28,28,29)	30 (30,30,31)	9 (9,9,10)	12 (12,12,12)
1e-06	13 (13,13,13)	10 (11,10,11)	12 (13,12,13)	11 (11,11,11)	20 (20,20,20)	19 (19,19,20)	28 (28,28,29)	30 (30,30,31)	10 (10,10,10)	12 (12,12,13)
1e-07	13 (13,13,13)	10 (11,10,11)	12 (13,13,13)	11 (11,11,11)	20 (20,20,20)	19 (20,19,20)	28 (29,28,29)	30 (31,30,31)	10 (10,10,11)	12 (12,12,13)
1e-08	13 (13,13,13)	10 (11,11,11)	12 (13,13,13)	11 (11,11,11)	20 (20,20,20)	19 (20,20,20)	28 (29,28,29)	30 (31,30,31)	10 (10,10,11)	12 (12,12,13)
1e-09	13 (13,13,13)	10 (11,11,11)	12 (13,13,13)	11 (11,11,11)	20 (20,20,20)	19 (20,20,20)	28 (29,29,29)	30 (31,31,31)	10 (10,10,11)	12 (12,12,13)
1e-10	13 (13,13,13)	10 (11,11,11)	12 (13,13,13)	11 (11,11,11)	20 (20,20,20)	19 (20,20,20)	28 (29,29,29)	30 (31,31,31)	10 (10,10,11)	12 (12,12,13)
1e-11	13 (13,13,13)	10 (11,11,11)	12 (13,13,13)	11 (11,11,11)	20 (20,20,20)	19 (20,20,20)	28 (29,29,29)	30 (31,31,31)	10 (11,10,11)	12 (13,12,13)
1e-12	13 (13,13,13)	10 (11,11,11)	12 (13,13,13)	11 (11,11,11)	20 (20,20,20)	19 (20,20,20)	29 (30,29,29)	31 (31,31,31)	10 (11,10,11)	12 (13,12,13)
1e-13	13 (13,13,13)	10 (11,11,11)	12 (13,13,13)	11 (11,11,11)	20 (20,20,20)	20 (20,20,20)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	13 (13,13,13)
1e-14	13 (13,13,13)	11 (11,11,11)	12 (13,13,13)	11 (11,11,11)	21 (20,21,22)	21 (21,20,22)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	13 (13,13,13)
1e-15	13 (13,13,13)	11 (11,11,11)	12 (13,13,13)	11 (11,11,11)	23 (**24**,23,**24**)	**24** (**24**,23,23)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	13 (13,13,13)
1e-16	13 (13,13,13)	11 (11,11,11)	13 (13,13,13)	11 (11,11,11)	23 (**24**,**24**,**24**)	**24** (**24**,**24**,**24**)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	13 (13,13,13)
1e-17	13 (13,13,13)	12 (12,11,12)	13 (13,13,13)	11 (11,11,11)	**24** (**24**,**24**,**24**)	**24** (**24**,**24**,**24**)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	13 (13,13,13)
1e-18	13 (13,13,13)	12 (12,12,12)	13 (13,13,13)	11 (11,11,11)	**24** (**24**,**24**,**24**)	**24** (**24**,**24**,**24**)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	13 (13,13,13)
1e-19	13 (13,13,13)	12 (12,12,13)	13 (13,13,13)	11 (11,11,11)	**24** (**24**,**24**,**24**)	**24** (**24**,**24**,**24**)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	13 (13,13,13)
1e-20	13 (13,13,13)	13 (13,13,13)	13 (13,13,13)	11 (11,11,11)	**24** (**24**,**24**,**24**)	**24** (**24**,**24**,**24**)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	12 (11,11,12)	13 (13,13,13)
Require	**15**	**15**	**15**	**15**	**24**	**24**	**33**	**33**	**15**	**15**

Ranks of identification tests with analytical derivatives for different tolerance levels tol, lags in autocovariogram $T = 100$, subintervalls $N = 10000$. Bold indicates full rank.

In parenthesis are the corresponding ranks computed with numerical derivatives given differentiation steps 1e-3, 1e-7 and 1e-11, respectively.

Table 7: Identification analysis of the An and Schorfheide (2007) model, 2nd-order approximation

