Identification analysis and higher-order approximation of DSGE models

Willi Mutschler

Introduction

Introduction

Identification Problem

- distinct parameter values do not lead to a distinct probability distribution of data (in particular moments and spectra)

$$p(Y|\theta) = p(Y|\theta_0) \nRightarrow \theta = \theta_0$$

- lack of identification leads to wrong conclusions from calibration, estimation and inference

Introduction

Identification Problem

- distinct parameter values do not lead to a distinct probability distribution of data (in particular moments and spectra)

$$p(Y|\theta) = p(Y|\theta_0) \nRightarrow \theta = \theta_0$$

- lack of identification leads to wrong conclusions from calibration, estimation and inference
- in practice, 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 asymptotics yield poor approximations

Introduction

Identification Problem

- distinct parameter values do not lead to a distinct probability distribution of data (in particular moments and spectra)

$$p(Y|\theta) = p(Y|\theta_0) \nRightarrow \theta = \theta_0$$

- lack of identification leads to wrong conclusions from calibration, estimation and inference
- in practice, 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 asymptotics yield poor approximations
- BUT: identifiability is a model property and can be analyzed without actually estimating the model
- lack of vs. strength of identification

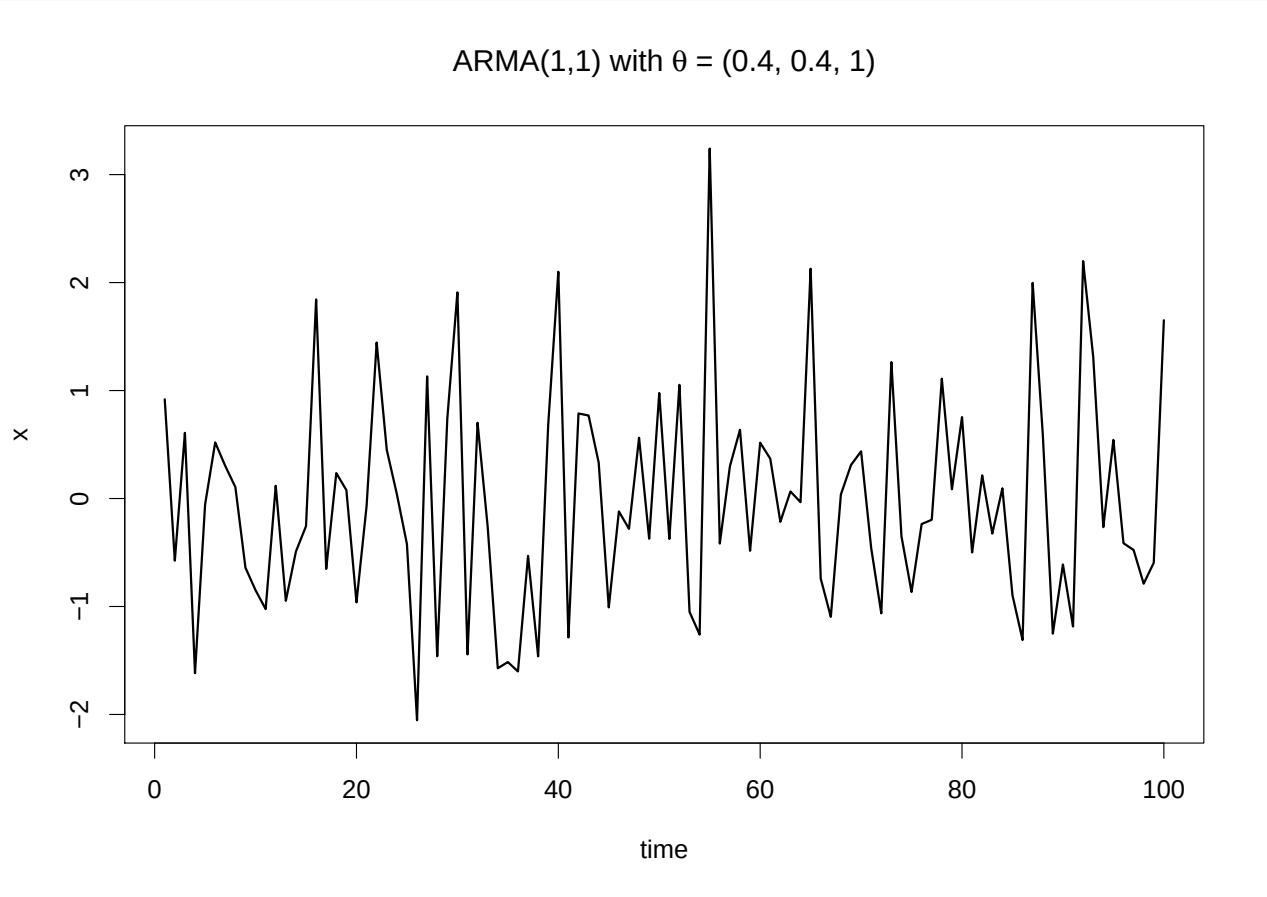
Example (1): ARMA(1,1)

