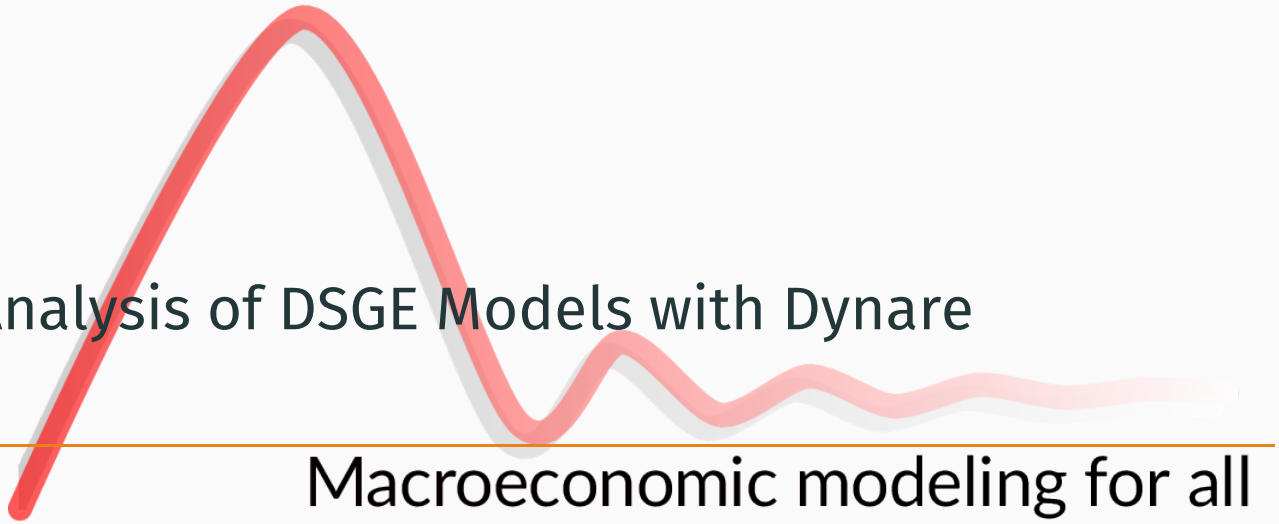
Identification Analysis of DSGE Models with Dynare



Macroeconomic modeling for all

Willi Mutschler & Marco Ratto

Dynare

Slides and source code available here

Table of contents

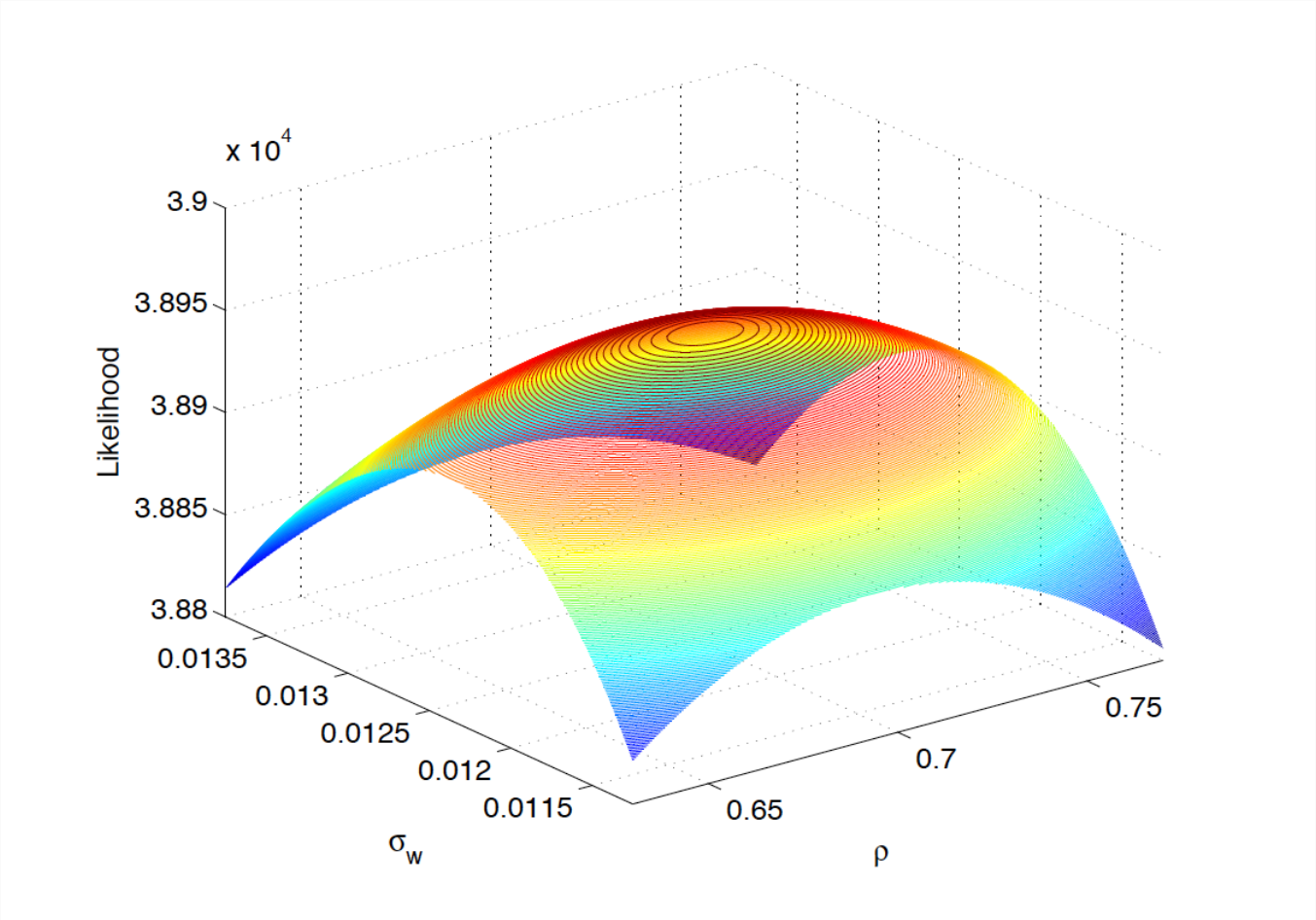
Introduction

Introduction

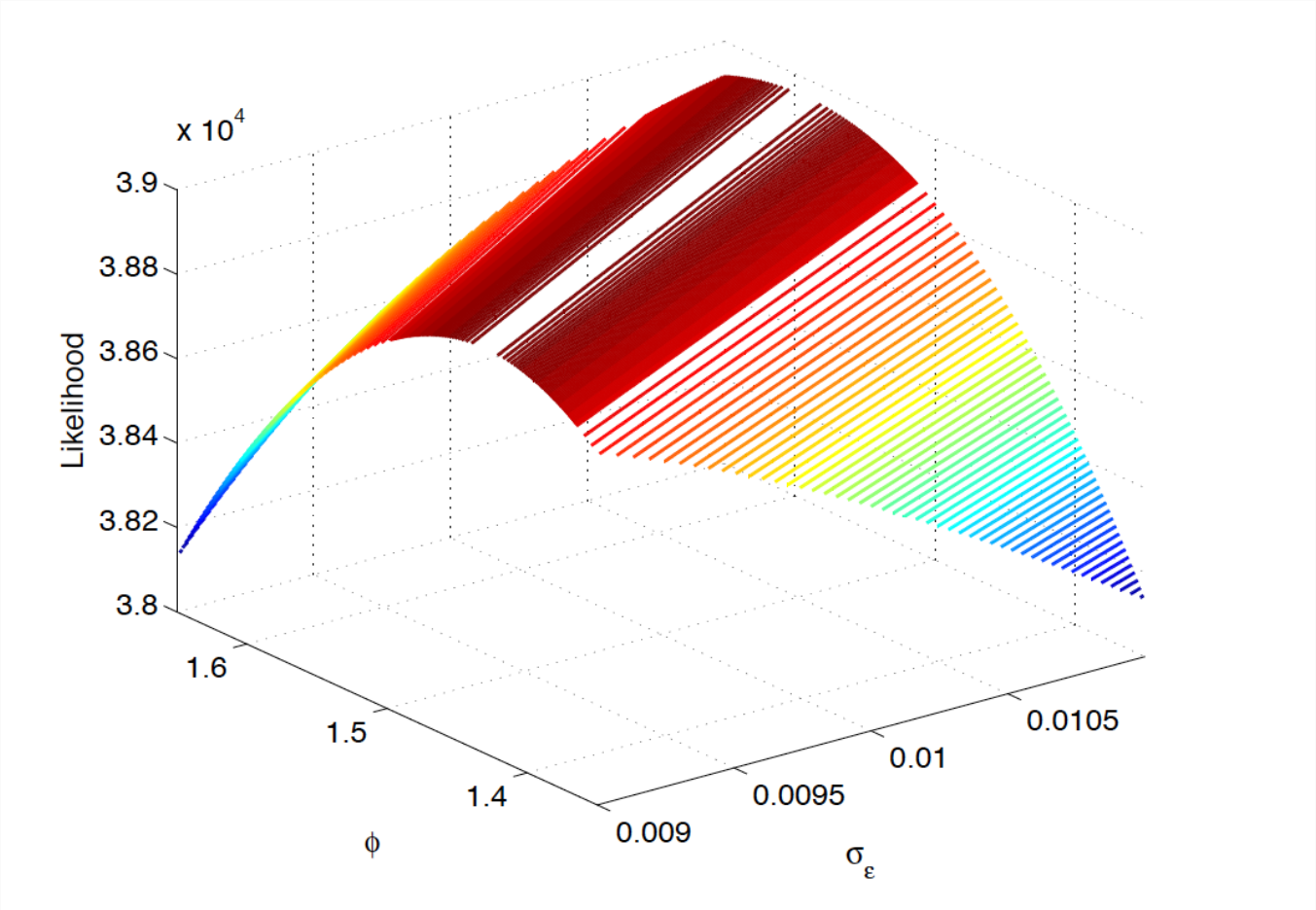
- Bayesian (and to some extent also Frequentist) estimation of DSGE models has rapidly progressed
- Study of **identifiability of parameters**, which should precede estimation (think about *regularity conditions*), is still a rather neglected topic in applied macro
- BUT: Parameter identification is a **model property** and can be readily assessed on a case-by-case basis before taking your model to data
- This talk:
  - Present well-established diagnostics and indicators to detect identification failures
  - Showcase the identification toolbox in Dynare
  - Provide applied tips on how to solve theoretical identification failures and improve the strength of DSGE model parameter identification

Introduction

Example (1): Likelihood shape



The likelihood function shows a clear peak in ρ and σ_{wt}.



The likelihood function has no unique maximum in σ_ε and ϕ, only a range of maxima

Introduction

Example (2): ARMA(1,1)

ARMA(1,1)

$$x_t - \phi_1 x_{t-1} = \varepsilon_t - \phi_2 \varepsilon_{t-1}, \text{ with } \varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$$

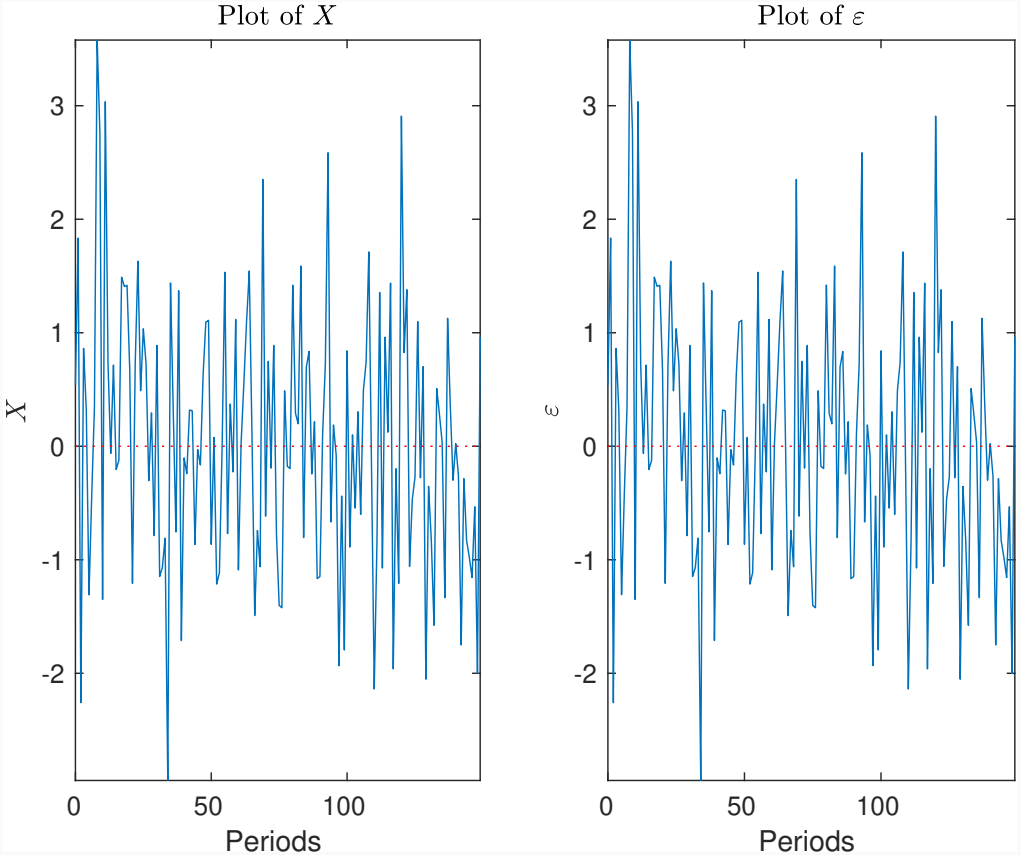


Figure 1: $(\phi_1, \phi_2, \sigma) = (0.4, 0.4, 1)$ [see `arma.mod`]