	Iskrev			*Komunjer/Ng*					*Qu/Tkachenko*	
Tol	J	M	$\overline{M}$	Δ_Λ	$\Delta_{\Lambda T}$	$\Delta_{\Lambda U}$	Δ	$\overline{\Delta}$	G	$\overline{G}$
1e-01	13 (13,13,13)	9 (9,9,9)	12 (12,12,12)	8 (8,8,8)	17 (17,17,17)	16 (16,16,16)	23 (23,23,23)	26 (26,26,26)	7 (7,7,7)	9 (9,9,9)
1e-02	**15** (**15**,**15**,**15**)	9 (9,9,9)	12 (12,12,12)	9 (9,9,9)	18 (18,18,18)	17 (17,17,17)	25 (25,25,25)	28 (28,28,28)	9 (9,9,9)	12 (12,12,12)
1e-03	**15** (**15**,**15**,**15**)	10 (10,10,11)	12 (12,12,13)	11 (11,11,11)	20 (20,20,20)	19 (19,19,19)	27 (27,27,27)	29 (29,29,29)	9 (9,9,9)	12 (12,12,12)
1e-04	**15** (**15**,**15**,**15**)	10 (10,10,11)	12 (12,12,13)	11 (11,11,11)	20 (20,20,20)	19 (19,19,20)	28 (28,28,28)	30 (30,30,30)	9 (9,9,9)	12 (12,12,12)
1e-05	**15** (**15**,**15**,**15**)	10 (11,10,11)	12 (13,12,13)	11 (11,11,11)	20 (20,20,20)	19 (19,19,20)	28 (28,28,29)	30 (30,30,31)	9 (9,9,10)	12 (12,12,12)
1e-06	**15** (**15**,**15**,**15**)	10 (11,10,11)	12 (13,12,13)	11 (11,11,11)	20 (20,20,20)	19 (19,19,20)	28 (28,28,29)	30 (30,30,31)	10 (10,10,10)	12 (12,12,13)
1e-07	**15** (**15**,**15**,**15**)	10 (11,10,11)	12 (13,13,14)	11 (11,11,11)	20 (20,20,20)	19 (20,19,20)	28 (29,28,29)	30 (31,30,31)	10 (10,10,11)	12 (12,12,13)
1e-08	**15** (**15**,**15**,**15**)	10 (11,11,11)	12 (13,13,14)	11 (11,11,11)	20 (20,20,20)	19 (20,20,20)	28 (29,28,29)	30 (31,30,32)	10 (10,10,11)	12 (12,12,13)
1e-09	**15** (**15**,**15**,**15**)	10 (11,11,11)	12 (13,13,14)	11 (11,11,11)	20 (20,20,20)	19 (20,20,20)	28 (29,29,29)	30 (31,31,32)	10 (10,10,11)	12 (12,12,13)
1e-10	**15** (**15**,**15**,**15**)	10 (11,11,11)	12 (13,13,14)	11 (11,11,11)	20 (20,20,20)	19 (20,20,20)	28 (29,29,29)	30 (31,31,32)	10 (10,10,11)	12 (12,12,13)
1e-11	**15** (**15**,**15**,**15**)	10 (11,11,11)	12 (13,14,14)	11 (11,11,11)	20 (20,20,20)	19 (20,20,20)	28 (29,29,29)	30 (31,31,32)	10 (11,10,11)	12 (13,12,13)
1e-12	**15** (**15**,**15**,**15**)	10 (11,11,11)	12 (13,14,14)	11 (11,11,11)	20 (20,20,20)	19 (20,20,20)	29 (30,29,29)	30 (31,32,32)	10 (11,10,11)	12 (13,12,13)
1e-13	**15** (**15**,**15**,**15**)	10 (11,11,11)	12 (13,14,14)	11 (11,11,11)	20 (20,20,20)	20 (20,20,20)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	13 (13,13,13)
1e-14	**15** (**15**,**15**,**15**)	11 (11,11,11)	12 (13,14,14)	11 (11,11,11)	21 (20,21,22)	21 (21,20,22)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	13 (13,13,14)
1e-15	**15** (**15**,**15**,**15**)	11 (11,11,11)	13 (14,14,14)	11 (11,11,11)	23 (**24**,23,**24**)	**24** (**24**,23,23)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	13 (13,13,14)
1e-16	**15** (**15**,**15**,**15**)	11 (11,11,11)	14 (**15**,**15**,**15**)	11 (11,11,11)	23 (**24**,**24**,**24**)	**24** (**24**,**24**,**24**)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	14 (14,14,**15**)
1e-17	**15** (**15**,**15**,**15**)	12 (12,11,12)	**15** (**15**,**15**,**15**)	11 (11,11,11)	**24** (**24**,**24**,**24**)	**24** (**24**,**24**,**24**)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	**15** (**15**,**15**,**15**)
1e-18	**15** (**15**,**15**,**15**)	12 (12,12,12)	**15** (**15**,**15**,**15**)	11 (11,11,11)	**24** (**24**,**24**,**24**)	**24** (**24**,**24**,**24**)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	**15** (**15**,**15**,**15**)
1e-19	**15** (**15**,**15**,**15**)	12 (12,12,13)	**15** (**15**,**15**,**15**)	11 (11,11,11)	**24** (**24**,**24**,**24**)	**24** (**24**,**24**,**24**)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	11 (11,11,11)	**15** (**15**,**15**,**15**)
1e-20	**15** (**15**,**15**,**15**)	13 (13,13,13)	**15** (**15**,**15**,**15**)	11 (11,11,11)	**24** (**24**,**24**,**24**)	**24** (**24**,**24**,**24**)	**33** (**33**,**33**,**33**)	**33** (**33**,**33**,**33**)	12 (11,11,12)	**15** (**15**,**15**,**15**)
Require	**15**	**15**	**15**	**15**	**24**	**24**	**33**	**33**	**15**	**15**

Ranks of identification tests with analytical derivatives for different tolerance levels tol, lags in autocovariogram $T = 100$, subintervalls $N = 10000$. Bold indicates full rank.

In parenthesis are the corresponding ranks computed with numerical derivatives given differentiation steps 1e-3, 1e-7 and 1e-11, respectively.

Table 8: Analysis of problematic parameters for the An and Schorfheide (2007) model

	Iskrev		*Komunjer/Ng*		*Qu/Tkachenko*	
Tol	$\overline{M}_{1^{st}}$	$\overline{M}_{2^{nd}}$	$\overline{\Delta}_{1^{st}}$	$\overline{\Delta}_{2^{nd}}$	$\overline{G}_{1^{st}}$	$\overline{G}_{2^{nd}}$
1e-03	v	$c/y, v, \psi_1, \psi_2, \pi^{(A)}$	$c/y, v, \psi_1, \psi_2, \rho_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \pi^{(A)}$	v	$v, \pi^{(A)}$
	c/y	$c/y, v, \psi_1, \psi_2, \pi^{(A)}$	$c/y, v, \psi_1, \psi_2, \rho_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \pi^{(A)}$	c/y	$c/y, \pi^{(A)}$
	ψ_1, ψ_2, ρ_R	$c/y, v, \psi_1, \psi_2, \pi^{(A)}, \rho_R$	$c/y, v, \psi_1, \psi_2, \rho_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \pi^{(A)}$	$\psi_1, \psi_2, \rho_R, \sigma_R$	$v, c/y$
				$c/y, v, \psi_1, \psi_2, \rho_R, \pi^{(A)}, \sigma_R, r^{(A)}, \sigma_z$		
1e-07	v	$c/y, v, \psi_1, \psi_2, \pi^{(A)}, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \pi^{(A)}, \sigma_R$	v	$v, \pi^{(A)}$
	c/y	$c/y, v, \psi_1, \psi_2, \pi^{(A)}, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \pi^{(A)}, \sigma_R$	c/y	$c/y, \pi^{(A)}$
	$\psi_1, \psi_2, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \pi^{(A)}, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \pi^{(A)}, \sigma_R$	$\psi_1, \psi_2, \rho_R, \sigma_R$	$v, c/y$
1e-11	v	$c/y, v, \psi_1, \psi_2, \pi^{(A)}, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \pi^{(A)}, \sigma_R, r^{(A)}$	v	$v, \pi^{(A)}$
	c/y	$c/y, v, \psi_1, \psi_2, \pi^{(A)}, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \pi^{(A)}, \sigma_R, r^{(A)}$	c/y	$c/y, \pi^{(A)}$
	$\psi_1, \psi_2, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \pi^{(A)}, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \sigma_R$	$c/y, v, \psi_1, \psi_2, \rho_R, \pi^{(A)}, \sigma_R, r^{(A)}$	$\psi_1, \psi_2, \rho_R, \sigma_R$	$v, c/y$
1e-17	v				v	
	c/y				c/y	

Parameter sets that are responsible for rank-deficiency at selected tolerance levels for first (1st) and second (2nd) order.

Method: analytical derivatives, lags in autocovariogram $T = 100$, subintervalls $N = 10000$.

come to the same conclusion. Our Matlab code is written model-independent and can be used easily to check identification of other models, as long as they can be represented in the proposed framework.

7 References

An, Sungbae and Frank Schorfheide (2007). "Bayesian Analysis of DSGE Models". In: *Econometric Reviews* 26.2-4, pp. 113–172 (cit. on pp. 1, 2, 21, 26–30, 40).

Anderson, Gary (2008). "Solving Linear Rational Expectations Models: A Horse Race". In: *Computational Economics* 31.2, pp. 95–113 (cit. on p. 5).

Beltran, Daniel O. and David Draper (2012). *Estimating dynamic macroeconomic models: How informative are the data?* Working Paper Presented at a Conference. Version January 24, 2012. Federal Reserve Board of Governors and University of California. URL: `https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=CEF2012\&paper_id=283` (visited on 09/17/2012) (cit. on pp. 1, 21).