Example (1): ARMA(1,1) (I)

- consider the following ARMA(1,1)-process

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

with parameter vector $\theta = (\phi_1, \phi_2, \sigma)'$:

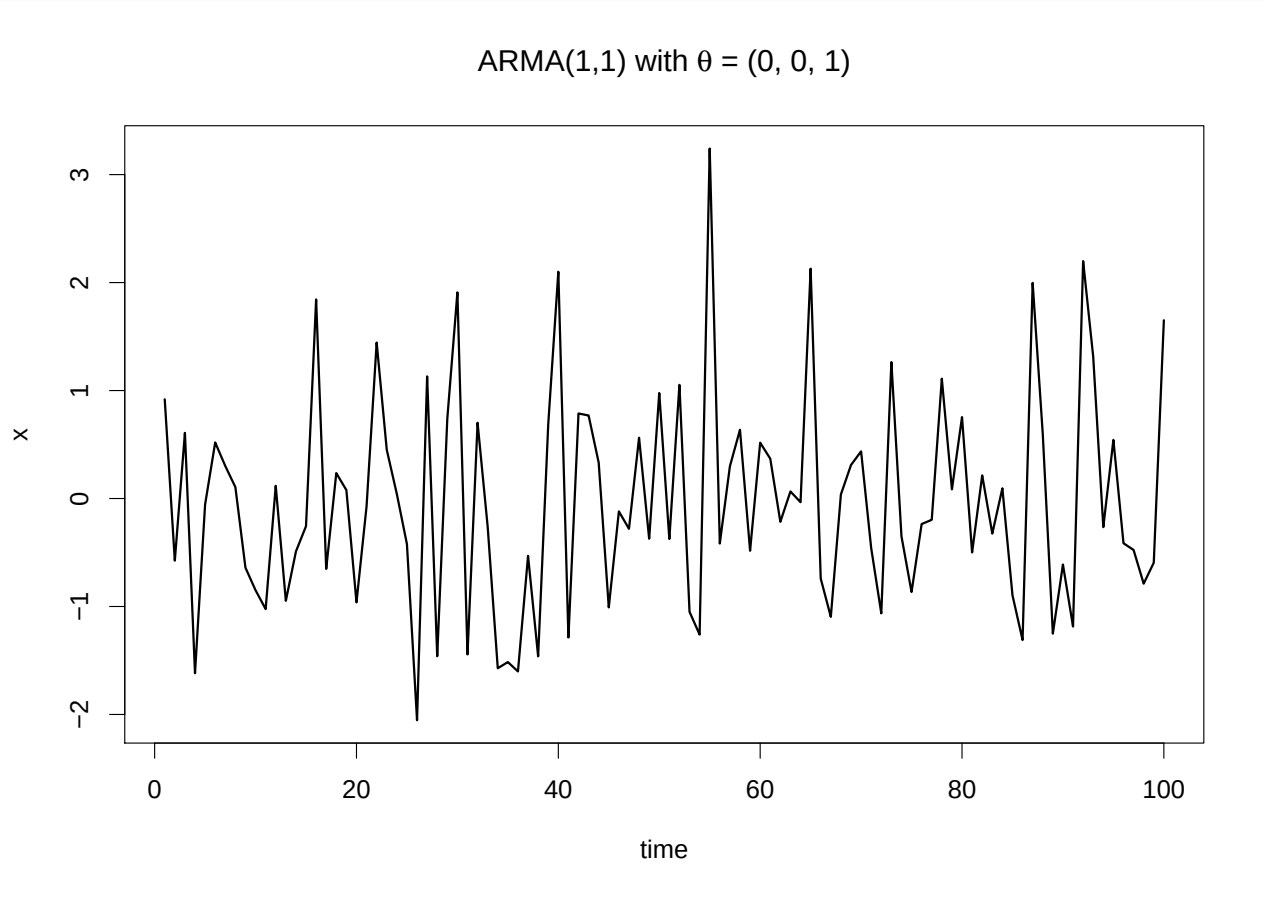


Example (1): ARMA(1,1) (I)