Autocovariance function

- Define $\gamma_j = E[x_t x_{t-j}]$, then for ARMA(1,1):

$$\gamma_0 = \frac{(1+\phi_2^2 - 2\phi_1\phi_2)\sigma^2}{1-\phi_1^2}, \gamma_1 = \frac{(\phi_1-\phi_2)(1-\phi_1\phi_2)\sigma^2}{1-\phi_1^2}, \gamma_h = \phi_1\gamma_{h-1}$$

- White noise: $(\phi_1 = \phi_2 = 0)$:

$$\gamma_0 = \sigma^2, \gamma_h = 0$$

- Equal coefficients $(\phi_1 = \phi_2)$:

$$\gamma_0 = \sigma^2, \gamma_h = 0$$

- Observational equivalent processes: θ_1 and θ_2 are not jointly identifiable

Introduction

Example (3): Forward-looking Model

Simple forward-looking model

$$r_t = \psi \pi_t + \varepsilon_t^M \qquad (TR)$$

$$x_t = E_t x_{t+1} - \frac{1}{\tau}(r_t - E_t \pi_{t+1}) + \varepsilon_t^D \qquad (IS)$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t + \varepsilon_t^S \qquad (PC)$$

Simple forward-looking model

$$r_t = \psi \pi_t + \varepsilon_t^M \qquad (TR)$$

$$x_t = E_t x_{t+1} - \frac{1}{\tau}(r_t - E_t \pi_{t+1}) + \varepsilon_t^D \qquad (IS)$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t + \varepsilon_t^S \qquad (PC)$$

or

$$\underbrace{\begin{pmatrix} 1 & 0 & -\psi \\ \frac{1}{\tau} & 1 & 0 \\ 0 & -\kappa & 1 \end{pmatrix}}_{A_0} \underbrace{\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix}}_{y_t} = \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & \frac{1}{\tau} \\ 0 & 0 & \beta \end{pmatrix}}_{A_1} \underbrace{\begin{pmatrix} E_t r_{t+1} \\ E_t x_{t+1} \\ E_t \pi_{t+1} \end{pmatrix}}_{E_t y_{t+1}} + \underbrace{\begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}}_{\varepsilon_t}$$

Stationary solution implies Eigenvalues of $A_0^{-1} A_1$ lie within unit circle:

Simple forward-looking model

$$r_t = \psi \pi_t + \varepsilon_t^M \qquad (TR)$$

$$x_t = E_t x_{t+1} - \frac{1}{\tau}(r_t - E_t \pi_{t+1}) + \varepsilon_t^D \qquad (IS)$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t + \varepsilon_t^S \qquad (PC)$$

or

$$\underbrace{\begin{pmatrix} 1 & 0 & -\psi \\ \frac{1}{\tau} & 1 & 0 \\ 0 & -\kappa & 1 \end{pmatrix}}_{A_0} \underbrace{\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix}}_{y_t} = \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & \frac{1}{\tau} \\ 0 & 0 & \beta \end{pmatrix}}_{A_1} \underbrace{\begin{pmatrix} E_t r_{t+1} \\ E_t x_{t+1} \\ E_t \pi_{t+1} \end{pmatrix}}_{E_t y_{t+1}} + \underbrace{\begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}}_{\varepsilon_t}$$

Stationary solution implies Eigenvalues of $A_0^{-1}A_1$ lie within unit circle:

$$y_t = A_0^{-1}A_1 E_t y_{t+1} + A_0^{-1}\varepsilon_t = \sum_{j=0}^{\infty} (A_0^{-1}A_1)^j A_0^{-1} E_t \varepsilon_{t+j} = A_0^{-1}\varepsilon_t$$

Simple forward-looking model

$$r_t = \psi \pi_t + \varepsilon_t^M \qquad (TR)$$

$$x_t = E_t x_{t+1} - \frac{1}{\tau}(r_t - E_t \pi_{t+1}) + \varepsilon_t^D \qquad (IS)$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t + \varepsilon_t^S \qquad (PC)$$

or

$$\underbrace{\begin{pmatrix} 1 & 0 & -\psi \\ \frac{1}{\tau} & 1 & 0 \\ 0 & -\kappa & 1 \end{pmatrix}}_{A_0} \underbrace{\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix}}_{y_t} = \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & \frac{1}{\tau} \\ 0 & 0 & \beta \end{pmatrix}}_{A_1} \underbrace{\begin{pmatrix} E_t r_{t+1} \\ E_t x_{t+1} \\ E_t \pi_{t+1} \end{pmatrix}}_{E_t y_{t+1}} + \underbrace{\begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}}_{\varepsilon_t}$$

Stationary solution implies Eigenvalues of $A_0^{-1}A_1$ lie within unit circle:

$$y_t = A_0^{-1}A_1 E_t y_{t+1} + A_0^{-1}\varepsilon_t = \sum_{j=0}^{\infty} (A_0^{-1}A_1)^j A_0^{-1} E_t \varepsilon_{t+j} = A_0^{-1}\varepsilon_t$$

[see `forward_looking.mod`]

Simple forward-looking model

Solution/data-generating-process/reduced-form

$$\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix} = \underbrace{\frac{1}{\frac{\kappa\psi}{\tau}+1}\begin{pmatrix} 1 & \kappa\psi & \psi \\ -\frac{1}{\tau} & 1 & -\frac{\psi}{\tau} \\ -\frac{\kappa}{\tau} & \kappa & 1 \end{pmatrix}}_{A_0^{-1}} \begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}$$

Some insights

Simple forward-looking model

Solution/data-generating-process/reduced-form

$$\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix} = \underbrace{\frac{1}{\frac{\kappa\psi}{\tau}+1}\begin{pmatrix} 1 & \kappa\psi & \psi \\ -\frac{1}{\tau} & 1 & -\frac{\psi}{\tau} \\ -\frac{\kappa}{\tau} & \kappa & 1 \end{pmatrix}}_{A_0^{-1}} \begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}$$

Some insights

- Some parameters (β) do not enter solution; thus, do not enter the likelihood (or any other objective)

Simple forward-looking model

Solution/data-generating-process/reduced-form

$$\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix} = \underbrace{\frac{1}{\frac{\kappa\psi}{\tau}+1}\begin{pmatrix} 1 & \kappa\psi & \psi \\ -\frac{1}{\tau} & 1 & -\frac{\psi}{\tau} \\ -\frac{\kappa}{\tau} & \kappa & 1 \end{pmatrix}}_{A_0^{-1}} \begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}$$

Some insights

- Some parameters (β) do not enter solution; thus, do not enter the likelihood (or any other objective)
- Identification depends on observables, e.g. when observing only x_t

Simple forward-looking model

Solution/data-generating-process/reduced-form

$$\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix} = \underbrace{\frac{1}{\frac{\kappa\psi}{\tau} + 1} \begin{pmatrix} 1 & \kappa\psi & \psi \\ -\frac{1}{\tau} & 1 & -\frac{\psi}{\tau} \\ -\frac{\kappa}{\tau} & \kappa & 1 \end{pmatrix}}_{A_0^{-1}} \begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}$$

Some insights

- Some parameters (β) do not enter solution; thus, do not enter the likelihood (or any other objective)
- Identification depends on observables, e.g. when observing only x_t
  - τ and ψ are pairwise collinear

Simple forward-looking model

Solution/data-generating-process/reduced-form

$$\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix} = \underbrace{\frac{1}{\frac{\kappa\psi}{\tau}+1}\begin{pmatrix} 1 & \kappa\psi & \psi \\ -\frac{1}{\tau} & 1 & -\frac{\psi}{\tau} \\ -\frac{\kappa}{\tau} & \kappa & 1 \end{pmatrix}}_{A_0^{-1}} \begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}$$

Some insights

- Some parameters (β) do not enter solution; thus, do not enter the likelihood (or any other objective)
- Identification depends on observables, e.g. when observing only x_t
  - τ and ψ are pairwise collinear
  - κ and ψ are pairwise collinear

Simple forward-looking model

Solution/data-generating-process/reduced-form

$$\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix} = \underbrace{\frac{1}{\frac{\kappa\psi}{\tau}+1}\begin{pmatrix} 1 & \kappa\psi & \psi \\ -\frac{1}{\tau} & 1 & -\frac{\psi}{\tau} \\ -\frac{\kappa}{\tau} & \kappa & 1 \end{pmatrix}}_{A_0^{-1}} \begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}$$

Some insights

- Some parameters (β) do not enter solution; thus, do not enter the likelihood (or any other objective)
- Identification depends on observables, e.g. when observing only x_t
  - τ and ψ are pairwise collinear
  - κ and ψ are pairwise collinear
  - We would need to fix ψ

Simple forward-looking model

Solution/data-generating-process/reduced-form

$$\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix} = \underbrace{\frac{1}{\frac{\kappa\psi}{\tau}+1}\begin{pmatrix} 1 & \kappa\psi & \psi \\ -\frac{1}{\tau} & 1 & -\frac{\psi}{\tau} \\ -\frac{\kappa}{\tau} & \kappa & 1 \end{pmatrix}}_{A_0^{-1}} \begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}$$

Some insights

- Some parameters (β) do not enter solution; thus, do not enter the likelihood (or any other objective)
- Identification depends on observables, e.g. when observing only x_t
  - τ and ψ are pairwise collinear
  - κ and ψ are pairwise collinear
  - We would need to fix ψ
- κ is already the product of several other structural parameters (Calvo or Rotemberg)

Simple forward-looking model

Solution/data-generating-process/reduced-form

$$\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix} = \underbrace{\frac{1}{\frac{\kappa\psi}{\tau}+1}\begin{pmatrix} 1 & \kappa\psi & \psi \\ -\frac{1}{\tau} & 1 & -\frac{\psi}{\tau} \\ -\frac{\kappa}{\tau} & \kappa & 1 \end{pmatrix}}_{A_0^{-1}} \begin{pmatrix} \varepsilon_t^M \\ \varepsilon_t^D \\ \varepsilon_t^S \end{pmatrix}$$

Some insights

- Some parameters (β) do not enter solution; thus, do not enter the likelihood (or any other objective)
- Identification depends on observables, e.g. when observing only x_t
  - τ and ψ are pairwise collinear
  - κ and ψ are pairwise collinear
  - We would need to fix ψ
- κ is already the product of several other structural parameters (Calvo or Rotemberg)
- Restrictions necessary to ensure regularity (Eigenvalues inside unit circle) imply bounds involving all parameters, i.e. parameter space is not *variation free*

Introduction

Example (4): Investment Adjustment Costs

Kim (2003): RBC model with two types of investment adjustment costs

Log utility, no labor, Cobb-Douglas production function, AR(1) technology

$$
\begin{aligned}
U_t &= \ln\left(C_t\right) \\
Y_t &= A_t K_{t-1}^{\alpha} \\
\ln\left(A_t\right) &= \rho_A \ln\left(A_{t-1}\right) + \varepsilon_t^A
\end{aligned}
$$

Kim (2003): RBC model with two types of investment adjustment costs

Log utility, no labor, Cobb-Douglas production function, AR(1) technology

$$U_t = \ln(C_t)$$
$$Y_t = A_t K_{t-1}^{\alpha}$$
$$\ln(A_t) = \rho_A \ln(A_{t-1}) + \varepsilon_t^A$$

Aggregate demand with multisectoral adjustment costs governed by θ

$$Y_t^d = \left[(1 - SAV)\left(\frac{C_t}{1 - SAV}\right)^{1+\theta} + SAV\left(\frac{I_t}{SAV}\right)^{1+\theta}\right]^{\frac{1}{1+\theta}}, \text{with } SAV = \frac{\bar{I}}{\bar{Y}^d}$$

Kim (2003): RBC model with two types of investment adjustment costs

Log utility, no labor, Cobb-Douglas production function, AR(1) technology

$$U_t = \ln(C_t)$$
$$Y_t = A_t K_{t-1}^{\alpha}$$
$$\ln(A_t) = \rho_A \ln(A_{t-1}) + \varepsilon_t^A$$

Aggregate demand with multisectoral adjustment costs governed by θ

$$Y_t^d = \left[(1 - SAV)\left(\frac{C_t}{1 - SAV}\right)^{1+\theta} + SAV\left(\frac{I_t}{SAV}\right)^{1+\theta}\right]^{\frac{1}{1+\theta}}, \text{with } SAV = \frac{\bar{I}}{\bar{Y}^d}$$

Capital accumulation with intertemporal adjustment costs governed by κ

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

Kim (2003): RBC model with two types of investment adjustment costs

Log utility, no labor, Cobb-Douglas production function, AR(1) technology

$$U_t = \ln(C_t)$$
$$Y_t = A_t K_{t-1}^{\alpha}$$
$$\ln(A_t) = \rho_A \ln(A_{t-1}) + \varepsilon_t^A$$

Aggregate demand with multisectoral adjustment costs governed by θ

$$Y_t^d = \left[(1-SAV)\left(\frac{C_t}{1-SAV}\right)^{1+\theta} + SAV\left(\frac{I_t}{SAV}\right)^{1+\theta}\right]^{\frac{1}{1+\theta}}, \text{with } SAV = \frac{\bar{I}}{\bar{Y}^d}$$

Capital accumulation with intertemporal adjustment costs governed by κ

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

[see `investadjust.mod`]

Baseline parametrization: $\theta = 1.5$ and $\kappa = 2$

THEORETICAL MOMENTS

VARIABLE	MEAN	STD. DEV.	VARIANCE
y	2.6827	1.8723	3.5054
yd	2.6827	1.8723	3.5054
c	2.0120	1.5077	2.2731
iv	0.6707	0.3649	0.1332
rk	0.0300	0.0207	0.0004
k	26.8270	3.5612	12.6824
lam	0.4970	0.4108	0.1688
q	1.0000	0.8094	0.6551
a	1.0000	0.6928	0.4800

No multisectoral costs: $\theta = 0$ and $\kappa = 1.4$

THEORETICAL MOMENTS

VARIABLE	MEAN	STD. DEV.	VARIANCE
y	2.6827	1.8723	3.5054
yd	2.6827	1.8723	3.5054
c	2.0120	1.5077	2.2731
iv	0.6707	0.3649	0.1332
rk	0.0300	0.0207	0.0004
k	26.8270	3.5612	12.6824
lam	0.4970	0.3724	0.1387
q	1.0000	0.7295	0.5322
a	1.0000	0.6928	0.4800