Blanchard, Olivier J. and Charles M. Kahn (1980). "The Solution of Linear Difference Models under Rational Expectations". In: *Econometrica* 48.5, pp. 1305–1311 (cit. on p. 5).

Caglar, Evren, Jagjit S. Chadha, and Katsuyuki Shibayama (2012). *Bayesian Estimation of DSGE Models: Is the Workhorse Model Identified?* Koç University-TUSIAD Economic Research Forum Working Papers 1205. Version February 2012. Koc University. URL: `http://eaf.ku.edu.tr/sites/eaf.ku.edu.tr/files/erf_wp_1205.pdf` (visited on 09/12/2012) (cit. on p. 20).

Canova, Fabio and Luca Sala (2009). "Back to square one: Identification issues in DSGE models". In: *Journal of Monetary Economics* 56.4, pp. 431–449 (cit. on p. 1).

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (2005). "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy". In: *Journal of Political Economy* 113.1, pp. 1–45 (cit. on pp. 21, 27).

Fisher, Franklin M. (1966). *The Identification Problem in Econometrics*. New York: McGraw-Hill (cit. on p. 1).

Garcia-Cicco, Javier, Roberto Pancrazi, and Martin Uribe (2010). "Real Business Cycles in Emerging Countries?" In: *American Economic Review* 100.5, pp. 2510–2531 (cit. on pp. 21, 27).

Gomme, Paul and Paul Klein (2011). "Second-order approximation of dynamic models without the use of tensors". In: *Journal of Economic Dynamics and Control* 35.4, pp. 604–615 (cit. on pp. 2, 3, 6).

Guerron-Quintana, Pablo A., Atsushi Inoue, and Lutz Kilian (2012). "Frequentist Inference in Weakly Identified DSGE Models". In: *Computational Economics*. Forthcoming (cit. on p. 1).

Hsiao, Cheng (1983). "Identification". In: *Handbook of Econometrics*. Ed. by Z. Griliches and Intriligator M. D. Vol. 1. Elsevier, pp. 223–283 (cit. on p. 1).

Iskrev, Nikolay I. (2008). "Evaluating the information matrix in linearized DSGE models". In: *Economics Letters* 99.3, pp. 607–610 (cit. on p. 7).

Iskrev, Nikolay I. (2010). "Local identification in DSGE models". In: *Journal of Monetary Economics* 57.2, pp. 189–202 (cit. on pp. 1, 9).

Kadane, Joseph B. (1974). "The Role of Identification in Bayesian Theory". In: *Studies in Bayesian econometrics and statistics: in honor of Leonard J. Savage*. Ed. by S.E Fienberg and A. Zellner. Contributions to economic analysis. Amsterdam: North-Holland Pub. Co., pp. 175–191 (cit. on p. 17).

Kim, Jinill (2003). "Functional equivalence between intertemporal and multisectoral investment adjustment costs". In: *Journal of Economic Dynamics and Control* 27.4, pp. 533–549 (cit. on pp. 2, 21, 23–25, 27, 39).

Kim, Jinill et al. (2008). "Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models". In: *Journal of Economic Dynamics and Control* 32.11, pp. 3397–3414 (cit. on pp. 21, 38).

Klein, Paul (2000). "Using the generalized Schur form to solve a multivariate linear rational expectations model". In: *Journal of Economic Dynamics and Control* 24.10, pp. 1405–1423 (cit. on p. 5).

Komunjer, Ivana (2012). "Global Identification in Nonlinear Models With Moment Restrictions". In: *Econometric Theory* 28.04, pp. 719–729. (Visited on 2012) (cit. on p. 16).

Komunjer, Ivana and Serena Ng (2011). "Dynamic Identification of Dynamic Stochastic General Equilibrium Models". In: *Econometrica* 79.6, pp. 1995–2032 (cit. on pp. 1, 10, 26).

Koop, Gary, M. Hashem Pesaran, and Ron P. Smith (2012). *On Identification of Bayesian DSGE Models*. Working Paper. Version August 2012. University of Strathclyde, University of Southern California and Trinity College, Cambridge, and Birkbeck College. URL: `http://www.econ.cam.ac.uk/faculty/pesaran/wp12/KPS6August12.pdf` (visited on 09/12/2012) (cit. on pp. 1, 17, 18).

Magnus, Jan R. (2010). "On the concept of matrix derivative". In: *Journal of Multivariate Analysis* 101.9, pp. 2200–2206 (cit. on p. 3).

Magnus, Jan R. and Heinz Neudecker (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. 2nd ed. Wiley (cit. on pp. 3, 5, 7, 8, 13, 35, 36).

Mutschler, Willi (2012). "Analyse, Lösung und Schätzung von DSGE-Modellen: Ein Vergleich der Methoden". Master thesis. Münster: University of Münster. URL: `http://www.mutschler.eu/masterarbeit.pdf` (visited on 09/18/2012) (cit. on p. 1).

Onatski, Alexei and Noah Williams (2010). "Empirical and policy performance of a forward-looking monetary model". In: *Journal of Applied Econometrics* 25.1, pp. 145–176 (cit. on p. 21).

Poirier, Dale J. (1998). "Revising Beliefs in Nonidentified Models". In: *Econometric Theory* 14.4, pp. 483–509 (cit. on pp. 17, 18).

Pollock, D. Stephen G. (2013). "On Kronecker products, tensor products and matrix differential calculus". In: *International Journal of Computer Mathematics*, pp. 1–15 (cit. on p. 3).

Qu, Zhongjun and Denis Tkachenko (2012). "Identification and frequency domain quasi-maximum likelihood estimation of linearized dynamic stochastic general equilibrium models". In: *Quantitative Economics* 3.1, pp. 95–132 (cit. on pp. 1, 26).

Ratto, Marco and Nikolay I. Iskrev (2011). *Identification analysis of DSGE models with DYNARE*. MONFISPOL 225149. Version August 2011. European Commission and Banco de Portugal. URL: `https://www.ifk-cfs.de/fileadmin/downloads/events/conferences/monfispol2011/RATTO_IdentifFinal.pdf` (visited on 09/12/2012) (cit. on p. 26).

Ríos-Rull, José-Víctor et al. (2012). "Methods versus substance: Measuring the effects of technology shocks". In: *Journal of Monetary Economics* 59.8, pp. 826–846 (cit. on p. 1).

Rothenberg, Thomas J. (1971). "Identification in Parametric Models". In: *Econometrica* 39.3, pp. 577–591 (cit. on pp. 9, 14).

Schmitt-Grohé, Stephanie and Martin Uribe (2004). "Solving dynamic general equilibrium models using a second-order approximation to the policy function". In: *Journal of Economic Dynamics and Control* 28.4, pp. 755–775 (cit. on pp. 1, 2, 5, 38).

Schmitt-Grohé, Stephanie and Martin Uribe (2012). "What's News In Business Cycles". In: *Econometrica* 80.6 (cit. on pp. 7, 8).

Sims, Christopher A. (2001). "Solving Linear Rational Expectations Models". In: *Computational Economics* 20.1-2, pp. 1–20 (cit. on p. 3).

Smets, Frank and Rafael Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach". In: *American Economic Review* 97.3, pp. 586–606 (cit. on pp. 20, 21, 27).