- consider the following ARMA(1,1)-process

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

with parameter vector $\theta = (\phi_1, \phi_2, \sigma)'$:



Example (1): ARMA(1,1) (II)

- autocovariance function: $\Gamma = (\gamma_0, \gamma_1, \ldots, \gamma_h)$ with

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

$$\gamma_1 = \frac{(\phi_1 - \phi_2)(1 - \phi_1\phi_2)\sigma^2}{1 - \phi_1^2},$$

$$\gamma_h = \phi_1\gamma_{h-1}$$

Example (1): ARMA(1,1) (II)

- autocovariance function: $\Gamma = (\gamma_0, \gamma_1, \ldots, \gamma_h)$ with

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

$$\gamma_1 = \frac{(\phi_1 - \phi_2)(1 - \phi_1\phi_2)\sigma^2}{1 - \phi_1^2},$$

$$\gamma_h = \phi_1\gamma_{h-1}$$

- rank of Jacobian of Γ w.r.t θ is not full at $\phi_1 = \phi_2$

Example (1): ARMA(1,1) (II)

- autocovariance function: $\Gamma = (\gamma_0, \gamma_1, \ldots, \gamma_h)$ with

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

$$\gamma_1 = \frac{(\phi_1 - \phi_2)(1 - \phi_1\phi_2)\sigma^2}{1 - \phi_1^2},$$

$$\gamma_h = \phi_1\gamma_{h-1}$$

- rank of Jacobian of Γ w.r.t θ is not full at $\phi_1 = \phi_2$
- similar argument applies to the spectral density matrix

Example (2): Simple DSGE Model

Example (2): Simple DSGE Model(I)

- consider a simple purely forward looking **log-linearized** model

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

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

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

Example (2): Simple DSGE Model(I)

- consider a simple purely forward looking **log-linearized** 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

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

Example (2): Simple DSGE Model(I)

- consider a simple purely forward looking **log-linearized** 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 of the model, i.e. Eigenvalues of $A_0^{-1}A_1$ lie within unit circle, implies:

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

Example (2): Simple DSGE Model(II)

- solution/data-generating-process/reduced-form

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

Example (2): Simple DSGE Model(II)

- solution/data-generating-process/reduced-form

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

- some insights

Example (2): Simple DSGE Model(II)

- solution/data-generating-process/reduced-form

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

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

Example (2): Simple DSGE Model(II)

- solution/data-generating-process/reduced-form

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

- some insights
  - some parameters (β) do not enter solution, thus do not enter the likelihood (or any other objective)
  - if we fix (or calibrate) β, all parameters are identifiable

Example (2): Simple DSGE Model(II)

- solution/data-generating-process/reduced-form

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

- some insights
  - some parameters (β) do not enter solution, thus do not enter the likelihood (or any other objective)
  - if we fix (or calibrate) β, all parameters are identifiable
  - some parameters enter as products, e.g. $\frac{\kappa\psi}{\tau}$, identifying them separately may be impossible (depending on choice of observables)

Example (2): Simple DSGE Model(II)

- solution/data-generating-process/reduced-form

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

- some insights
  - some parameters (β) do not enter solution, thus do not enter the likelihood (or any other objective)
  - if we fix (or calibrate) β, all parameters are identifiable
  - some parameters enter as products, e.g. $\frac{\kappa\psi}{\tau}$, identifying them separately may be impossible (depending on choice of observables)
  - κ is already the product of several other structural parameters (Calvo or Rotemberg)

Example (2): Simple DSGE Model(II)

- solution/data-generating-process/reduced-form

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

- some insights
  - some parameters (β) do not enter solution, thus do not enter the likelihood (or any other objective)
  - if we fix (or calibrate) β, all parameters are identifiable
  - some parameters enter as products, e.g. $\frac{\kappa\psi}{\tau}$, identifying them separately may be impossible (depending on choice of observables)
  - 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 structural parameters, i.e. parameter space is not *variation free*

Example (2): Simple DSGE Model (I)

concerned with *injectivity* of two mappings

- uniqueness of solution
  - $\hookrightarrow$ from the structural parameters to the reduced-form parameters
- uniqueness of probability distribution
  - $\hookrightarrow$ from the solution to observable data