No intertemporal costs: $\theta = -3.5$ and $\kappa = 0$

THEORETICAL MOMENTS

VARIABLE	MEAN	STD. DEV.	VARIANCE
y	2.6827	1.8723	3.5054
yd	2.6827	1.8723	3.5054
c	2.0120	1.5077	2.2731
iv	0.6707	0.3649	0.1332
rk	0.0300	0.0207	0.0004
k	26.8270	3.5612	12.6824
lam	0.4970	0.2832	0.0802
q	1.0000	0.5448	0.2968
a	1.0000	0.6928	0.4800

Insights

- At first-order, DSGE solution is linear Gaussian state-space system
  - First two moments characterize whole distribution (moments, impulse responses, variance decomposition, etc)
- Different parameters yield same theoretical moments for y, c, iv, rk and a; and therefore also observational equivalent models
  - Even with an infinite sample size, no way to uniquely recover true model structure
- Observing lam or q might work, but unrealistic

Identification problem

Identification problem

Theory and Practice

Identification Problem in Theory

1. **Observational Equivalence**: mapping between structural parameters and objective function has no unique maximum
   - structural models with potentially different economic interpretations may be indistinguishable
2. **Under-identification**: objective function is independent of certain structural parameters, e.g. because they disappear from rational expectations solution
3. **Partial identification** with two or more structural parameters entering objective function only proportionally, making them separately unrecoverable
4. **Weak identification**: parameter theoretically identified, but curvature may be small in certain regions of the parameter space

Identification Problem in Practice

- Lack of identification leads to wrong conclusions from calibration, estimation and inference
- Many caveats due to identifiability issues and/or an unfortunate choice of observables
  - Difficult to maximize likelihood/posterior or minimize some (moment) objective function
  - Estimators often lie on the boundary of theoretically admissible space
  - Gaussian asymptotic theory yields poor approximations
- "Unidentifiability causes no real difficulties in the Bayesian approach" (Lindley, 1971, p. 46) is misleading, because if parameters are not identifiable
  - prior becomes extremely influential and needs to be informative for a proper posterior
  - comparison of prior and posterior for non-identified parameters can be misleading
- Weak identification is likely a more serious concern for applied researchers

Identification Problem in Practice

- Lack of identification leads to wrong conclusions from calibration, estimation and inference
- Many caveats due to identifiability issues and/or an unfortunate choice of observables
  - Difficult to maximize likelihood/posterior or minimize some (moment) objective function
  - Estimators often lie on the boundary of theoretically admissible space
  - Gaussian asymptotic theory yields poor approximations
- ”Unidentifiability causes no real difficulties in the Bayesian approach” (Lindley, 1971, p. 46) is misleading, because if parameters are not identifiable
  - prior becomes extremely influential and needs to be informative for a proper posterior
  - comparison of prior and posterior for non-identified parameters can be misleading
- Weak identification is likely a more serious concern for applied researchers

Is there a systematic way to detect such issues?

Identification problem

Definitions

Theoretical Lack of Identification

Distinct parameter values do not lead to distinct objective functions of data (in the injective sense)

Theoretical Lack of Identification

Distinct parameter values do not lead to distinct objective functions of data (in the injective sense)

- $\theta \in \Theta$: (unknown) vector of model parameters
- Θ: admissible parameter space for unique and stable solution
- Y_T: matrix of observables with sample size T
- $p(\theta; Y_T)$: objective function generated by a DSGE model (probability distribution, likelihood, posterior or moment's distance)

Theoretical Lack of Identification

Distinct parameter values do not lead to distinct objective functions of data (in the injective sense)

- $\theta \in \Theta$: (unknown) vector of model parameters
- Θ: admissible parameter space for unique and stable solution
- Y_T: matrix of observables with sample size T
- $p(\theta; Y_T)$: objective function generated by a DSGE model (probability distribution, likelihood, posterior or moment's distance)

Definition: Global identification (Rothenberg, 1971)

A point $\theta_0 \in \Theta$ is said to be globally identified if for all Y_T:

$$p(\theta_0; Y_T) = p(\theta_1; Y_T) \text{ implies } \theta_0 = \theta_1 \tag{1}$$

Theoretical Lack of Identification

Distinct parameter values do not lead to distinct objective functions of data (in the injective sense)

- $\theta \in \Theta$: (unknown) vector of model parameters
- Θ: admissible parameter space for unique and stable solution
- Y_T: matrix of observables with sample size T
- $p(\theta; Y_T)$: objective function generated by a DSGE model (probability distribution, likelihood, posterior or moment's distance)

Definition: Global identification (Rothenberg, 1971)

A point $\theta_0 \in \Theta$ is said to be globally identified if for all Y_T:

$$p(\theta_0; Y_T) = p(\theta_1; Y_T) \text{ implies } \theta_0 = \theta_1 \tag{1}$$

Definition: Local identification (Rothenberg, 1971)

If (1) is true only for values $\tilde{\theta}$ in an open neighborhood of θ_0, then θ_0 is said to be locally identified.

Empirical Strength of Identification

- How much information can be extracted from a specific Y_T to estimate model parameters precisely?
- More formally: What happens to the precision of estimates with a growing sample size.

Identification problem

Literature

Global Identification for Linearized Gaussian DSGE Models

- Minimizing Kullback-Leibler discrepancy (Qu and Tkachenko, 2017)
- Exploiting link between observationally equivalent state space representations and model solution constraints (Kocicecki & Kolasa, 2018)

Local Identification for Linearized Gaussian DSGE Models

- Moments (Iskrev, 2010)
- Spectral density (Qu and Tkachenko, 2012)
- Control theory for minimal systems (Komunjer and Ng, 2011)

Local Identification for non-linear and/or non-Gaussian DSGE Models

- Cumulants and Polyspectra (Mutschler, 2015)

Weak Identification for Linearized Gaussian DSGE Models

- Asymptotic information matrix (Ratto and Iskrev, 2011; Andrle, 2010)
- Bayesian indicators (Koop, Pesaran and Smith, 2013)
- Indirect inference on VAR approximation (Le, Meenagh, Minford and Wickens, 2017)
- Score test on Gaussian likelihood (Qu, 2014)

Global Identification for Linearized Gaussian DSGE Models

- Minimizing Kullback-Leibler discrepancy (Qu and Tkachenko, 2017)
- Exploiting link between observationally equivalent state space representations and model solution constraints (Kocicecki & Kolasa, 2018)

Local Identification for Linearized Gaussian DSGE Models

- Moments (Iskrev, 2010)
- Spectral density (Qu and Tkachenko, 2012)
- Control theory for minimal systems (Komunjer and Ng, 2011)

Local Identification for non-linear and/or non-Gaussian DSGE Models

- Cumulants and Polyspectra (Mutschler, 2015)

Weak Identification for Linearized Gaussian DSGE Models

- Asymptotic information matrix (Ratto and Iskrev, 2011; Andrle, 2010)
- Bayesian indicators (Koop, Pesaran and Smith, 2013)
- Indirect inference on VAR approximation (Le, Meenagh, Minford and Wickens, 2017)
- Score test on Gaussian likelihood (Qu, 2014)

Local identification in linearized DSGE models

Local identification in linearized DSGE models

Linear state space system

Dynare's linearized state-space form

$$z_t - \bar{z} = ghx \cdot (x_{t-1} - \bar{x}) + ghu \cdot u_t$$

can be cast into the more familiar state-space form

Dynare's linearized state-space form

$$z_t - \bar{z} = ghx \cdot (x_{t-1} - \bar{x}) + ghu \cdot u_t$$

can be cast into the more familiar state-space form

$$x_t - \bar{x} = A(\theta)(x_{t-1} - \bar{x}) + B(\theta)u_t \qquad [STATES]$$
$$y_t - \bar{y} = C(\theta)(x_{t-1} - \bar{x}) + D(\theta)u_t \qquad [VAROBS]$$

where x_t are state and y_t observable variables.

Dynare's linearized state-space form

$$z_t - \bar{z} = ghx \cdot (x_{t-1} - \bar{x}) + ghu \cdot u_t$$

can be cast into the more familiar state-space form

$$x_t - \bar{x} = A(\theta)(x_{t-1} - \bar{x}) + B(\theta)u_t \qquad [STATES]$$
$$y_t - \bar{y} = C(\theta)(x_{t-1} - \bar{x}) + D(\theta)u_t \qquad [VAROBS]$$

where x_t are state and y_t observable variables.

As $u_t \sim \mathcal{N}(0, \Sigma_u)$, the unconditional first and second moments are given by:

$$E[y_t] \equiv \mu_y = \bar{y}, \quad E(x_t) \equiv \mu_x = \bar{x}$$
$$E[(y_t - \bar{y})(y_t - \bar{y}')] \equiv \Sigma_y(0) = C\Sigma_x(0)C' + D\Sigma_u D'$$
$$E[(x_t - \bar{x})(x_t - \bar{x})'] \equiv \Sigma_x(0) = A\Sigma_x(0)A' + B\Sigma_u B'$$

where the latter is the fixed point of the Lyapunov-equation.

Dynare's linearized state-space form

$$z_t - \bar{z} = ghx \cdot (x_{t-1} - \bar{x}) + ghu \cdot u_t$$

can be cast into the more familiar state-space form

$$x_t - \bar{x} = A(\theta)(x_{t-1} - \bar{x}) + B(\theta)u_t \qquad [STATES]$$
$$y_t - \bar{y} = C(\theta)(x_{t-1} - \bar{x}) + D(\theta)u_t \qquad [VAROBS]$$

where x_t are state and y_t observable variables.

As $u_t \sim \mathcal{N}(0, \Sigma_u)$, the unconditional first and second moments are given by:

$$E[y_t] \equiv \mu_y = \bar{y}, \quad E(x_t) \equiv \mu_x = \bar{x}$$
$$E[(y_t - \bar{y})(y_t - \bar{y}')] \equiv \Sigma_y(0) = C\Sigma_x(0)C' + D\Sigma_u D'$$
$$E[(x_t - \bar{x})(x_t - \bar{x})'] \equiv \Sigma_x(0) = A\Sigma_x(0)A' + B\Sigma_u B'$$

where the latter is the fixed point of the Lyapunov-equation.

From this, the **theoretical autocovariogram** ($\Sigma_x(j)$ and $\Sigma_y(j)$) as well as the **spectral density** ($\mathcal{S}_{2,x}(\omega)$ and $\mathcal{S}_{2,y}(\omega)$, $\omega \in [-\pi; \pi]$) can be (easily) computed

Some notes

- In general, only observing y_t would be insufficient to fully characterize the distribution of x_t (and of z_t)
- Fortunately, our model implies restrictions through θ
- Mapping from structural parameters θ to (A, B, C, D) is highly nonlinear and only implicitly available

Local identification in linearized DSGE models

Diagnostics based on moments

Proposition Moments (Iskrev, 2010)

Suppose that

$$m(q) := \left[\mu_y', vech(\Sigma_y(0))', vec(\Sigma_y(1)), \ldots, vec(\Sigma_y(T-1))'\right]$$

is a continuously differentiable function of $\theta \in \Theta$. Let $\theta_0 \in \Theta$ be a regular point, θ is then locally identifiable at a point θ_0 from the first two moments of y_t, if the Jacobian matrix

$$J = \frac{\partial m(q)}{\partial \theta'}$$

has full column rank at θ_0 for $q \leq T$. This condition is both necessary and sufficient when $q = T$ if u_t is Gaussian.

Intuition

- Mapping from population moments to structural parameter should be unique, check injectivity by looking at Jacobian matrix
- Rank condition: Check whether derivative of the **theoretical mean, variance and autocovariogram** of observables w.r.t structural parameters has full rank
- Order condition: At least as many moments as parameters
- Very helpful in detecting observational equivalence (columns of zeros and linear dependence between parameters)
- Checking global identification is hard $\rightarrow$ check local identification for relevant parameter range
- *time domain approach*

Decompose the Jacobian

$$J = \frac{\partial m(T)}{\partial \tau'} \frac{\partial \tau}{\partial \theta'}$$

where $\tau := [\bar{z}', vec(ghx)', vech(ghu \cdot \Sigma_u \cdot ghu')']'$ denote reduced form parameters

- $\frac{\partial \tau}{\partial \theta'}$ shows how the parameters θ affect the (non-constant) model solution parts τ
- $\frac{\partial m(T)}{\partial \tau'}$ shows how the model solution maps into the observed data moments

Decompose the Jacobian

$$J = \frac{\partial m(T)}{\partial \tau'} \frac{\partial \tau}{\partial \theta'}$$

where $\tau := [\bar{z}', vec(ghx)', vech(ghu \cdot \Sigma_u \cdot ghu')']'$ denote reduced form parameters

- $\frac{\partial \tau}{\partial \theta'}$ shows how the parameters θ affect the (non-constant) model solution parts τ
- $\frac{\partial m(T)}{\partial \tau'}$ shows how the model solution maps into the observed data moments

Corollary

The point θ_0 is locally identifiable only if $\frac{\partial \tau}{\partial \theta'}$ at θ_0 has full rank.

- Necessary condition as parameters only affect distribution of observables through their effect on model solution
- It is not sufficient unless all states are observed

Local identification in linearized DSGE models

Diagnostics based on spectrum

Proposition Spectrum (Qu and Tkachenko, 2012)

Suppose that the spectral density $\mathcal{S}_{2,y}$ is continuous in $\omega \in [-\pi; \pi]$ and continuous and differentiable in $\theta \in \Theta$. Let

$$\bar{G} = \left(\frac{\partial \mu_y}{\partial \theta'}\right)' \left(\frac{\partial \mu_y}{\partial \theta'}\right) + \int_{-\pi}^{\pi} \left(\frac{\partial \mathcal{S}_{2,y}}{\partial \theta'}\right)^* \left(\frac{\partial \mathcal{S}_{2,y}}{\partial \theta'}\right) \mathrm{d}\omega$$

and $\theta_0 \in \Theta$ be a regular point. Furthermore, assume there is an open neighborhood of θ_0 in which $\bar{G}$ has a constant rank. Then θ is locally identifiable at a point θ_0 from the mean and spectrum of y_t, if and only if $\bar{G}$ is non-singular and equal to the number of parameters.