Stock, James H., Jonathan H. Wright, and Motohiro Yogo (2002). "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments". In: *Journal of Business & Economic Statistics* 20.4, pp. 518–29 (cit. on p. 19).

Tkachenko, Denis and Zhongjun Qu (2012). *Frequency Domain Analysis of Medium Scale DSGE Models with Application to Smets and Wouters (2007)*. Working Paper Series WP2011-060. Version April 30, 2012. Boston University: Department of Economics. URL: `http://people.bu.edu/qu/dsge3/paper.pdf` (visited on 09/26/2012) (cit. on pp. 1, 14).

A Magnus-Neudecker definition of Hessian

Define the steady state as $\overline{z} := (\overline{x}', \overline{y}', \overline{x}', \overline{y}')' = (\overline{x}', \overline{y}')'$, then the Jacobian $\mathcal{D}f(\overline{z})$ and Hessian $\mathcal{H}f(\overline{z})$ of f evaluated at the steady-state are defined as:

$$f(\overline{z}) = \begin{pmatrix} f^1(\overline{z}) \\ \vdots \\ f^n(\overline{z}) \end{pmatrix}$$

$$\mathcal{D}f(\overline{z}) := \begin{pmatrix} \frac{\partial f(\overline{z})}{\partial x'_{t+1}} & \frac{\partial f(\overline{z})}{\partial y'_{t+1}} & \frac{\partial f(\overline{z})}{\partial x'_t} & \frac{\partial f(\overline{z})}{\partial y'_t} \end{pmatrix} = \begin{pmatrix} \mathcal{D}f^1(\overline{z}) \\ \vdots \\ \mathcal{D}f^n(\overline{z}) \end{pmatrix} = \begin{pmatrix} \frac{\partial f^1(\overline{z})}{\partial x'_{t+1}} & \frac{\partial f^1(\overline{z})}{\partial y'_{t+1}} & \frac{\partial f^1(\overline{z})}{\partial x'_t} & \frac{\partial f^1(\overline{z})}{\partial y'_t} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f^n(\overline{z})}{\partial x'_{t+1}} & \frac{\partial f^n(\overline{z})}{\partial y'_{t+1}} & \frac{\partial f^n(\overline{z})}{\partial x'_t} & \frac{\partial f^n(\overline{z})}{\partial y'_t} \end{pmatrix}$$

$$\mathcal{H}f(\overline{z}) := \mathcal{D}vec((\mathcal{D}f(\overline{z}))') = \begin{pmatrix} \mathcal{H}f^1(\overline{z}) \\ \vdots \\ \mathcal{H}f^n(\overline{z}) \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 f^1(\overline{z})}{\partial x_{t+1} \partial x_{t+1}'} & \frac{\partial^2 f^1(\overline{z})}{\partial x_{t+1} \partial y_{t+1}'} & \frac{\partial^2 f^1(\overline{z})}{\partial x_{t+1} \partial x_t'} & \frac{\partial^2 f^1(\overline{z})}{\partial x_{t+1} \partial y_t'} \\ \frac{\partial^2 f^1(\overline{z})}{\partial y_{t+1} \partial x_{t+1}'} & \frac{\partial^2 f^1(\overline{z})}{\partial y_{t+1} \partial y_{t+1}'} & \frac{\partial^2 f^1(\overline{z})}{\partial y_{t+1} \partial x_t'} & \frac{\partial^2 f^1(\overline{z})}{\partial y_{t+1} \partial y_t'} \\ \frac{\partial^2 f^1(\overline{z})}{\partial x_t \partial x_{t+1}'} & \frac{\partial^2 f^1(\overline{z})}{\partial x_t \partial y_{t+1}'} & \frac{\partial^2 f^1(\overline{z})}{\partial x_t \partial x_t'} & \frac{\partial^2 f^1(\overline{z})}{\partial x_t \partial y_t'} \\ \frac{\partial^2 f^1(\overline{z})}{\partial y_t \partial x_{t+1}'} & \frac{\partial^2 f^1(\overline{z})}{\partial y_t \partial y_{t+1}'} & \frac{\partial^2 f^1(\overline{z})}{\partial y_t \partial x_t'} & \frac{\partial^2 f^1(\overline{z})}{\partial y_t \partial y_t'} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 f^n(\overline{z})}{\partial x_{t+1} \partial x_{t+1}'} & \frac{\partial^2 f^n(\overline{z})}{\partial x_{t+1} \partial y_{t+1}'} & \frac{\partial^2 f^n(\overline{z})}{\partial x_{t+1} \partial x_t'} & \frac{\partial^2 f^n(\overline{z})}{\partial x_{t+1} \partial y_t'} \\ \frac{\partial^2 f^n(\overline{z})}{\partial y_{t+1} \partial x_{t+1}'} & \frac{\partial^2 f^n(\overline{z})}{\partial y_{t+1} \partial y_{t+1}'} & \frac{\partial^2 f^n(\overline{z})}{\partial y_{t+1} \partial x_t'} & \frac{\partial^2 f^n(\overline{z})}{\partial y_{t+1} \partial y_t'} \\ \frac{\partial^2 f^n(\overline{z})}{\partial x_t \partial x_{t+1}'} & \frac{\partial^2 f^n(\overline{z})}{\partial x_t \partial y_{t+1}'} & \frac{\partial^2 f^n(\overline{z})}{\partial x_t \partial x_t'} & \frac{\partial^2 f^n(\overline{z})}{\partial x_t \partial y_t'} \\ \frac{\partial^2 f^n(\overline{z})}{\partial y_t \partial x_{t+1}'} & \frac{\partial^2 f^n(\overline{z})}{\partial y_t \partial y_{t+1}'} & \frac{\partial^2 f^n(\overline{z})}{\partial y_t \partial x_t'} & \frac{\partial^2 f^n(\overline{z})}{\partial y_t \partial y_t'} \end{pmatrix}.$$

f is of dimension $n \times 1$, the Jacobian $Df(\overline{z})$ of dimension $n \times (2n_x + 2n_y)$ and the Hessian $Hf(\overline{z})$ of dimension $n(2n_x + 2n_y) \times (2n_x + 2n_y)$.

B Deriving analytical derivatives

In order to calculate the derivatives of the solution matrices, we will use repeatedly the commutation matrix $K_{m,n}$ which transforms the $m \times n$ matrix A such that $K_{m,n}vec(A) = vec(A')$, and the following useful results from matrix differential calculus:

Theorem 1 (Derivative of products). *Let A be a $(m \times n)$ matrix, B a $(n \times o)$ matrix, C a*

$(o \times p)$ matrix and D a $(p \times q)$ matrix, then the derivative of $vec(ABCD)$ with respect to θ is given by

$$d(ABCD) = (D'C'B' \otimes I_m)dA + (D'C' \otimes A)dB + (D' \otimes AB)dC + (I_q \otimes ABC)dD$$

Proof: Magnus and Neudecker (1999, p. 175). Note that $dX := \frac{\partial vec(X)}{\partial \theta'}$.