Example (2): Simple DSGE Model (II)

Local Identification for Linearized Gaussian DSGE Models

- autocovariogram (Iskrev, 2010)
- spectral density (Qu and Tkachenko, 2012)
- control theory for minimal systems (Komunjer and Ng, 2011)

Global Identification for Linearized Gaussian DSGE Models

- Kullback-Leibler discrepancy (Qu and Tkachenko, 2017)

Weak Identification for Linearized Gaussian DSGE Models

- 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)

Example (2): Simple DSGE Model (II)

Local Identification for Linearized Gaussian DSGE Models

- autocovariogram (Iskrev, 2010)
- spectral density (Qu and Tkachenko, 2012)
- control theory for minimal systems (Komunjer and Ng, 2011)

Global Identification for Linearized Gaussian DSGE Models

- Kullback-Leibler discrepancy (Qu and Tkachenko, 2017)

Weak Identification for Linearized Gaussian DSGE Models

- 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)

Example (2): Simple DSGE Model (III)

VERY SIMILAR TO EXAMPLE (1)

Iskrev (2010)'s approach

- check whether derivative of the **theoretical mean, variance and autocovariogram** of observables w.r.t structural parameters has full rank

↪ *time domain approach*

Qu and Tkachenko (2012)'s approach

- check whether derivative of the **theoretical mean and spectrum** of observables w.r.t structural parameters has full rank

↪ *frequency domain approach*

Example (2): Simple DSGE Model (IV)

Koop, Pesaran and Smith (2013)'s approach

- 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*

Example (3): An and Schorfheide (2007)

Example (3): An and Schorfheide (2007) (I)

prototypical small-scale business cycle model

$$\underbrace{\left(\frac{C_t}{A_t}\right)^{-\tau}\frac{1}{A_t}}_{U_t^c} = \beta E_t\left[\frac{R_t}{\pi_{t+1}}\underbrace{\left(\frac{C_{t+1}}{A_{t+1}}\right)^{-\tau}\frac{1}{A_{t+1}}}_{U_{t+1}^c}\right] \tag{1}$$

$$1 = \frac{1}{\nu}\left[1-\left(\frac{C_t}{A_t}\right)^{\tau}\right] + \phi(\pi_t-\pi)\left[\left(1-\frac{1}{2\nu}\right)\pi_t + \frac{\pi}{2\nu}\right] - \phi\beta E_t\left[\left(\frac{C_{t+1}/A_{t+1}}{C_t/A_t}\right)^{-\tau}\frac{Y_{t+1}/A_{t+1}}{Y_t/A_t}(\pi_{t+1}-\pi)\pi_{t+1}\right] \tag{2}$$

$$Y_t = C_t + G_t + \frac{\phi}{2}(\pi_t-\pi)^2 Y_t \tag{3}$$

$$Y_t^* = (1-\nu)^{\frac{1}{\tau}} A_t \underbrace{\left(\frac{Y_t}{Y_t-G_t}\right)}_{g_t} \tag{4}$$

Example (3): An and Schorfheide (2007) (II)

- stochastic processes

$$\ln A_t = \ln \gamma + \ln A_{t-1} + \ln z_t$$
$$\ln z_t = \rho_z \ln z_{t-1} + \sigma_z \epsilon_{z,t} \tag{5}$$
$$\ln(g_t) = (1-\rho_g)\ln(g) + \rho_g \ln(g_{t-1}) + \sigma_g \epsilon_{g,t} \tag{6}$$

- monetary policy modeled by a Taylor Rule

$$R_t = R_t^{*\,1-\rho_r} R_{t-1}^{\rho_r} e^{\sigma_r \epsilon_{r,t}} \tag{7}$$

- two specifications for R_t^* are considered

$$R_t^* = \begin{cases} r\pi^* \left(\frac{\pi_t}{\pi^*}\right)^{\psi_1} \left(\frac{Y_t}{Y_t^*}\right)^{\psi_2} & (\mathcal{TR}_1 \text{ flex-price rule}) \\ r\pi^* \left(\frac{\pi_t}{\pi^*}\right)^{\psi_1} \left(\frac{Y_t}{\gamma Y_{t-1}}\right)^{\psi_2} & (\mathcal{TR}_2 \text{ output-growth rule}) \end{cases}$$

- $\epsilon_{z,t} \sim \mathcal{N}(0,1)$, $\epsilon_{g,t} \sim \mathcal{N}(0,1)$ and $\epsilon_{r,t} \sim \mathcal{N}(0,1)$