Intuition

- Mapping from population mean and spectral density to structural parameters should be unique, check injectivity by looking at Jacobian matrix
- Rank condition: Check whether derivative of the **theoretical mean and spectrum** of observables w.r.t structural parameters has full rank
- Very helpful in detecting observational equivalence (columns of zeros and linear dependence between parameters)
- Gram matrix structure numerically facilitates rank computations, no order condition required
- Checking global identification is hard $\rightarrow$ check local identification for relevant parameter range

$\hookrightarrow$ *frequency domain approach*

Local identification in linearized DSGE models

Diagnostics based on control theory for minimal systems

Minimal State-Space System

- Dynamics are entirely driven by the smallest possible dimension of the state vector (and shocks)
- Definition of minimality:
  - Controllability: For any initial state, it is always possible to design an input sequence that puts the system in the desired final state
  - Observability: Given the evolution of the input it is always possible to reconstruct the initial state by observing the evolution of the output
- Solution in Dynare is not based on the minimal state representation
- Numerical procedures (pole-zero cancellation) do not necessarily output minimal states with economic meaning

Proposition Minimal System (Komunjer and Ng, 2011)

Consider the minimal linearized DSGE model solution. Assume that the vector containing the minimal solution matrices

$$\Lambda := \left(vec(\tilde{A})', vec(\tilde{B})', vec(\tilde{C})', vec(\tilde{D})', vech(\tilde{\Sigma_u})'\right)'$$

is continuously differentiable on Θ. Two triples $(\theta_0, I_{\tilde{x}}, I_u)$ and (θ_1, T, U) are observationally equivalent if

$$\tilde{A}(\theta_1) = T\tilde{A}(\theta_0)T^{-1}, \tilde{B}(\theta_1) = T\tilde{B}(\theta_0)U^{-1}, \tilde{C}(\theta_1) = \tilde{C}(\theta_0)T^{-1}, \Sigma_u(\theta_1) = U^{-1}\Sigma_u(\theta_0)U^{-1}$$

with similarity transformation matrices T and U being full rank matrices.

Let $\theta_0 \in \Theta$ be a regular point, then θ is locally identifiable at a point θ_0 from the mean, autocovariances, and spectrum of y_t if and only if:

$$\bar{\Delta} := \begin{pmatrix} \frac{\partial \mu_y}{\partial \theta'} & 0 & 0 \\ \frac{\partial vec(\widetilde{\mathcal{A}})}{\partial \theta'} & \widetilde{\mathcal{A}}' \otimes I - I \otimes \widetilde{\mathcal{A}} & 0 \\ \frac{\partial vec(\widetilde{\mathcal{B}})}{\partial \theta'} & \widetilde{\mathcal{B}}' \otimes I & I \otimes \widetilde{\mathcal{B}} \\ \frac{\partial vec(\widetilde{\mathcal{C}})}{\partial \theta'} & -I \otimes \widetilde{\mathcal{C}} & 0 \\ \frac{\partial vec(\widetilde{\mathcal{D}})}{\partial \theta'} & 0 & I \otimes \widetilde{\mathcal{D}} \\ \frac{\partial vec(\Sigma_u)}{\partial \theta'} & 0 & -2[\Sigma_u \otimes I] \end{pmatrix}$$

has full column rank.

Intuition

- Based upon identification results from control theory for minimal systems
- Derive restrictions implied by equivalent model dynamics without computing any autocovariances or the spectral density
  - for given size of shocks Σ_u, each transfer function is potentially obtained from a multitude of quadruples (A, B, C, D)
  - many pairs of transfer functions and size of shocks Σ_u that jointly generate the same spectral density or autocovariogram
- Rank condition: Check injectivity of these restrictions using a Jacobian matrix
- Order condition and some other useful diagnostics for IRFs

Local identification in linearized DSGE models

Dynare implementation

Dynare command `identification;`

Triggers the local identification tests and has two modes of operation:

- Point identification check (default)
- Monte Carlo mode (`prior_mc>1`):
  - Draw θ from Θ (ensuring stability and determinacy)
  - Compute point identification checks and repeat many times

If there is an `estimated_params` block

- with prior information: the program performs the local identification checks for the estimated parameters at the prior mean (prior mode, posterior mean and posterior mode are also alternative options)
- for ML estimation (no prior definition): local identification checks are performed for the estimated parameters at the actual or initial value declared for estimation (ML value is also possible)

Otherwise all model parameters and stderr parameters are considered

WARNING: Use best practices in your mod file!

- Dynare's symbolic preprocessor interprets and implements the model definitions as expressed in the Dynare file
- It will not reflect all parameter definitions which may be hidden e.g. in a custom `_steadystate.m` file (use `steady_state_model` instead)
- Try to avoid declaring auxiliary parameters, but use the '`#`' syntax in the `model` block of the Dynare file to declare endogenous parameters instead

Tracking singularities

Whenever some of the Jacobian matrices are rank deficient, the code tries to diagnose the subset of parameters responsible for the rank deficiency:

1. ranks are computed using the singular value decomposition
2. if there are columns of zeros in the Jacobian matrix, the associated parameter is printed on the MATLAB command window
3. compute pairwise- and multi-correlation coefficients for each column of the Jacobian matrix: if there are parameters with correlation coefficients equal to unity, these are printed on the MATLAB command window

Alternatively, a brute-force approach to check rank conditions for all possible combinations can be triggered (`checks_via_subsets=1`)

Example 'Point Identification Mode': `investadjust.mod`

```
======== Identification Analysis ========
...
Testing calibration
...
REDUCED-FORM:
!!!WARNING!!!
The rank of Tau (Jacobian of steady state and reduced-form solution matrices) is deficient!

RA is not identified!

MINIMAL SYSTEM (KOMUNJER AND NG, 2011):
!!!WARNING!!!
The rank of Deltabar (Jacobian of steady state and minimal system) is deficient!

RA is not identified!
[KAPPA,THETA] are PAIRWISE collinear!

SPECTRUM (QU AND TKACHENKO, 2012):
!!!WARNING!!!
The rank of Gbar (Jacobian of mean and spectrum) is deficient!

RA is not identified!
[KAPPA,THETA] are PAIRWISE collinear!

MOMENTS (ISKREV, 2010):
!!!WARNING!!!
The rank of J (Jacobian of first two moments) is deficient!

RA is not identified!
[KAPPA,THETA] are PAIRWISE collinear!

==== Identification analysis completed ====
```

Example 'Monte Carlo Mode': `investadjust.mod`

```
======== Identification Analysis ========
...
Monte Carlo Testing

Testing MC sample
...
REDUCED-FORM:
!!!WARNING!!!
The rank of Tau (Jacobian of steady state and reduced-form solution matrices) is deficient for 20 out of 20 MC r

RA is not identified for 100% of MC runs!

MINIMAL SYSTEM (KOMUNJER AND NG, 2011):
!!!WARNING!!!
The rank of Deltabar (Jacobian of steady state and minimal system) is deficient for 20 out of 20 MC runs!

RA is not identified for 100% of MC runs!
[KAPPA,THETA] are PAIRWISE collinear for 100% of MC runs!

SPECTRUM (QU AND TKACHENKO, 2012):
!!!WARNING!!!
The rank of Gbar (Jacobian of mean and spectrum) is deficient for 20 out of 20 MC runs!

RA is not identified for 100% of MC runs!
[KAPPA,THETA] are PAIRWISE collinear for 90% of MC runs!

MOMENTS (ISKREV, 2010):
!!!WARNING!!!
The rank of J (Jacobian of first two moments) is deficient for 20 out of 20 MC runs!

RA is not identified for 100% of MC runs!
[KAPPA,THETA] are PAIRWISE collinear for 100% of MC runs!

==== Identification analysis completed ====
```

Local identification in linearized DSGE models

Computational Remarks

Settings

```
Note that differences in the criteria could be due to numerical settings,
numerical errors or the method used to find problematic parameter sets.

Settings:
Derivation mode for Jacobians:                          Analytic using sylvester equations
Method to find problematic parameters:                  Nullspace and multicorrelation coefficients
Normalize Jacobians:                                    Yes
Tolerance level for rank computations:                  robust
Tolerance level for selecting nonzero columns:          1e-08
Tolerance level for selecting nonzero singular values: 1e-03
```

- Key issue: distinguish possible weak identification, that is, near linear dependence, form true perfect collinearity
- Errors of numerical differentiations: the rank test for singularity is much more sensitive to the significance threshold set the by user in checking the rank when numerical derivatives are used
- Whenever possible the code uses closed-form expressions to compute the Jacobians analytically

Analytic Jacobians

- Basic idea: view f as a function of θ and of the steady-state vector $\bar{z}(\theta)$, which is also a function of θ

Analytic Jacobians

- Basic idea: view f as a function of θ and of the steady-state vector $\bar{z}(\theta)$, which is also a function of θ
- Thus, implicitly we have $f(\bar{z}(\theta), \theta) = 0$

Analytic Jacobians

- Basic idea: view f as a function of θ and of the steady-state vector $\bar{z}(\theta)$, which is also a function of θ
- Thus, implicitly we have $f(\bar{z}(\theta), \theta) = 0$
- Differentiating yields parameter derivatives of dynamic model Jacobian (`g_1`) and dynamic Hessian (`g_2`) and also higher-order derivatives of dynamic model

Analytic Jacobians

- Basic idea: view f as a function of θ and of the steady-state vector $\bar{z}(\theta)$, which is also a function of θ
- Thus, implicitly we have $f(\bar{z}(\theta), \theta) = 0$
- Differentiating yields parameter derivatives of dynamic model Jacobian (`g_1`) and dynamic Hessian (`g_2`) and also higher-order derivatives of dynamic model
- Then it is straightforward (but very tedious) to derive closed-forms for the derivatives of first-order, second-order and third-order solution matrices wrt parameters

Analytic Jacobians

- Basic idea: view f as a function of θ and of the steady-state vector $\bar{z}(\theta)$, which is also a function of θ
- Thus, implicitly we have $f(\bar{z}(\theta), \theta) = 0$
- Differentiating yields parameter derivatives of dynamic model Jacobian (`g_1`) and dynamic Hessian (`g_2`) and also higher-order derivatives of dynamic model
- Then it is straightforward (but very tedious) to derive closed-forms for the derivatives of first-order, second-order and third-order solution matrices wrt parameters

  ↪ `get_perturbation_params_derivs.m`

Analytic Jacobians

- Basic idea: view f as a function of θ and of the steady-state vector $\bar{z}(\theta)$, which is also a function of θ
- Thus, implicitly we have $f(\bar{z}(\theta), \theta) = 0$
- Differentiating yields parameter derivatives of dynamic model Jacobian (`g_1`) and dynamic Hessian (`g_2`) and also higher-order derivatives of dynamic model
- Then it is straightforward (but very tedious) to derive closed-forms for the derivatives of first-order, second-order and third-order solution matrices wrt parameters

  ↪ `get_perturbation_params_derivs.m`
- Once we have that, we can compute identification Jacobians in closed-form (no need to rely on numerical differentiation)

Analytic Jacobians

- Basic idea: view f as a function of θ and of the steady-state vector $\bar{z}(\theta)$, which is also a function of θ
- Thus, implicitly we have $f(\bar{z}(\theta), \theta) = 0$
- Differentiating yields parameter derivatives of dynamic model Jacobian (`g_1`) and dynamic Hessian (`g_2`) and also higher-order derivatives of dynamic model
- Then it is straightforward (but very tedious) to derive closed-forms for the derivatives of first-order, second-order and third-order solution matrices wrt parameters

  ↪ `get_perturbation_params_derivs.m`
- Once we have that, we can compute identification Jacobians in closed-form (no need to rely on numerical differentiation)
- Closed-form expressions using either Kronecker products or (more efficient) generalized Sylvester equations (default)

Identification Strength

Identification Strength

Intuition

Weak identification

- Even though all parameters are locally identifiable (enter objective function separately and it has a unique extremum), its curvature may be small in certain regions of the parameter space (especially in small samples)
- Diagnostics are based on **precision** of parameter estimates computed via
  - Inverse of (asymptotic) Fisher Information Matrix (Andrle, 2010; Ratto and Iskrev, 2011)
  - Bayesian Learning Rate Indicator for growing sample sizes (Koop, Pesaran, Smith, 2013)

Identification Strength

Asymptotic Information Matrix

The Information Matrix Revisited

The precision of parameter estimates at the mode is given (asymptotically) by the inverse of the **Fisher Information Matrix**

$$\mathcal{I}(\theta) = E\left[\left(\frac{\partial \log\left(p(y|\theta)\right)}{\partial \theta'}\right)' \left(\frac{\partial \log\left(p(y|\theta)\right)}{\partial \theta'}\right)\right]$$

Typically, non-singularity of this matrix is sufficient for local identification

The Information Matrix Revisited

The precision of parameter estimates at the mode is given (asymptotically) by the inverse of the **Fisher Information Matrix**

$$\mathcal{I}(\theta) = E\left[\left(\frac{\partial \log\left(p(y|\theta)\right)}{\partial \theta'}\right)' \left(\frac{\partial \log\left(p(y|\theta)\right)}{\partial \theta'}\right)\right]$$

Typically, non-singularity of this matrix is sufficient for local identification

Can be decomposed

$$\mathcal{I}(\theta) = \Delta^{\frac{1}{2}} \tilde{\mathcal{I}}(\theta) \Delta^{\frac{1}{2}}$$

into a variance matrix

$$\Delta = diag(\mathcal{I}(\theta))$$

and a correlation matrix

$$\tilde{\mathcal{I}}(\theta)$$

Identification Strength Intuition

Here, we can again see our two reasons for non-identifiability:

1. The likelihood does not change when parameter i changes:

$$\frac{\partial \log\left(p(y|\theta)\right)}{\partial \theta_i} = 0 \Leftrightarrow \Delta_i = 0$$