Theorem 2 (Derivative of Kronecker products). *Let X be a $(n \times q)$ matrix, Y a $(p \times r)$ matrix and $K_{r,n}$ the commutation matrix of order (r, n), then the derivative of $vec(X \otimes Y)$ with respect to θ is given by*

$$d(X \otimes Y) = (I_q \otimes K_{r,n} \otimes I_p)\left[(I_{nq} \otimes vec(Y))dX + (vec(X) \otimes I_{pr})dY\right]$$

Proof: Magnus and Neudecker (1999, p. 185). Note that $dX := \frac{\partial vec(X)}{\partial \theta'}$.

Moreover, we will make use of the following algorithm:

Algorithm 1 (Derivative of partitioned matrix). *Let X be a $(m \times n)$ matrix, that is partitioned such that $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$, with X_1 being $(m \times n_1)$ and X_2 being $(m \times n_2)$, $n = n_1 + n_2$.*

1. *Derive dX_1 and dX_2; dX_1 is of dimension $(mn_1 \times n_\theta)$ and dX_2 of dimension $(mn_2 \times n_\theta)$.*
2. *For $i = 1, \ldots, n_\theta$*
   (a) *Denote the i-th column of dX_1 and dX_2 as dX_1^i and dX_2^i respectively. dX_1^i is of dimension $(mn_1 \times 1)$ and dX_2^i of dimension $(mn_2 \times 1)$.*
   (b) *Reshape dX_1^i into a $(m \times n_1)$ matrix $[dX_1^i]_{(m \times n_1)}$ and dX_2^i into a $(m \times n_2)$ matrix $[dX_2^i]_{(m \times n_2)}$.*
   (c) *Store $vec(\begin{bmatrix}[dX_1^i]_{(m \times n_1)} & [dX_2^i]_{(m \times n_2)}\end{bmatrix})$ into the i-th column of a matrix dX.*
3. *dX is the derivative of X with respect to θ and is of dimension $(mn \times n_\theta)$.*

Note that $dX := \frac{\partial vec(X)}{\partial \theta'}$.

Now we are able to derive the expressions for Q^{-1}, A, B^{-1} and C:

Derivative of $\mathbf{Q^{-1}}$ Notice that Q is partitioned into $Q = [Q_1\ Q_2]$,

$$Q_1 = h_x' \otimes f_2 \otimes h_x' + I_{n_x} \otimes f_4 \otimes I_{n_x},$$
$$Q_2 = I_{n_x} \otimes (f_1 + f_2 g_x) \otimes I_{n_x}.$$

Deriving $d(f_2 g_x)$ using theorem 1 and mechanically applying theorem 2 repeatedly, we obtain the derivatives dQ_1 and dQ_2. Now we can use algorithm 1 to compute dQ. However, we are interested in dQ^{-1}, thus in step 2(b) we also compute the derivative of the inverse using $-Q^{-1}\begin{bmatrix}[dQ_1^i] & [dQ_2^i]\end{bmatrix}Q^{-1}$ (Magnus and Neudecker 1999, p. 184) and store it in step 2(c) in the i-th column of dQ^{-1}.

Derivative of A Regarding the derivative of A we first have to derive dM. This can be done in the same fashion, since M is partitioned into $M = (h_x, g_x h_x, I_{n_x}, g_x)'$. dh_x and dg_x are known, whereas $d(g_x h_x)$ can be derived using theorem 1. Applying algorithm 1 we get dM, whereas for the transpose we have the following relationship $dM' = K_{2(n_x+n_y),n_x} dM$. Now we are able to compute the derivative of A using theorems 1 and 2.

Derivative of $\mathbf{B^{-1}}$ Since B is similarly partitioned as Q, i.e. $B = [B_1\ B_2]$, the derivative dB^{-1} can be calculated analogously to dQ^{-1}.

Derivative of C C is the sum of two matrices, for which we will derive the derivatives separately. Consider the first part, $f_2 \cdot trm[(I_{n_y} \otimes (\eta_x \eta_x{}'))g_{xx}]$. Since the derivatives of $(\eta_x \eta_x{}')$ and g_{xx} are known, it is straightforward to compute $d((I_{n_y} \otimes (\eta_x \eta_x{}'))g_{xx})$ applying theorems 1 and 2. The only slightly difficult part is the matrix trace function. However, algorithm 1 can be used to overcome this difficulty. In fact, we only have one partition, for which we know the derivative. Now taking the trm of the reshaped matrix in step 2(b) and storing this in 2(c), we get $d(trm[(I_{n_y} \otimes (\eta_x \eta_x{}'))g_{xx}])$. Theorem 1 then yields the derivative of $f_2 \cdot trm[(I_{n_y} \otimes (\eta_x \eta_x{}'))g_{xx}]$. The same steps can be used to derive the derivative of the second part, $trm[(I_{n_x+n_y} \otimes N')HN(\eta_x \eta_x{}')]$. However, we first have to derive an expression for dN and dN'. Since N is partitioned, we can use algorithm 1 to derive dN and $dN' = K_{2(n_x+n_y),n_x} dN$.

Derivative of $\mathbf{\Lambda_x}$ The first sum of Λ_x consists of two parts, one Kronecker product and the matrix $vec(hxx')$. First we apply theorem 2 to get the derivative of the Kronecker product. Then we notice that $d(vec(h'_{xx})) = d(h'_{xx}) = K_{n_x^2,n_x} dh_{xx}$. Thus, we are able to calculate the derivative of the first sum using theorem 1. The derivative of the second sum, $\sigma^2 h_{\sigma\sigma}$ is equal to $2\sigma h_{\sigma\sigma} d\sigma + \sigma^2 dh_{\sigma\sigma}$.

Derivative of $\mathbf{\Lambda_y}$ The procedure is fully analogous to the derivation of Λ_x, the only difference is using $K_{n_x n_y,n_x}$ to derive $d(vec(g'_{xx}))$.

Derivative of $\mathbf{\Sigma_x}$ Starting from $\Sigma_x = h_x \Sigma_x h'_x + \sigma^2 \eta_x \eta_x{}'$, applying theorem 1 and rearranging we get $d\Sigma_x = \left[I_{n_x^2} - (h_x \otimes h_x)\right]^{-1} \left[(h_x \Sigma_x \otimes I_{n_x}) dh_x + (I_{n_x} \otimes h_x \Sigma_x) dh'_x + d(\sigma^2 \eta_x \eta_x{}')\right]$.

Derivative of $\boldsymbol{H_\varepsilon(e^{-i\omega_s};\theta)}$ H_ε is given by two sums $Dg_x(I_{n_x} - h_x e^{-i\omega_s})^{-1} \sigma \eta_x$ and η_d. The derivatives of $\sigma\eta_x$ and η_d can be calculated symbolically given the model structure. Closed form expressions for dg_x and dh_x are given in chapter 2.4. Thus, we only need the derivative of the inverted expression which is given by

$$d\left((I_{n_x} - h_x e^{-i\omega_s})^{-1}\right) = \left(-(I_{n_x} - h_x e^{-i\omega_s})'^{-1} \otimes (I_{n_x} - h_x e^{-i\omega_s})^{-1}\right)(-dh_x \cdot e^{-i\omega_s})$$

where we used $d(X^{-1}) = (-(X')^{-1} \otimes X^{-1})dX$, see Magnus and Neudecker (1999, p. 184). Thus, computing dH_ε is a straightforward application of theorem 1. The derivative of the conjugate transpose is given by $dH^*_\varepsilon(e^{-i\omega_s};\theta) = K_{n_d,n_\varepsilon} conj(dH_\varepsilon(e^{-i\omega_s};\theta))$, where $conj$ returns the complex conjugate.