Example (3): An and Schorfheide (2007) (III)

SIMILAR TO EXAMPLE (2)

- log-linearized model with $\widehat{y}_t$, $\widehat{\pi}_t$ and $\widehat{r}_t$ observable

$$\widehat{y}_t = E_t[\widehat{y}_{t+1}] + \widehat{g}_t - E_t[\widehat{g}_{t+1}] - \frac{1}{\tau}(\widehat{r}_t - E_t[\widehat{\pi}_{t+1}] - E_t[\widehat{z}_{t+1}])$$

$$\widehat{\pi}_t = \beta E_t[\widehat{\pi}_{t+1}] + \underbrace{\tau \frac{1-\nu}{\nu \pi^{*2} \phi}}_{\kappa}(\widehat{y}_t - \widehat{g}_t)$$

$$\widehat{g}_t = \rho_g \widehat{g}_{t-1} + \sigma_g \epsilon_{g,t}$$

$$\widehat{z}_t = \rho_z \widehat{z}_{t-1} + \sigma_z \epsilon_{z,t}$$

$$\widehat{r}_t = \rho_r \widehat{r}_{t-1} + (1-\rho_r)\psi_1 \widehat{\pi}_t + (1-\rho_r)\psi_2 \begin{cases} (\widehat{y}_t - \widehat{g}_t) \\ (\widehat{y}_t - \widehat{y}_{t-1} + \widehat{z}_t) \end{cases} + \sigma_r \epsilon_{r,t}$$

- $\theta = (\tau, \phi, \nu, \psi_1, \psi_2, \rho_r, \rho_g, \rho_z, \beta, \pi^*, \gamma, \sigma_r, \sigma_g, \sigma_z)$

Example (3): An and Schorfheide (2007) (III)

SIMILAR TO EXAMPLE (2)

- log-linearized model with $\widehat{y}_t$, $\widehat{\pi}_t$ and $\widehat{r}_t$ observable

$$\widehat{y}_t = E_t[\widehat{y}_{t+1}] + \widehat{g}_t - E_t[\widehat{g}_{t+1}] - \frac{1}{\tau}(\widehat{r}_t - E_t[\widehat{\pi}_{t+1}] - E_t[\widehat{z}_{t+1}])$$

$$\widehat{\pi}_t = \beta E_t[\widehat{\pi}_{t+1}] + \underbrace{\tau \frac{1-\nu}{\nu \pi^{*2} \phi}}_{\kappa}(\widehat{y}_t - \widehat{g}_t)$$

$$\widehat{g}_t = \rho_g \widehat{g}_{t-1} + \sigma_g \epsilon_{g,t}$$

$$\widehat{z}_t = \rho_z \widehat{z}_{t-1} + \sigma_z \epsilon_{z,t}$$

$$\widehat{r}_t = \rho_r \widehat{r}_{t-1} + (1-\rho_r)\psi_1 \widehat{\pi}_t + (1-\rho_r)\psi_2 \begin{cases} (\widehat{y}_t - \widehat{g}_t) \\ (\widehat{y}_t - \widehat{y}_{t-1} + \widehat{z}_t) \end{cases} + \sigma_r \epsilon_{r,t}$$

- $\theta = (\tau, \phi, \nu, \psi_1, \psi_2, \rho_r, \rho_g, \rho_z, \beta, \pi^*, \gamma, \sigma_r, \sigma_g, \sigma_z)$