2. The effect on the likelihood is offset due to perfect correlation:

$$\rho_i \equiv corr\left(\frac{\partial \log\left(p(y|\theta)\right)}{\partial \theta_i}, \frac{\partial \log\left(p(y|\theta)\right)}{\partial \theta_{-i}}\right) = 1$$

This suggests that we can use this curvature information also for **weak identification** as in that case $\Delta_i \approx 0$ and/or $\rho_i \approx 1$

Identification Strength Formally

- If $\mathcal{I}(\theta)$ is not singular, Cramer-Rao bound of the uncertainty of estimated θ_i:

$$\operatorname{std}(\theta_i) \geq \sqrt{\{\mathcal{I}(\theta)^{-1}\}_{(ii)}}$$

- **Identification strength** can therefore be measured using (Andrle, 2010):

$$s_i(\theta_i) \equiv \operatorname{std}(\theta_i)^{-1} = \frac{1}{\sqrt{\{\mathcal{I}(\theta)^{-1}\}_{(ii)}}} = \sqrt{\Delta_i(1-\rho_i^2)}$$

- As this precision measure only uses the population objective function (expected log-likelihood), it is an a-priori measure
- At the same time, it can be interpreted as the lower bound on the estimation uncertainty in an unbiased finite sample estimator

Identification Strength

Dynare implementation

Identification Strength Measure in Dynare

- Problem: measure is percentage change in likelihood for unit change in parameter $\rightarrow$ not unit free
- Solution: multiply by the value of θ_i to get a re-normalized version of the 'curvature':

$$s_i^{dyn}(\theta_i) \equiv \theta_i \times s_i(\theta_i) = \sqrt{\frac{{\theta_i}^2}{\{\mathcal{I}(\theta)^{-1}\}_{(ii)}}} = \sqrt{{\theta_i}^2 \Delta_i} \times \sqrt{(1-\rho_i^2)} \qquad \text{[STRENGTH]}$$

Identification Strength Measure in Dynare

- Problem: measure is percentage change in likelihood for unit change in parameter $\rightarrow$ not unit free
- Solution: multiply by the value of θ_i to get a re-normalized version of the 'curvature':

$$s_i^{dyn}(\theta_i) \equiv \theta_i \times s_i(\theta_i) = \sqrt{\frac{{\theta_i}^2}{\{\mathcal{I}(\theta)^{-1}\}_{(ii)}}} = \sqrt{{\theta_i}^2 \Delta_i} \times \sqrt{(1-\rho_i^2)} \qquad \text{[STRENGTH]}$$

- Can be interpreted as an 'a-priori t-test'
- Taking square of θ assures it to be positive

Identification Strength Measure in Dynare

- Problem: measure is percentage change in likelihood for unit change in parameter $\rightarrow$ not unit free
- Solution: multiply by the value of θ_i to get a re-normalized version of the 'curvature':

$$s_i^{dyn}(\theta_i) \equiv \theta_i \times s_i(\theta_i) = \sqrt{\frac{{\theta_i}^2}{\{\mathcal{I}(\theta)^{-1}\}_{(ii)}}} = \sqrt{{\theta_i}^2 \Delta_i} \times \sqrt{(1-\rho_i^2)} \qquad \text{[STRENGTH]}$$

- Can be interpreted as an 'a-priori t-test'
- Taking square of θ assures it to be positive
- The sensitivity component contained in this measure (as opposed to the correlation component) is defined as

$$\Delta_i^{dyn} = \sqrt{\theta_i^2 \Delta_i} = \sqrt{\theta_i^2 \mathcal{I}(\theta)_{(i,i)}} \qquad \text{[SENSITIVITY]}$$

is the elasticity of the likelihood function w.r.t. θ_i (keeping all other parameters constant)

Alternative Normalization

- Identification is often checked over the prior region, e.g. at the prior mean
- Because some parameters may have a prior mean of zero, alternative normalization using the **prior standard deviation** $\sigma(\theta_i)$:

Alternative Normalization

- Identification is often checked over the prior region, e.g. at the prior mean
- Because some parameters may have a prior mean of zero, alternative normalization using the **prior standard deviation** $\sigma(\theta_i)$:

$$s_i^{prior} = \sigma(\theta_i) / \sqrt{(\mathcal{I}(\theta)^{-1})_{(i,i)}} \qquad \text{[STRENGTH]}$$

$$\Delta_i^{prior} = \sigma(\theta_i) \cdot \sqrt{\mathcal{I}(\theta)_{(i,i)}} \qquad \text{[SENSITIVITY]}$$

Identification Strength Plots

The identification toolbox shows, after the check of rank conditions, the **plots of the strength of identification and of the sensitivity component** for all estimated parameters.

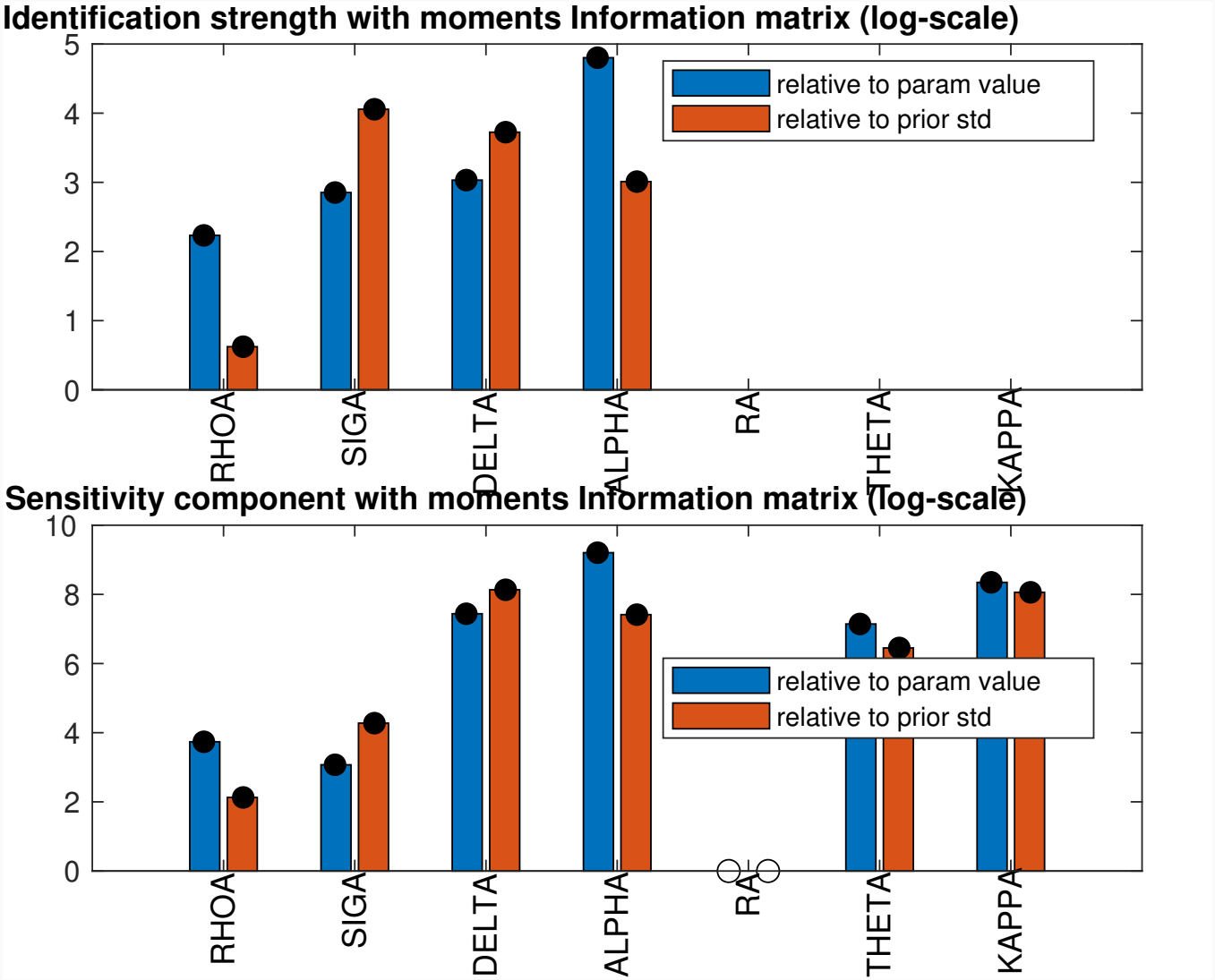


Figure 2: see `investadjust.mod`

Numerical Remark: Computing $\mathcal{I}(\theta)$

Asymptotic Information Matrix

Given a sample size T, the Fischer information matrix $\mathcal{I}(\theta)$ can be computed analytically as discussed in Iskrev (2009). This also provides an estimate of the Cramer-Rao lower bound of uncertainty of θ.

Simulated moments

Perform stochastic simulations for T periods and compute sample moments of observed variables; repeat for N_r replicas, getting N_r samples of simulated moments; take the covariance matrix $\Sigma(m_T)$ of (first and second) simulated moments.

A 'moment information matrix' can be defined as $\mathcal{I}(\theta|m_T) = J \cdot \Sigma(m_T)^{-1} \cdot J'$. This does NOT provide an estimate of the Cramer-Rao lower bound. Used with stochastic singularity and whenever the Asymptotic Information Matrix cannot be computed.

Identification Strength

Bayesian Learning Rate Indicator

Koop, Pesaran and Smith (2013)'s approach

- Basic idea: strength of identification becomes better as more data becomes available, i.e. parameters are estimated more precisely
- Suppose θ_2 is identified, whereas θ_1 is weakly identified such that the rank of the reduced-form parameters depends on the sample size T
- For growing T the posterior precision of
  - θ_1 divided by the sample size will go to **zero**.
  - θ_2 divided by the sample size will go to a **constant**.

$\hookrightarrow$ *Bayesian simulation approach*

Implementation

- Not yet implemented as an option in the toolbox, but actually very easy to do:
- Simulate large dataset, estimate with Bayesian MCMC using first e.g. 75, 150, 300, 600,... observations
- Look at either average posterior precisions or convergence ratios:
  - ratio of two subsequent estiamted posterior precision values, e.g. at T=150 and T=300, and check whether this ratio is close to the rate at which T increases, i.e. close to 300/150=2
- Slice with double rotation (50x200x2500 draws) works very well as it requires almost no fine-tuning (and no mode-finding step)
- Use parallel option to get more chains

Average Posterior Precisions

T	α	R^A	δ	ρ_A	σ_A	θ	κ
100	5.02419	0.16006	68.06352	2.35966	2.23979	0.01789	0.00460
300	1.66330	0.05318	10.19060	1.66202	0.37313	0.00576	0.00227
900	0.60317	0.01777	4.79942	1.37207	0.24600	0.00194	0.00066
2700	0.23070	0.00594	2.50897	1.28453	0.16133	0.00066	0.00015
8100	0.10292	0.00199	1.48975	1.29131	0.06690	0.00023	0.00005

Average Posterior Precisions

T	α	R^A	δ	ρ_A	σ_A	θ	κ
100	5.02419	0.16006	68.06352	2.35966	2.23979	0.01789	0.00460
300	1.66330	0.05318	10.19060	1.66202	0.37313	0.00576	0.00227
900	0.60317	0.01777	4.79942	1.37207	0.24600	0.00194	0.00066
2700	0.23070	0.00594	2.50897	1.28453	0.16133	0.00066	0.00015
8100	0.10292	0.00199	1.48975	1.29131	0.06690	0.00023	0.00005

Convergence Ratios

dT	α	R^A	δ	ρ_A	σ_A	θ	κ
300/100	0.993	0.997	0.449	2.113	0.500	0.966	1.479
900/300	1.088	1.003	1.413	2.477	1.978	1.009	0.868
2700/900	1.147	1.002	1.568	2.809	1.967	1.017	0.693
8100/2700	1.338	1.004	1.781	3.016	1.244	1.063	0.927

Average Posterior Precisions

T	α	R^A	δ	ρ_A	σ_A	θ	κ
100	5.02419	0.16006	68.06352	2.35966	2.23979	0.01789	0.00460
300	1.66330	0.05318	10.19060	1.66202	0.37313	0.00576	0.00227
900	0.60317	0.01777	4.79942	1.37207	0.24600	0.00194	0.00066
2700	0.23070	0.00594	2.50897	1.28453	0.16133	0.00066	0.00015
8100	0.10292	0.00199	1.48975	1.29131	0.06690	0.00023	0.00005

Convergence Ratios

dT	α	R^A	δ	ρ_A	σ_A	θ	κ
300/100	0.993	0.997	0.449	2.113	0.500	0.966	1.479
900/300	1.088	1.003	1.413	2.477	1.978	1.009	0.868
2700/900	1.147	1.002	1.568	2.809	1.967	1.017	0.693
8100/2700	1.338	1.004	1.781	3.016	1.244	1.063	0.927

Insights

- Only ρ_A is strongly identifiable
- Estimating non-identified models yields severe problems in the estimation of other, actually identified model parameters, as we do not fix either θ or κ
- Fixing theoretical lack of identification is also important from a Bayesian perspective

The advanced option of the toolbox

Point mode

- Different Sensitivity Measure
- Analysis of the LRE form (steady-state and dynamic model derivatives) for all model variables
- Analysis of the reduced form (steady-state and solution matrices) for all model variables
- Analysis of identification patterns based on regressions and SVD

Point mode

- Different Sensitivity Measure
- Analysis of the LRE form (steady-state and dynamic model derivatives) for all model variables
- Analysis of the reduced form (steady-state and solution matrices) for all model variables
- Analysis of identification patterns based on regressions and SVD

Monte Carlo mode

- Analysis of the condition number of the Jacobians of moments, reduced-form and Linear Rational Expectations model and detection of the parameters that mostly drive large condition numbers (i.e. weaker identification)
- Analysis of the identification patterns across the Monte Carlo sample
- Detailed point-estimate (identification strength and collinearity analysis) of the parameters set having the smallest/largest condition number
- When some singularity (rank condition failure) is detected for some elements of the Monte Carlo sample, detailed point-estimates are performed for such critical points