C Deriving numerical derivatives

In order to derive the Jacobian of a function or matrix $F(\theta)$ at a point θ_0 with respect to θ, we use a two-sided finite difference method (also known as central differences). That is:
For each $j = 1, \ldots, n_\theta$

(i) Select a step size h_j.

(ii) Solve the DSGE model twice using $\overline{\theta} = \theta_0 + e_j h_j$ and $\underline{\theta} = \theta_0 - e_j h_j$ with e_j a unit vector with the jth element equal to 1.

(iii) Compute

$$dF^j := \frac{\partial vec(F(\theta_0))}{\partial \theta_j} \approx vec\left(\frac{F(\theta_0 + e_j h_j) - F(\theta_0 - e_j h_j)}{2h_j}\right)$$

(iv) Store dF^j as the j-th column of dF.

D Deriving the minimal state

Given the linear solution (5) and (6) of the first order approximation, we will first derive the canonical $\mathcal{ABCD}$-representation of the DSGE model, i.e.

$$\begin{aligned} \mathrm{z_t} &= \overline{\mathrm{z}} + \mathcal{A}(\mathrm{z_{t-1}} - \overline{\mathrm{z}}) + \mathcal{B}e_t \\ d_t &= \overline{d} + \mathcal{C}(\mathrm{z_{t-1}} - \overline{\mathrm{z}}) + \mathcal{D}e_t \end{aligned}$$

with $\mathrm{z_t} = (y_t', x_t')'$ collecting all model variables and $e_t = \begin{pmatrix} \sigma\eta_u u_t \\ \eta_v v_t \end{pmatrix}$ with $E(e_t) = 0$ and $E(e_t e_t') = \begin{pmatrix} \Sigma_u & 0 \\ 0 & \Sigma_v \end{pmatrix} =: \Sigma_e$. We do this by appending auxiliary equations and variables for the shocks and measurement errors ($e_{t+1} = E_t \varepsilon_{t+1} = 0$), that is we consider the state vector $x_t = (x_{1,t-1}', x_{2,t-1}', e_t')'$ with $x_{1,t}$ denoting endogenous and $x_{2,t}$ exogenous states. Also we add the measurement equations to the model and consider the control vector $\widetilde{y}_t = (y_t', d_t')'$. The solution then becomes

$$\underbrace{\begin{pmatrix} x_{1,t} \\ {\scriptstyle n_{x_1}\times 1} \\ x_{2,t} \\ {\scriptstyle n_{x_2}\times 1} \\ e_{t+1} \\ {\scriptstyle n_\varepsilon \times 1} \end{pmatrix}}_{x_{t+1}} = \underbrace{\begin{pmatrix} \overline{x}_1 \\ {\scriptstyle n_{x_1}\times 1} \\ \overline{x}_2 \\ {\scriptstyle n_{x_2}\times 1} \\ 0 \\ {\scriptstyle n_\varepsilon \times 1} \end{pmatrix}}_{\overline{x}} + \underbrace{\begin{pmatrix} 0 & h_{x_1} & h_{e_1} \\ {\scriptstyle n_{x_1}\times n_{x_1}} & {\scriptstyle n_{x_1}\times n_{x_2}} & {\scriptstyle n_{x_1}\times n_\varepsilon} \\ 0 & h_{x_2} & h_{e_2} \\ {\scriptstyle n_{x_2}\times n_{x_1}} & {\scriptstyle n_{x_2}\times n_{x_2}} & {\scriptstyle n_{x_2}\times n_\varepsilon} \\ 0 & 0 & 0 \\ {\scriptstyle n_\varepsilon \times n_{x_1}} & {\scriptstyle n_\varepsilon \times n_{x_2}} & {\scriptstyle n_\varepsilon \times n_\varepsilon} \end{pmatrix}}_{h_x} \underbrace{\begin{pmatrix} x_{1,t-1} - \overline{x}_1 \\ {\scriptstyle n_{x_1}\times 1} \\ x_{2,t-2} - \overline{x}_2 \\ {\scriptstyle n_{x_2}\times 1} \\ e_t - 0 \\ {\scriptstyle n_\varepsilon \times 1} \end{pmatrix}}_{x_t - \overline{x}} + \sigma \underbrace{\begin{pmatrix} 0 \\ {\scriptstyle n_{x_1}\times n_\varepsilon} \\ 0 \\ {\scriptstyle n_{x_2}\times n_\varepsilon} \\ \begin{bmatrix} \eta_u & 0 \\ {\scriptstyle n_u\times n_u} & {\scriptstyle n_u\times n_v} \\ 0 & \eta_v/\sigma \\ {\scriptstyle n_v\times n_n} & {\scriptstyle n_v\times n_v} \end{bmatrix} \end{pmatrix}}_{\eta_x} \underbrace{\begin{pmatrix} u_{t+1} \\ {\scriptstyle n_u\times 1} \\ v_{t+1} \\ {\scriptstyle n_v\times 1} \end{pmatrix}}_{\varepsilon_{t+1}}$$

$$\underbrace{\begin{pmatrix} y_t \\ {\scriptstyle n_y\times 1} \\ d_t \\ {\scriptstyle n_d\times 1} \end{pmatrix}}_{\widetilde{y}_t} = \underbrace{\begin{pmatrix} \overline{y} \\ {\scriptstyle n_y\times 1} \\ \overline{d} \\ {\scriptstyle n_d\times 1} \end{pmatrix}}_{\overline{\widetilde{y}}} + \underbrace{\begin{pmatrix} g_{y_1} & g_{y_2} & g_{y_e} \\ {\scriptstyle n_y\times n_{x_1}} & {\scriptstyle n_y\times n_{x_2}} & {\scriptstyle n_y\times n_\varepsilon} \\ g_{d_1} & g_{d_2} & g_{d_e} \\ {\scriptstyle n_d\times n_{x_1}} & {\scriptstyle n_d\times n_{x_2}} & {\scriptstyle n_d\times n_\varepsilon} \end{pmatrix}}_{g_x} \underbrace{\begin{pmatrix} x_{1,t-1} - \overline{x}_1 \\ {\scriptstyle n_{x_1}\times 1} \\ x_{2,t-1} - \overline{x}_2 \\ {\scriptstyle n_{x_2}\times 1} \\ e_t - 0 \\ {\scriptstyle n_\varepsilon \times 1} \end{pmatrix}}_{x_t - \overline{x}}$$