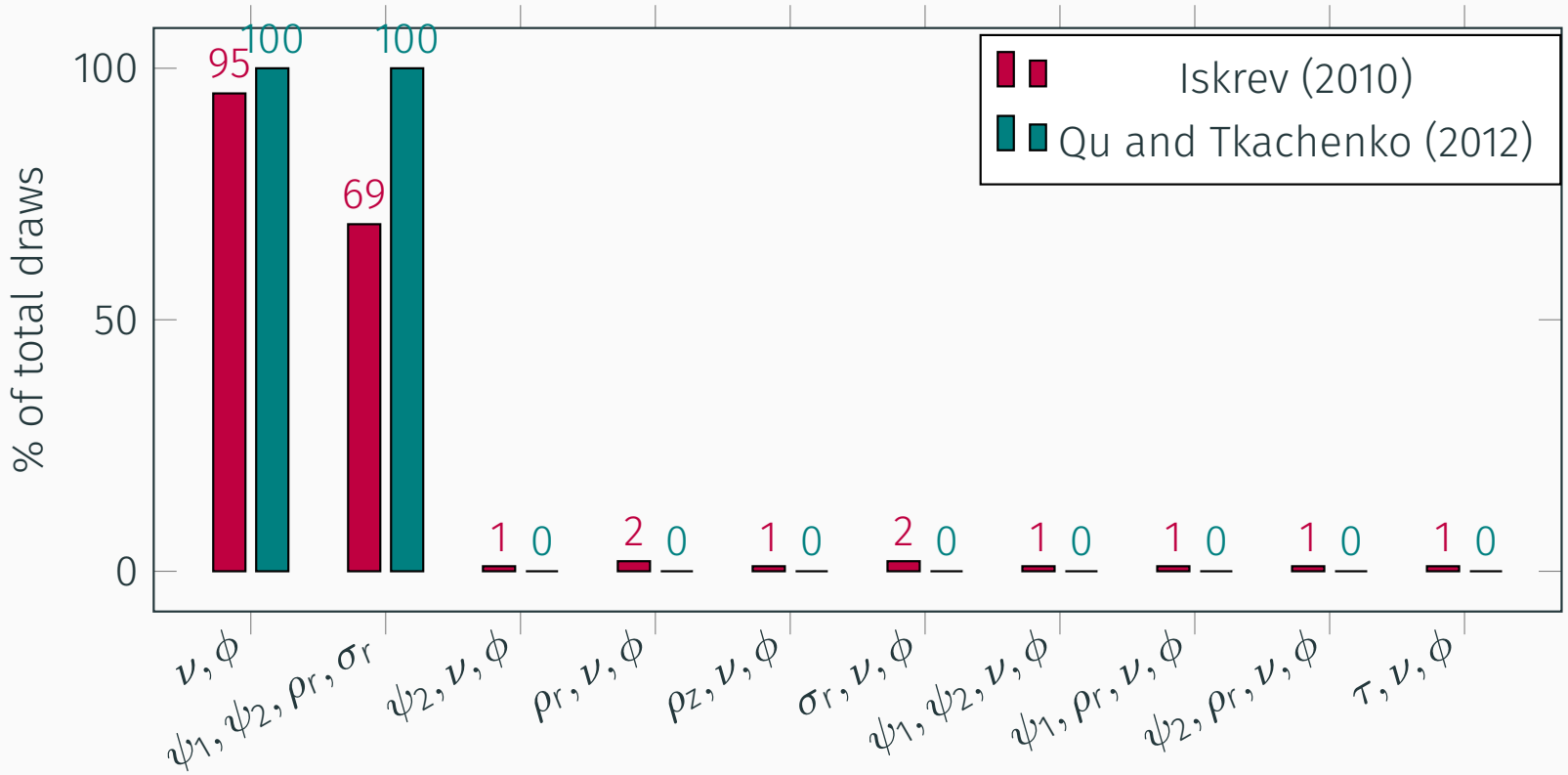
Obvious identification failure

$\hookrightarrow$ elasticity of demand $\frac{1}{\nu}$ and price stickiness ϕ are not jointly identifiable, but only composite parameter κ

Example (3): Identification Analysis

Example (3): Identification Analysis (I)