A Different Sensitivity Measure

- Another sensitivity measure is how changes in the elements of the parameter vector θ impact the model moments, the reduced form solution, and the dynamic model
- Can be measured locally using the corresponding Jacobian with certain normalizations to account for
  - different parameter uncertainty by ascribing more importance to more variable parameters
  - differently volatile moments, solution matrices, or dynamic Jacobians
- Norm of the columns of the standardized Jacobian yields single aggregate sensitivity measure over all moments, solution matrices, or dynamic model Jacobians for each parameter

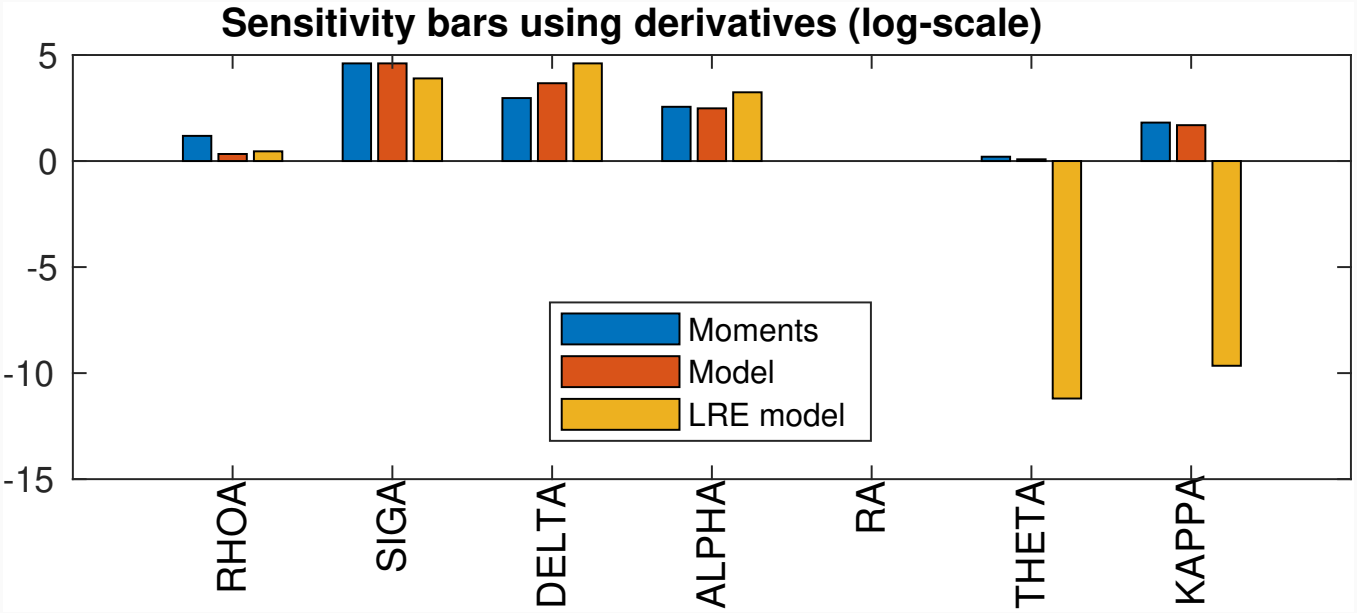


Figure 3: Sensitivity measure, see `investadjust.mod`

The respective Jacobian matrices refer to the

1. moments matrix ($\partial m_T / \partial \theta'$), indicating how well a parameter can be identified due the strength of its impact on the observed moments
2. model solution matrices ($\partial \tau / \partial \theta'$), indicating how well a parameter could in principle be identified if all state variables were observed
3. Linear Rational Expectations model ($\partial LRE / \partial \theta'$), indicating trivial cases of non-identifiability due to e.g. some parameters always showing up as a product in the model equations

Analyzing identification patterns

1. Iskrev (2010): check which group of one, two or more parameters is most capable to mimic (replace) the effect of each parameter. A brute force search is done for each column of $J_{(j)}$ to detect the group of columns $J_{(l \not\supseteq j)}$, having the highest explanatory power for $J_{(j)}$ by a linear regression
2. Andrle (2010): the identification patterns are shown by taking the singular value decomposition of $\mathcal{I}(\theta)$ or of the J matrix and displaying the eigenvectors corresponding to the smallest (or highest) singular values

Correlation Component (Iskrev, 2010)

- Even if the likelihood itself is sensitive, there might be (perfect) co-linearity
- Need to look at correlation component of columns
- This analysis is conducted via brute force: For each single parameter, a set of regressions is run of the column of the Jacobian corresponding to the parameter in the row on all possible combinations of other Jacobian columns
- Aim: finding the column (and thus parameter) combination with the highest R^2.

Correlation Component of Columns, 1 parameter

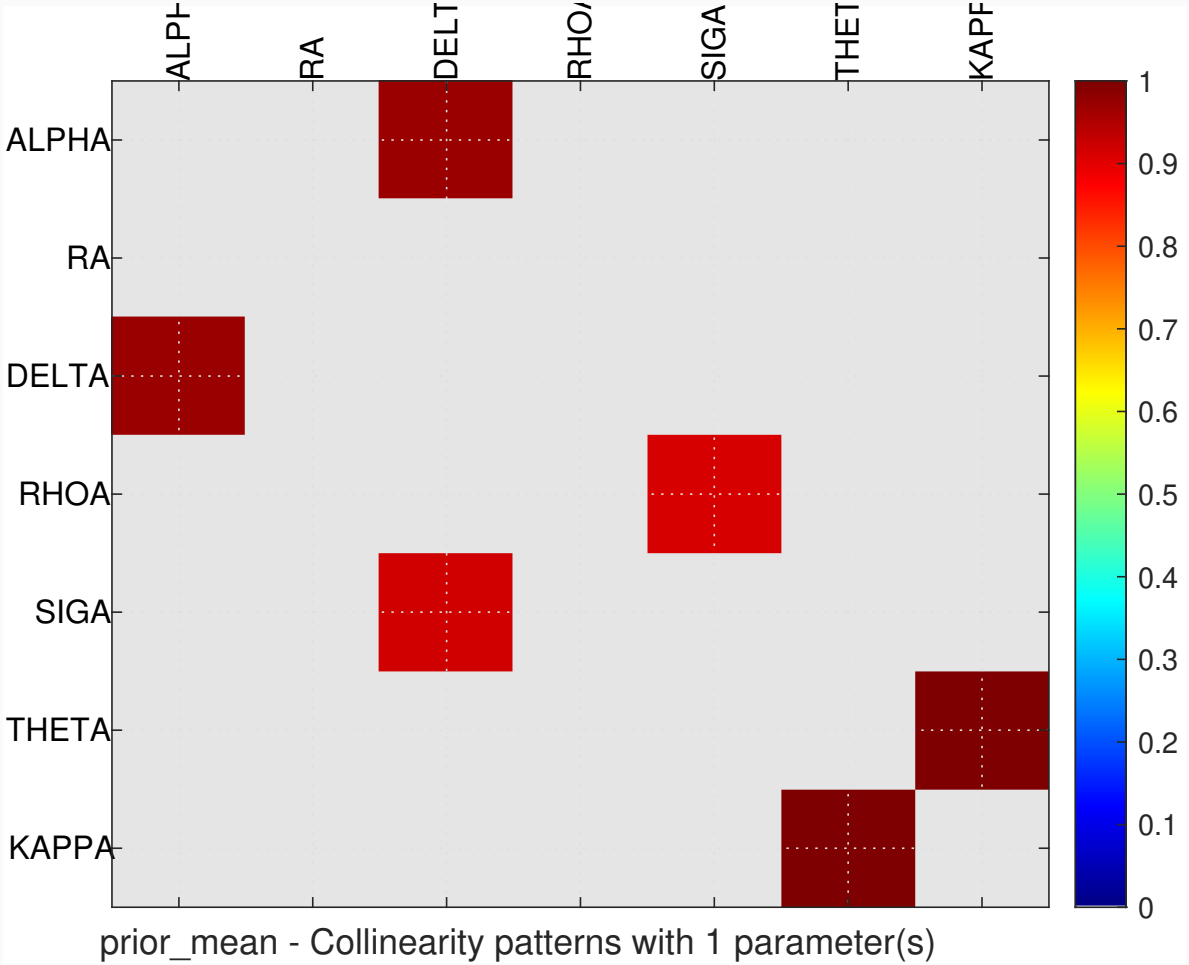


Figure 4: see `investadjust.mod`

Correlation Component of Columens, 2 parameter

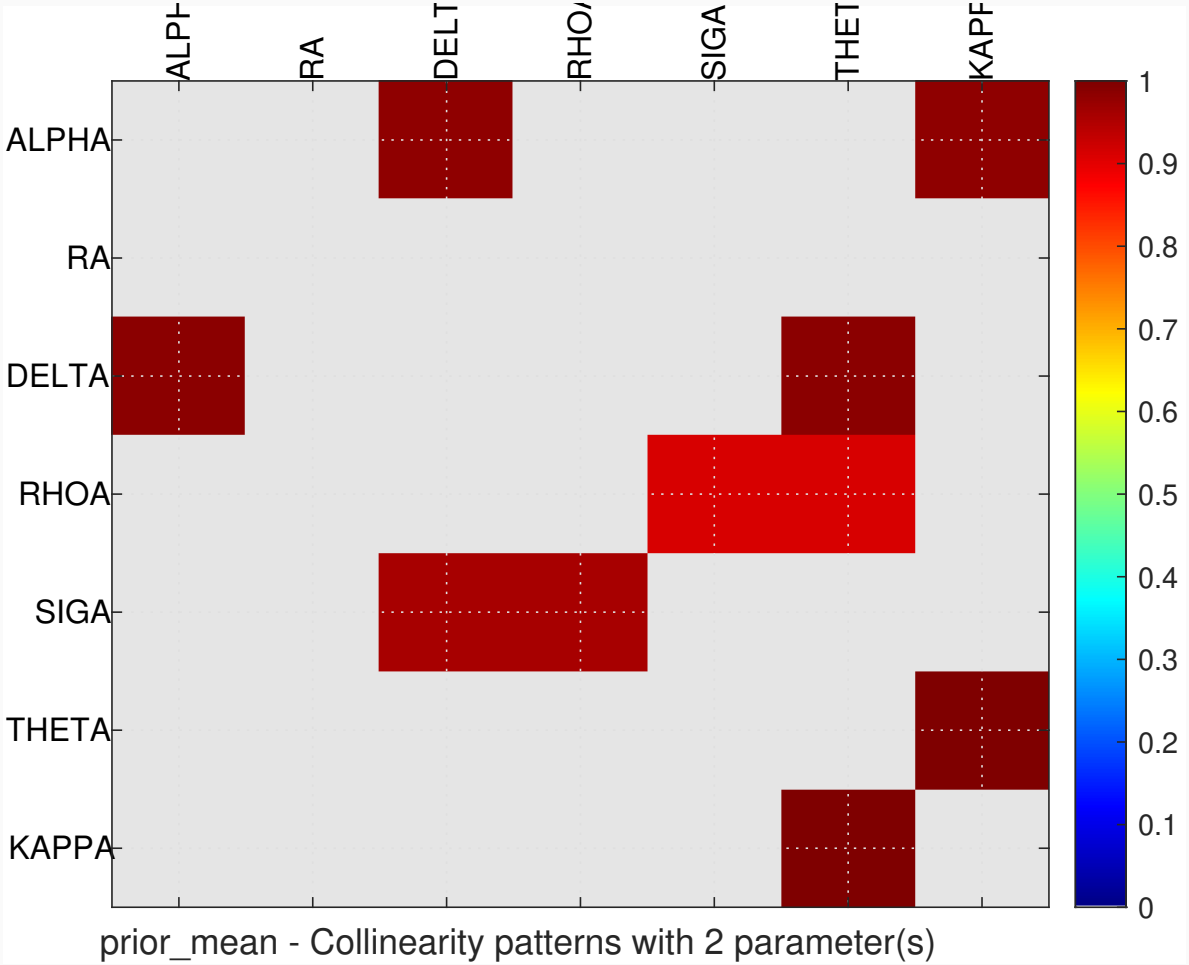
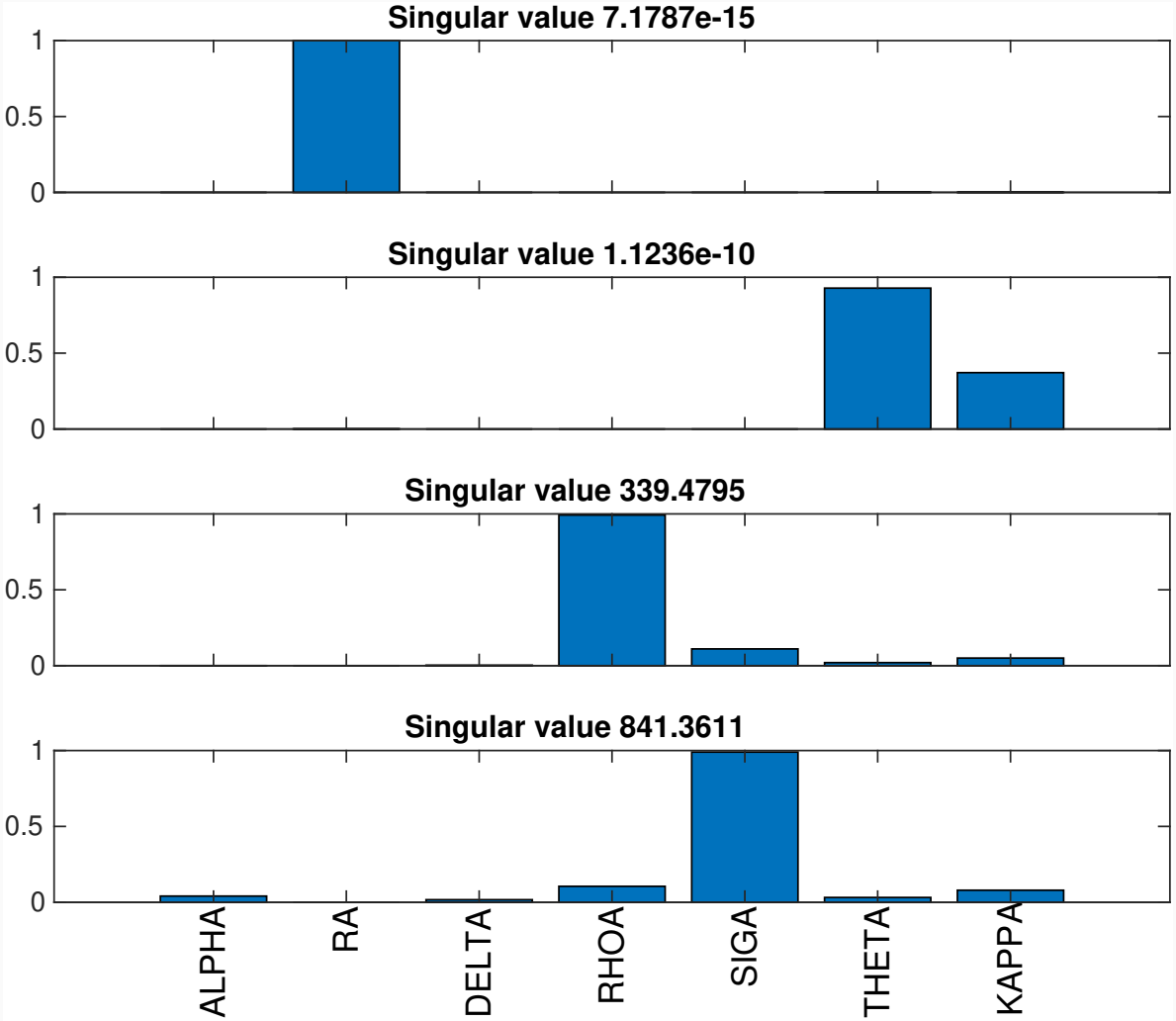