Note that $Dg_x = \begin{pmatrix} g_{d_1} & g_{d_2} & g_{d_e} \end{pmatrix}$, since D picks the last n_d entries of $\widetilde{y}_t$. Given that $e_t = \begin{pmatrix} \sigma\eta_u u_t \\ \eta_v v_t \end{pmatrix}$, this can be rearranged to get

$$\underbrace{\begin{pmatrix} y_t \\ x_{1,t} \\ x_{2,t} \end{pmatrix}}_{z_t} = \underbrace{\begin{pmatrix} \overline{y} \\ \overline{x}_1 \\ \overline{x}_2 \end{pmatrix}}_{\overline{z}} + \underbrace{\begin{pmatrix} 0 & g_{y_1} & g_{y_2} \\ 0 & 0 & h_{x_1} \\ 0 & 0 & h_{x_2} \end{pmatrix}}_{\mathcal{A}} \underbrace{\begin{pmatrix} y_{t-1} - \overline{y} \\ x_{1,t-1} - \overline{x}_1 \\ x_{2,t-1} - \overline{x}_2 \end{pmatrix}}_{z_{t-1}-\overline{z}} + \underbrace{\begin{pmatrix} g_{y_e} \\ h_{e_1} \\ h_{e_2} \end{pmatrix}}_{\mathcal{B}} e_t$$

$$d_t = \overline{d} + \underbrace{\begin{pmatrix} 0 & g_{d_1} & g_{d_2} \end{pmatrix}}_{\mathcal{C}} \underbrace{\begin{pmatrix} y_{t-1} - \overline{y} \\ x_{1,t-1} - \overline{x}_1 \\ x_{2,t-1} - \overline{x}_2 \end{pmatrix}}_{z_{t-1}-\overline{z}} + \underbrace{g_{d_e}}_{\mathcal{D}} e_t$$

Obviously, the driving force of the model is the vector of exogenous states $x_{2,t}$, which we call the minimal state vector. Together with the evolution of the stochastic innovations ε_t it determines the evolution of the endogenous states, the control and the observable variables. The minimal representation is thus given by

$$x_{2,t} = \overline{x}_2 + \underbrace{h_{x_2}}_{\widetilde{\mathcal{A}}}(x_{2,t-1} - \overline{x}_2) + \underbrace{h_{e_2}}_{\widetilde{\mathcal{B}}} e_t$$

$$d_t = \overline{d} + \underbrace{g_{d_2}}_{\widetilde{\mathcal{C}}}(x_{2,t-1} - \overline{x}_2) + \underbrace{g_{d_e}}_{\widetilde{\mathcal{D}}} e_t$$

Some practical issues: For small and medium-sized DSGE models the distinction between endogenous and exogenous states is given through theory, some variables are clearly endogenous (like output) and some are clearly exogenous (like technology). Thus, we order the state vector as mentioned in the text. First, we check the rank conditions for minimality and observability given the full state vector. If the conditions fail, we remove state variables from the top until the conditions pass. Note that we remove the entries from all first-order solution matrices as well as from the derivatives corresponding to redundant state variables.

For big DSGE models the distinction of endogenous and exogenous states is often not as clear. A failsafe approach deriving the minimal state vector is to consider all possible subsets of combinations of the state vector and check the rank conditions for minimality and controllability in each case. For a different (computational) approach handling the minimal state in big DSGE models see Kim et al. (2008).

E Example Models

The neoclassical growth model

We will consider the simple neoclassical growth model (see e.g. Schmitt-Grohé and Uribe (2004)) and transform the model using $X_t = e^{log(X_t)} =: e^{x_t}$ in order to focus on log-deviations from

steady-state. Thus, f is given by the following eight equations.

$$
\begin{aligned}
0 &= -e^{y_t} + e^{a_{t-1}+\alpha k_{t-1}}, & 0 &= -c_t^{obs} + c_t + v_{c,t},\\
0 &= -e^{k_t} + (1-\delta)e^{k_{t-1}} + e^{y_t} - e^{c_t}, & 0 &= -y_t^{obs} + y_t,\\
0 &= -e^{-\gamma c_t} + \beta e^{-\gamma E_t c_{t+1}}(\alpha e^{a_t+(\alpha-1)k_t} + 1 - \delta), & 0 &= E_t u_{a,t+1},\\
0 &= -a_t + \rho a_{t-1} + u_{a,t}, & 0 &= E_t v_{c,t+1}.
\end{aligned}
$$

There are two exogenous states k_t and a_t, and no endogenous states. The controls are c_t and y_t, and the observables are c_t^{obs} and y_t^{obs}. There is one shock on technology $u_{a,t}$, and c_t^{obs} is observed with measurement error $v_{c,t}$. Thus, given our definition and ordering of variables we have

$$
\varepsilon_t = (u_{a,t}, v_{c,t})', \qquad x_t = (k_{t-1}, a_{t-1}, \varepsilon_t)', \qquad y_t = (c_t, y_t, c_t^{obs}, y_t^{obs})', \qquad D = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
$$

Further we set the perturbation parameter equal to the standard deviation of the shock on technology, then we have

$$
\sigma = \sigma_A, \qquad \eta_u = 1, \qquad \eta_v = \sigma_c, \qquad \eta_x = \begin{bmatrix} 0_{2\times 1} & 0_{2\times 1} \\ \eta_u & 0 \\ 0 & \eta_v/\sigma \end{bmatrix}, \qquad \eta_d = \begin{bmatrix} 0 & \eta_v \\ 0 & 0 \end{bmatrix}.
$$

The steady-state of this model is given by

$$
a = log[1], \quad k = log\left[\left(\frac{1/\beta + \delta - 1}{\alpha}\right)^{\frac{1}{\alpha-1}}\right], \quad c = log\left[e^{a+\alpha k} - \delta e^{k}\right], \quad y = a + \alpha k, \quad \varepsilon = 0.
$$

We will consider identification of the parameter vector θ at the local point θ_0:

$$
\begin{aligned}
\theta &= (\alpha, \quad \rho, \quad \beta, \quad \delta, \quad \gamma, \quad \sigma_a, \quad \sigma_c)',\\
\theta_0 &= (0.3, \quad 0.9, \quad 0.95, \quad 1, \quad 2, \quad 1, \quad 1)'.
\end{aligned}
$$

The Kim (2003) model

First we define an auxiliary parameter and variable:

$$
s = \frac{\beta\delta\alpha}{1 - \beta + \delta\beta}, \qquad \lambda_t = \frac{(1-s)^{\theta}}{(1+\theta)c_t^{1+\theta}}.
$$