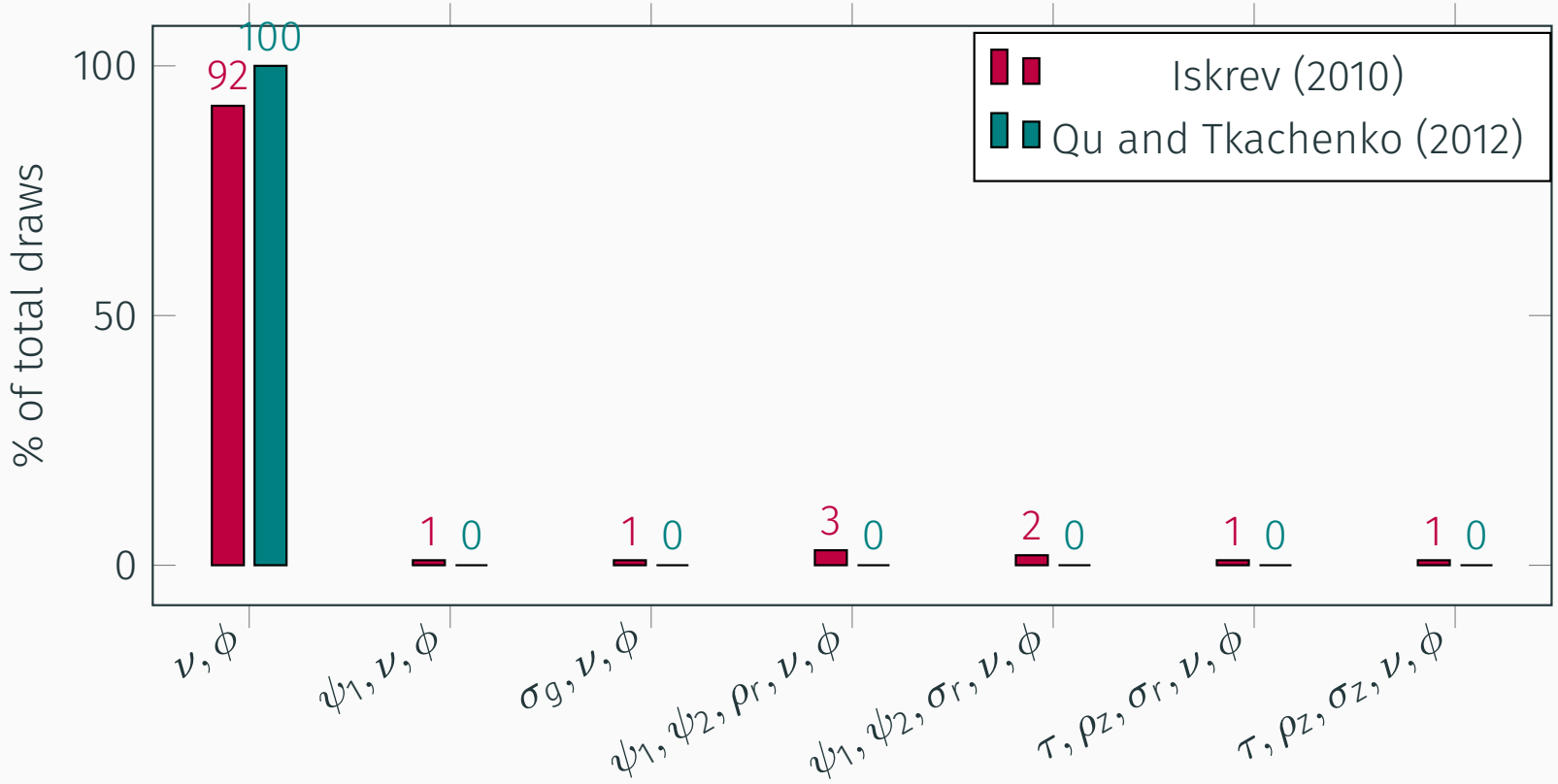
Figure 1: Non-Identification Sets: Linearized Model $\mathcal{TR}_1$



Notes: Identification results for Iskrev (2010) and Qu and Tkachenko (2012) for 100 draws from the prior domain using analytical derivatives with robust tolerance level, 30 lags and 10000 subintervalls. Sets by brute-force method. Source: Mutschler (2015)
$\mathcal{TR}_1$ corresponds to the output-gap specification of the Taylor-rule.

Example (3): Identification Analysis (II)

Figure 2: Non-Identification Sets: Linearized Model $\mathcal{TR}_2$



Notes: Identification results for Iskrev (2010) and Qu and Tkachenko (2012) for 100 draws from the prior domain using analytical derivatives with robust tolerance level, 30 lags and 10000 subintervalls. Sets by brute-force method. Source: Mutschler (2015)
$\mathcal{TR}_2$ corresponds to the output-growth specification of the Taylor-rule.

Example (3): Identification Analysis (III)