Figure 5: see `investadjust.mod`

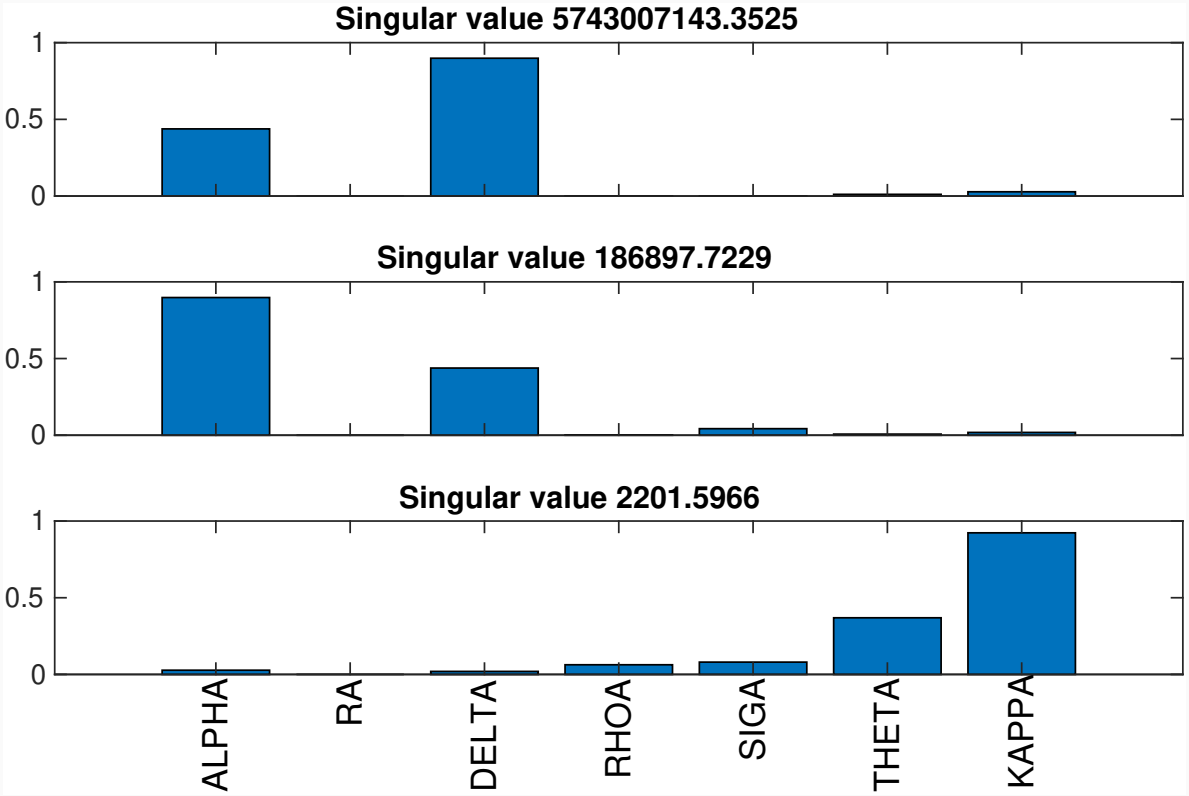
Identification Patterns Based on Singular Values

- Following Andrle (2010), identification can also be judged from a singular value decomposition (SVD) of the information matrix
- Provides the size of the singular values and the associated eigenvectors (i.e. parameters)
- Parameter combinations associated with the smallest singular values are closest to being perfectly collinear and thus redundant
- Singular value of 0 implies that the parameter is completely unidentified as it is responsible for the information matrix being rank deficient

Identification patterns – Smallest SV



Identification patterns – Largest SV



Identification of nonlinear and non-Gaussian DSGE models

Identification of nonlinear and non-Gaussian DSGE models

Pruned-State-Space

General DSGE model

$$E_t f\left(z_{t+1}, z_t, z_{t-1}, u_t|\theta\right) = 0,$$

$$z_t = g(z_{t-1}, u_t|\theta), \qquad y_t = \tilde{g}(x_t, u_t|\theta),$$

where y_t are observables, x_t states, u_t shocks and z_t all endogenous

General DSGE model

$$E_t f\left(z_{t+1}, z_t, z_{t-1}, u_t|\theta\right) = 0,$$
$$z_t = g(z_{t-1}, u_t|\theta), \qquad y_t = \tilde{g}(x_t, u_t|\theta),$$

where y_t are observables, x_t states, u_t shocks and z_t all endogenous

Solution method: Perturbation

- Taylor-approximation around the non-stochastic steady-state:

$$\begin{aligned} z_t &= \bar{z} + g_x(x_{t-1} - \bar{x}) + g_u u_t \\ &+ \frac{1}{2}\left[g_{xx}(x_{t-1} - \bar{x}) \otimes (x_{t-1} - \bar{x}) + 2g_{xu}(x_{t-1} \otimes u_t) + g_{uu}(u_t \otimes u_t) + g_{\sigma\sigma}\sigma^2\right] \\ &+ \frac{1}{6}[\ldots] + \ldots \end{aligned}$$

Problem of higher-order Taylor approximations

- Possibility of explosive behavior in higher-order approximations
- Model may not be stationary or does not have an ergodic probability distribution

Problem of higher-order Taylor approximations

- Possibility of explosive behavior in higher-order approximations
- Model may not be stationary or does not have an ergodic probability distribution

Solution: Pruning

- Idea: Leaving out terms in the solution that have higher-order effects than the approximation order
- Kim, Kim, Schaumburg and Sims (2008) and Andreasen, Fernández-Villaverde and Rubio-Ramírez (2016) show that pruned state space is stationary and ergodic
- Lombardo and Uhlig (2014) or Lan and Meyer-Gohde (2013) provide theoretical foundation for this seemingly *ad-hoc* procedure

Example with simple univariate model

- $x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t, |g_x| < 1, g_{xx} > 0$

Example with simple univariate model

- $x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t, |g_x| < 1, g_{xx} > 0$
- Two fixed points: $\bar{x} = 0$ and $\bar{x} = (1 - g_x)/g_{xx}$

Example with simple univariate model

- $x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t, |g_x| < 1, g_{xx} > 0$
- Two fixed points: $\bar{x} = 0$ and $\bar{x} = (1 - g_x)/g_{xx}$
  - $\hookrightarrow$ Once the model passes the second steady state it explodes

Example with simple univariate model

- $x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t, |g_x| < 1, g_{xx} > 0$
- Two fixed points: $\bar{x} = 0$ and $\bar{x} = (1 - g_x)/g_{xx}$
  - $\hookrightarrow$ Once the model passes the second steady state it explodes
- Pruning: Decompose state vector into 1st- and 2nd-order effects

$$x_t = x_t^f + x_t^s = g_x x_{t-1}^f + g_x x_{t-1}^s + g_{xx} \left(x_{t-1}^f\right)^2 + 2 g_{xx} \left(x_{t-1}^f x_{t-1}^s\right)^2 + g_{xx} \left(x_{t-1}^s\right)^2 + g_u u_t$$

Example with simple univariate model

- $x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t, |g_x| < 1, g_{xx} > 0$
- Two fixed points: $\bar{x} = 0$ and $\bar{x} = (1 - g_x)/g_{xx}$

  $\hookrightarrow$ Once the model passes the second steady state it explodes
- Pruning: Decompose state vector into 1st- and 2nd-order effects

$$x_t = x_t^f + x_t^s = g_x x_{t-1}^f + g_x x_{t-1}^s + g_{xx} \left(x_{t-1}^f\right)^2 + 2 g_{xx} \left(x_{t-1}^f x_{t-1}^s\right)^2 + g_{xx} \left(x_{t-1}^s\right)^2 + g_u u_t$$

- Stable solution: Prune terms that contain $x_t^f x_t^s$ and $(x_t^s)^2$

$$x_t^f = g_x x_{t-1}^f + g_u u_t, \quad x_t^s = g_x x_{t-1}^s + g_{xx} (x_{t-1}^f)^2,$$

Example with simple univariate model

- $x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t$, $|g_x| < 1, g_{xx} > 0$
- Two fixed points: $\bar{x} = 0$ and $\bar{x} = (1 - g_x)/g_{xx}$

  $\hookrightarrow$ Once the model passes the second steady state it explodes

- Pruning: Decompose state vector into 1st- and 2nd-order effects

$$x_t = x_t^f + x_t^s = g_x x_{t-1}^f + g_x x_{t-1}^s + g_{xx} \left(x_{t-1}^f\right)^2 + 2 g_{xx} \left(x_{t-1}^f x_{t-1}^s\right)^2 + g_{xx} \left(x_{t-1}^s\right)^2 + g_u u_t$$

- Stable solution: Prune terms that contain $x_t^f x_t^s$ and $(x_t^s)^2$

$$x_t^f = g_x x_{t-1}^f + g_u u_t, \quad x_t^s = g_x x_{t-1}^s + g_{xx} (x_{t-1}^f)^2,$$

- Pruned solution can be rewritten as a stable state-space system

$$\underbrace{\begin{pmatrix} \\ \end{pmatrix}}_{} = \underbrace{\begin{pmatrix} \\ \end{pmatrix}}_{} \underbrace{\begin{pmatrix} \\ \end{pmatrix}}_{} + \underbrace{\begin{pmatrix} \\ \end{pmatrix}}_{} \underbrace{\begin{pmatrix} \\ \end{pmatrix}}_{} \underbrace{\qquad}_{}$$

Example with simple univariate model

- $x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t, |g_x| < 1, g_{xx} > 0$
- Two fixed points: $\bar{x} = 0$ and $\bar{x} = (1 - g_x)/g_{xx}$

  $\hookrightarrow$ Once the model passes the second steady state it explodes
- Pruning: Decompose state vector into 1st- and 2nd-order effects

$$x_t = x_t^f + x_t^s = g_x x_{t-1}^f + g_x x_{t-1}^s + g_{xx} \left(x_{t-1}^f\right)^2 + 2 g_{xx} \left(x_{t-1}^f x_{t-1}^s\right)^2 + g_{xx} \left(x_{t-1}^s\right)^2 + g_u u_t$$

- Stable solution: Prune terms that contain $x_t^f x_t^s$ and $(x_t^s)^2$

$$x_t^f = g_x x_{t-1}^f + g_u u_t, \quad x_t^s = g_x x_{t-1}^s + g_{xx} (x_{t-1}^f)^2,$$

- Pruned solution can be rewritten as a stable state-space system

$$\underbrace{\begin{pmatrix} x_t^f \\ \end{pmatrix}}_{} = \underbrace{\begin{pmatrix} g_x & 0 & 0 \\ & & \end{pmatrix}}_{} \underbrace{\begin{pmatrix} x_{t-1}^f \\ \end{pmatrix}}_{} + \underbrace{\begin{pmatrix} g_u & 0 & 0 \\ & & \end{pmatrix}}_{} \underbrace{\begin{pmatrix} u_t \\ \end{pmatrix}}_{} \underbrace{\qquad}_{}$$

Example with simple univariate model

- $x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t, |g_x| < 1, g_{xx} > 0$
- Two fixed points: $\bar{x} = 0$ and $\bar{x} = (1 - g_x)/g_{xx}$

  $\hookrightarrow$ Once the model passes the second steady state it explodes
- Pruning: Decompose state vector into 1st- and 2nd-order effects

$$x_t = x_t^f + x_t^s = g_x x_{t-1}^f + g_x x_{t-1}^s + g_{xx} \left(x_{t-1}^f\right)^2 + 2 g_{xx} \left(x_{t-1}^f x_{t-1}^s\right)^2 + g_{xx} \left(x_{t-1}^s\right)^2 + g_u u_t$$

- Stable solution: Prune terms that contain $x_t^f x_t^s$ and $(x_t^s)^2$

$$x_t^f = g_x x_{t-1}^f + g_u u_t, \quad x_t^s = g_x x_{t-1}^s + g_{xx} (x_{t-1}^f)^2,$$