Then the model is given by the following eight equations f:

$$
0 = -\lambda_t(1+\theta)\left(\frac{i_t}{s}\right)^{\theta}\left(\frac{i_t}{\delta k_t}\right)^{\phi} + \beta E_t\lambda_{t+1}\left[\alpha(1+\theta)a_t^{1+\theta}k_t^{\alpha(1+\theta)-1} + (1-\delta)(1+\theta)\left(\frac{E_t i_{t+1}}{\delta k_t}\right)^{\phi}\left(\frac{E_t i_{t+1}}{s}\right)^{\theta}\right]
$$

$$
0 = \left[(1-s)\left(\frac{c_t}{1-s}\right)^{1+\theta} + s\left(\frac{i_t}{s}\right)^{1+\theta}\right]^{\frac{1}{1+\theta}} - a_{t-1}k_{t-1}^{\alpha}, \qquad 0 = k_t - \left[\delta\left(\frac{i_t}{\delta}\right)^{1-\phi} + (1-\delta)\left(k_{t-1}\right)^{1-\phi}\right]^{\frac{1}{1-\phi}}
$$

$$
\begin{aligned}
0 &= -a_t + a_0 + u_{a,t}, & 0 &= E_t u_{a,t+1}, & 0 &= E_t v_{c,t+1},\\
0 &= -c_t^{obs} + c_t + v_{c,t}, & 0 &= -i_t^{obs} + i_t.
\end{aligned}
$$

There are two exogenous states k_t and a_t, and no endogenous states. The controls are c_t and i_t, and the observables are c_t^{obs} and i_t^{obs}. There is one shock on technology $u_{a,t}$, and c_t^{obs} is observed with measurement error $v_{c,t}$. Thus, given our definition and ordering of variables we have

$$\varepsilon_t = (u_{a,t}, v_{c,t})', \qquad x_t = (k_{t-1}, a_{t-1}, \varepsilon_t)', \qquad y_t = (c_t, i_t, c_t^{obs}, i_t^{obs})', \qquad D = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Further we set the perturbation parameter equal to the standard deviation of the shock on technology, then we have

$$\sigma = \sigma_A, \qquad \eta_u = 1, \qquad \eta_v = \sigma_c, \qquad \eta_x = \begin{bmatrix} 0_{2\times 1} & 0_{2\times 1} \\ \eta_u & 0 \\ 0 & \eta_v/\sigma \end{bmatrix}, \qquad \eta_d = \begin{bmatrix} 0 & \eta_v \\ 0 & 0 \end{bmatrix}.$$

The steady-state of this model is given by

$$a = a_0, \quad k = \left(\frac{\delta}{sa}\right)^{\frac{1}{\alpha-1}}, \quad i = \delta k, \quad c = (1-s)\left[\frac{(\alpha k^{\alpha})^{1+\theta} - s\left(\frac{i}{s}\right)^{1+\theta}}{1-s}\right]^{\frac{1}{1+\theta}}, \quad \varepsilon = 0.$$

We will consider identification of the parameter vector θ at the local point θ_0:

$$\begin{array}{rcccccccc} \theta &=& (\alpha, & \beta, & \delta, & a_0, & \theta, & \phi, & \sigma_a, & \sigma_c)', \\ \theta_0 &=& (0.6, & 0.99, & 0.0125, & 0.1, & 1, & 2, & 1, & 1)'. \end{array}$$

The An and Schorfheide (2007) model

First we define auxiliary parameters:

$$\beta = \frac{1}{1+\frac{r^{(A)}}{400}}, \qquad \pi^* = 1 + \frac{\pi^{(A)}}{400}, \qquad \phi = \frac{\tau(1-\nu)}{\nu\kappa\pi^{*2}}, \qquad g^* = \frac{1}{(c/y)^*}.$$

Then the model f consists of thirteen equations:

$$0 = \frac{1-\nu}{\nu\phi\pi^{*2}}\left(e^{\tau c_t} - 1\right) - \left(e^{\pi_t} - 1\right)\left[\left(1 - \frac{1}{2\nu}\right)e^{\pi_t} + \frac{1}{2\nu}\right] + \beta\left(e^{E_t\pi_{t+1}} - 1\right)e^{-\tau E_t c_{t+1} + \tau c_t + E_t dy_{t+1} + E_t\pi_{t+1}},$$

$$0 = 1 - e^{-\tau E_t c_{t+1} + \tau c_t + R_t - \rho_z z_t - E_t\pi_{t+1}}, \qquad 0 = e^{c_t - y_t} - e^{-g_t} + \frac{\phi\pi^{*2}g^*}{2}\left(e^{\pi_t} - 1\right)^2,$$

$$0 = R_t - \rho_R R_{t-1} - (1-\rho_R)\psi_1\pi_t - (1-\rho_R)\psi_2\left(y_t - g_t\right) - u_{R,t},$$

$$0 = dy_t - y_t + y_{t-1}, \qquad 0 = g_t - \rho_g g_{t-1} - u_{g,t}, \qquad 0 = z_t - \rho_z z_{t-1} - u_{z,t},$$

$$0 = YGR_t - \gamma^{(Q)} - 100(dy_t + z_t), \qquad 0 = INFL_t - \pi^{(A)} - 400\pi_t,$$

$$0 = INT_t - \pi^{(A)} - r^{(A)} - 4\gamma^{(Q)} - 400R_t,$$

$$0 = E_t u_{R,t+1}, \qquad 0 = E_t u_{g,t+1}, \qquad 0 = E_t u_{z,t+1}.$$

There are three exogenous states R_t, g_t and z_t, and one endogenous state variable y_t. The controls are c_t, dy_t and π_t, and the observables are $YGR_t, INFL_t$ and INT_t. There are three shocks: a monetary $u_{R,t}$, a fiscal $u_{g,t}$ and a technological shock $u_{z,t}$. Further, there are no measurement

errors in the model. Thus, given our definition and ordering of variables we have

$$\varepsilon_t = (u_{R,t}, u_{g,t}, u_{z,t})', \qquad x_t = (y_{t-1}, R_{t-1}, g_{t-1}, z_{t-1}, \varepsilon_t)', \qquad D = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$
$$y_t = (c_t, dy_t, \pi_t, YGR_t, INFL_t, INT_t)',$$

Further we set the perturbation parameter equal to the standard deviation of the shock on technology, then we have

$$\sigma = \sigma_z, \qquad \eta_u = \begin{bmatrix} \sigma_R/\sigma_z & 0 & 0 \\ 0 & \sigma_g/\sigma_z & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \eta_v = [\,], \qquad \eta_x = \begin{bmatrix} 0_{4\times 3} \\ \eta_u \end{bmatrix}, \qquad \eta_d = 0_{3\times 3}.$$

The steady-state of this model is given by

$$y = R = g = z = \varepsilon = c = dy = \pi = 0,$$
$$YGR = \gamma^{(Q)}, \quad INFL = \pi^{(A)}, \quad INT = \pi^{(A)} + r^{(A)} + 4\gamma^{(Q)}.$$

We will consider identification of the parameter vector θ at the local point θ_0:

θ =	$(\tau,$	$\kappa,$	$\psi_1,$	$\psi_2,$	$\rho_R,$	$\rho_g,$	$\rho_z,$	$r^{(A)},$	$\pi^{(A)},$	$\gamma^{(Q)},$	$\sigma_R,$	$\sigma_g,$	$\sigma_z,$	$\nu,$	$(c/y)^*)'$
θ_0 =	(2,	0.33,	1.5,	0.125,	0.75,	0.95,	0.9,	1,	3.2,	0.55,	0.002,	0.006,	0.003,	0.1,	$0.85)'$.