- Pruned solution can be rewritten as a stable state-space system

$$\underbrace{\begin{pmatrix} x_t^s \end{pmatrix}}_{} = \underbrace{\begin{pmatrix} 0 & g_x & g_{xx} \end{pmatrix}}_{} \underbrace{\begin{pmatrix} x_{t-1}^s \\ x_{t-1}^{f^2} \end{pmatrix}}_{} + \underbrace{\begin{pmatrix} 0 & 0 & 0 \end{pmatrix}}_{} \underbrace{\begin{pmatrix} \end{pmatrix}}_{} \underbrace{\qquad}_{}$$

Example with simple univariate model

- $x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t, |g_x| < 1, g_{xx} > 0$
- Two fixed points: $\bar{x} = 0$ and $\bar{x} = (1 - g_x)/g_{xx}$

  $\hookrightarrow$ Once the model passes the second steady state it explodes
- Pruning: Decompose state vector into 1st- and 2nd-order effects

$$x_t = x_t^f + x_t^s = g_x x_{t-1}^f + g_x x_{t-1}^s + g_{xx} \left(x_{t-1}^f\right)^2 + 2 g_{xx} \left(x_{t-1}^f x_{t-1}^s\right)^2 + g_{xx} \left(x_{t-1}^s\right)^2 + g_u u_t$$

- Stable solution: Prune terms that contain $x_t^f x_t^s$ and $(x_t^s)^2$

$$x_t^f = g_x x_{t-1}^f + g_u u_t, \quad x_t^s = g_x x_{t-1}^s + g_{xx} (x_{t-1}^f)^2,$$
$$(x_t^f)^2 = g_x^2 \left(x_{t-1}^f\right)^2 + 2 g_x g_u x_{t-1}^f u_t + g_u^2 u_t^2$$

- Pruned solution can be rewritten as a stable state-space system

$$\underbrace{\begin{pmatrix} \\ \\ x_t^{f^2} \end{pmatrix}}_{} = \underbrace{\begin{pmatrix} & & \\ & & \\ 0 & 0 & g_x^2 \end{pmatrix}}_{} \underbrace{\begin{pmatrix} \\ \\ x_{t-1}^{f^2} \end{pmatrix}}_{} + \underbrace{\begin{pmatrix} & & \\ & & \\ 0 & 2 g_x g_u & g_u^2 \end{pmatrix}}_{} \underbrace{\begin{pmatrix} x_{t-1}^f u_t \\ \\ u_t^2 \end{pmatrix}}_{} \underbrace{\qquad}_{}$$

Example with simple univariate model

- $x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t, |g_x| < 1, g_{xx} > 0$
- Two fixed points: $\bar{x} = 0$ and $\bar{x} = (1 - g_x)/g_{xx}$

  $\hookrightarrow$ Once the model passes the second steady state it explodes
- Pruning: Decompose state vector into 1st- and 2nd-order effects

$$x_t = x_t^f + x_t^s = g_x x_{t-1}^f + g_x x_{t-1}^s + g_{xx} \left(x_{t-1}^f\right)^2 + 2 g_{xx} \left(x_{t-1}^f x_{t-1}^s\right)^2 + g_{xx} \left(x_{t-1}^s\right)^2 + g_u u_t$$

- Stable solution: Prune terms that contain $x_t^f x_t^s$ and $(x_t^s)^2$

$$x_t^f = g_x x_{t-1}^f + g_u u_t, \quad x_t^s = g_x x_{t-1}^s + g_{xx} (x_{t-1}^f)^2,$$
$$(x_t^f)^2 = g_x^2 \left(x_{t-1}^f\right)^2 + 2 g_x g_u x_{t-1}^f u_t + g_u^2 u_t^2$$

- Pruned solution can be rewritten as a stable state-space system

$$\underbrace{\begin{pmatrix} x_t^f \\ x_t^s \\ x_t^{f^2} \end{pmatrix}}_{} = \underbrace{\begin{pmatrix} g_x & 0 & 0 \\ 0 & g_x & g_{xx} \\ 0 & 0 & g_x^2 \end{pmatrix}}_{} \underbrace{\begin{pmatrix} x_{t-1}^f \\ x_{t-1}^s \\ x_{t-1}^{f^2} \end{pmatrix}}_{} + \underbrace{\begin{pmatrix} g_u & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 2 g_x g_u & g_u^2 \end{pmatrix}}_{} \underbrace{\begin{pmatrix} u_t \\ x_{t-1}^f u_t \\ u_t^2 \end{pmatrix}}_{} \underbrace{\qquad}_{}$$

Example with simple univariate model

- $x_t = g_x x_{t-1} + g_{xx} x_{t-1}^2 + g_u u_t, |g_x| < 1, g_{xx} > 0$
- Two fixed points: $\bar{x} = 0$ and $\bar{x} = (1 - g_x)/g_{xx}$

  $\hookrightarrow$ Once the model passes the second steady state it explodes
- Pruning: Decompose state vector into 1st- and 2nd-order effects

$$x_t = x_t^f + x_t^s = g_x x_{t-1}^f + g_x x_{t-1}^s + g_{xx} \left(x_{t-1}^f\right)^2 + 2 g_{xx} \left(x_{t-1}^f x_{t-1}^s\right)^2 + g_{xx} \left(x_{t-1}^s\right)^2 + g_u u_t$$

- Stable solution: Prune terms that contain $x_t^f x_t^s$ and $(x_t^s)^2$

$$x_t^f = g_x x_{t-1}^f + g_u u_t, \quad x_t^s = g_x x_{t-1}^s + g_{xx} (x_{t-1}^f)^2,$$
$$(x_t^f)^2 = g_x^2 \left(x_{t-1}^f\right)^2 + 2 g_x g_u x_{t-1}^f u_t + g_u^2 u_t^2$$

- Pruned solution can be rewritten as a stable state-space system

$$\underbrace{\begin{pmatrix} x_t^f \\ x_t^s \\ x_t^{f^2} \end{pmatrix}}_{z_t} = \underbrace{\begin{pmatrix} g_x & 0 & 0 \\ 0 & g_x & g_{xx} \\ 0 & 0 & g_x^2 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_{t-1}^f \\ x_{t-1}^s \\ x_{t-1}^{f^2} \end{pmatrix}}_{z_{t-1}} + \underbrace{\begin{pmatrix} g_u & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 2 g_x g_u & g_u^2 \end{pmatrix}}_{B} \underbrace{\begin{pmatrix} u_t \\ x_{t-1}^f u_t \\ u_t^2 - \sigma_u^2 \end{pmatrix}}_{\xi_t} + \underbrace{\begin{pmatrix} 0 \\ 0 \\ g_u^2 \sigma_u^2 \end{pmatrix}}_{c}$$

Pruned state-space

Given an extended state vector z_t and an extended vector of innovations ξ_t, the pruned solution of a DSGE model can be rewritten as a linear time-invariant state-space system:

$$z_t = c + Az_{t-1} + B\xi_t$$
$$y_t = \bar{y} + d + Cz_{t-1} + D\xi_t$$

Pruned state-space

Given an extended state vector z_t and an extended vector of innovations ξ_t, the pruned solution of a DSGE model can be rewritten as a linear time-invariant state-space system:

$$z_t = c + Az_{t-1} + B\xi_t$$
$$y_t = \bar{y} + d + Cz_{t-1} + D\xi_t$$

- Same procedure for higher-order approximations

Pruned state-space

Given an extended state vector z_t and an extended vector of innovations ξ_t, the pruned solution of a DSGE model can be rewritten as a linear time-invariant state-space system:

$$z_t = c + Az_{t-1} + B\xi_t$$
$$y_t = \bar{y} + d + Cz_{t-1} + D\xi_t$$

- Same procedure for higher-order approximations
- Straightforward to compute moments, cumulants and polyspectra

Pruned state-space

Given an extended state vector z_t and an extended vector of innovations ξ_t, the pruned solution of a DSGE model can be rewritten as a linear time-invariant state-space system:

$$z_t = c + Az_{t-1} + B\xi_t$$
$$y_t = \bar{y} + d + Cz_{t-1} + D\xi_t$$

- Same procedure for higher-order approximations
- Straightforward to compute moments, cumulants and polyspectra
- Note: Even if u_t is Gaussian, ξ_t is not!

Pruned state-space

Given an extended state vector z_t and an extended vector of innovations ξ_t, the pruned solution of a DSGE model can be rewritten as a linear time-invariant state-space system:

$$z_t = c + Az_{t-1} + B\xi_t$$
$$y_t = \bar{y} + d + Cz_{t-1} + D\xi_t$$

- Same procedure for higher-order approximations
- Straightforward to compute moments, cumulants and polyspectra
- Note: Even if u_t is Gaussian, ξ_t is not!

$\hookrightarrow$ Higher-order statistics (HOS) may contain additional information for estimation and identification

Identification of nonlinear and non-Gaussian DSGE models

Diagnostics

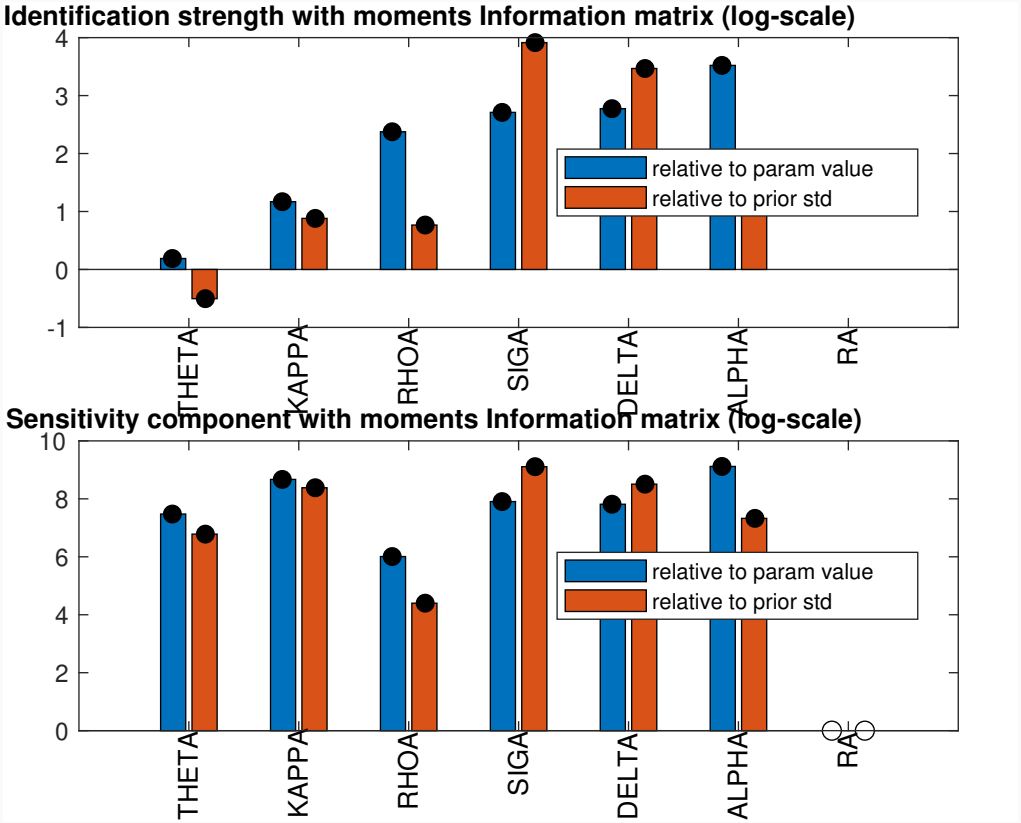
Local identification diagnostics

- We can use the same diagnostics based on moments or the spectrum to detect local identification issues
- Note that first and second moments and the power spectrum are computed from the pruned state-space system
- Higher-order moments and polyspectra might be also considered (not yet in Dynare)
- Identification Strength
  - Asymptotic information matrix is not available analytically, but covariance of moments is approximated by simulation from pruned state space
  - Bayesian Learning Rate indicator is readily available; however, one needs to use a nonlinear Kalman filter or a particle filter to evaluate the likelihood

investadjust.mod: identification(order=2)

```
======== Identification Analysis ========
Based on Pruned State Space System (order=2)
...
Testing prior mean
...
REDUCED-FORM:
!!!WARNING!!!
The rank of Tau (Jacobian of steady state and reduced-form solution matrices) is deficient!

RA is not identified!

MINIMAL SYSTEM (KOMUNJER AND NG, 2011):
!!!WARNING!!!
The rank of Deltabar (Jacobian of first-order minimal system and second-
order accurate mean) is deficient!

RA is not identified!

SPECTRUM (MUTSCHLER, 2015):
!!!WARNING!!!
The rank of Gbar (Jacobian of mean and spectrum) is deficient!

RA is not identified!

MOMENTS (MUTSCHLER, 2015):
!!!WARNING!!!
The rank of Mbar (Jacobian of first two moments) is deficient!

RA is not identified!

==== Identification analysis completed ====
```

Identification Strength based on Simulated Moments from Pruned State Space



Concluding Remarks

Concluding Remarks

- Identifiability is a model property that depends on the choices a modeler makes: (1) choice of observables, (2) functional specifications, (3) model features and (4) choice of structural shocks (Ivashchenko and Mutschler, 2019)
- Fix obvious identification failures: calibrate or (better) reparameterize!
- Change your model slighty, e.g. `investadjust.mod` model:
  - use intertemporal adjustment costs based on investment growth

$$K_t = (1-\delta)K_{t-1} + I_t\left(1 - S\left(\frac{I_t}{I_t}\right)\right)$$

  - add capital utilization
  - add investment-specific technological shock (best for strength)
  - include labor choice
- Larger models tend to be theoretically identified, but suffer from weak identification
  - Robust inference under possible weak identification (Dufour et al., 2009, 2013; Kleibergen and Mavroeidis, 2009; Mavroeidis, 2005, 2010; Guerron-Quintana et al.,2013; Andrews and Mikusheva, 2014; Qu, 2014)
  - wide confidence intervals for weakly identified parameters can be accompanied by narrow bands for IRFs
- Nonlinear or non-Gaussian approach might enrich identifiability and model dynamics (but comes at a price)