Table 1: Average Posterior Precisions: Linearized Model $\mathcal{TR}_1$

Obs	PARAMETERS									
	τ	ψ_1	ψ_2	$r^{(A)}$	ρ_g	ρ_z	$100\sigma_r$	$100\sigma_g$	$100\sigma_z$	ν
	HESSIAN METHOD - GAUSSIAN PRIORS									
20	5.07	5.09	8.58	5.01	3.80	98.85	36.40	9.61	17.40	368.01
50	2.04	2.04	4.12	2.11	4.88	46.96	36.25	7.70	10.68	307.86
100	1.03	1.03	3.05	1.15	3.33	56.86	43.00	7.13	11.71	264.77
1000	0.16	0.20	5.49	0.17	7.96	73.15	47.64	5.97	14.07	100.05
10000	0.03	0.15	5.04	0.19	5.71	28.69	46.59	4.87	9.36	29.69
	MCMC METHOD - GAUSSIAN PRIORS									
20	5.28	5.03	10.84	5.38	10.69	113.62	30.75	8.70	19.30	364.50
50	2.03	2.07	5.15	2.18	10.68	51.34	33.28	6.87	11.07	329.11
100	1.01	1.06	3.10	1.11	3.71	50.30	39.19	6.36	10.34	283.05
1000	0.14	0.10	1.40	0.15	7.54	65.27	43.20	5.55	12.93	84.67
10000	0.06	0.01	0.29	0.03	8.17	74.32	46.40	5.35	15.91	43.79

Notes: ϕ and ρ_r are fixed at true values. $\beta = \exp\left\{-r^{(A)}/400\right\}$.

Nelder-Mead simplex optimization routine for posterior mode and Hessian. MCMC method uses variances from draws of marginal posteriors, i.e. Random-Walk Metropolis-Hastings algorithm with 3 chains á 20000 draws. Gaussian proposal density initialized at mode and Hessian with scale parameter equal to 0.6, acceptance ratios lie in between 20%-35%. Gaussian priors correspond to using truncated independent normal distributions with mean set to the true value and standard deviation equal to 0.1.

Example (3): Identification Analysis (III)

Table 1: Average Posterior Precisions: Linearized Model $\mathcal{TR}_1$

Obs	PARAMETERS									
	τ	ψ_1	ψ_2	$r^{(A)}$	ρ_g	ρ_z	$100\sigma_r$	$100\sigma_g$	$100\sigma_z$	ν
	HESSIAN METHOD - GAUSSIAN PRIORS									
20	5.07	5.09	8.58	5.01	3.80	98.85	36.40	9.61	17.40	368.01
50	2.04	2.04	4.12	2.11	4.88	46.96	36.25	7.70	10.68	307.86
100	1.03	1.03	3.05	1.15	3.33	56.86	43.00	7.13	11.71	264.77
1000	0.16	0.20	5.49	0.17	7.96	73.15	47.64	5.97	14.07	100.05
10000	0.03	0.15	5.04	0.19	5.71	28.69	46.59	4.87	9.36	29.69
	MCMC METHOD - GAUSSIAN PRIORS									
20	5.28	5.03	10.84	5.38	10.69	113.62	30.75	8.70	19.30	364.50
50	2.03	2.07	5.15	2.18	10.68	51.34	33.28	6.87	11.07	329.11
100	1.01	1.06	3.10	1.11	3.71	50.30	39.19	6.36	10.34	283.05
1000	0.14	0.10	1.40	0.15	7.54	65.27	43.20	5.55	12.93	84.67
10000	0.06	0.01	0.29	0.03	8.17	74.32	46.40	5.35	15.91	43.79

Notes: ϕ and ρ_r are fixed at true values. $\beta = \exp\left\{-r^{(A)}/400\right\}$.

Nelder-Mead simplex optimization routine for posterior mode and Hessian. MCMC method uses variances from draws of marginal posteriors, i.e. Random-Walk Metropolis-Hastings algorithm with 3 chains á 20000 draws. Gaussian proposal density initialized at mode and Hessian with scale parameter equal to 0.6, acceptance ratios lie in between 20%-35%. Gaussian priors correspond to using truncated independent normal distributions with mean set to the true value and standard deviation equal to 0.1.

Computational and numerical issues

- Identification criteria generally come to the same conclusion, yet there are computational and numerical issues
  - Numerical instability of solution algorithm
  - Size of matrices: use different or robust tolerance level for ranks
  - If feasible: Use analytical rather than numerical derivatives and robust tolerance levels
  - Lag length (ISK), subintervals for integral of spectra (QT), filtering and speed (KPS)

Identification of nonlinear and non-Gaussian DSGE models

General model framework and solution

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 model framework and solution

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:

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

Pruning

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

Pruning

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 $\bar{x} = (1 - g_x)/g_{xx}$
- 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 $\bar{x} = (1 - g_x)/g_{xx}$
- 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 $\bar{x} = (1 - g_x)/g_{xx}$
- 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 $\bar{x} = (1 - g_x)/g_{xx}$
- 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 $\bar{x} = (1 - g_x)/g_{xx}$
- 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 $\bar{x} = (1 - g_x)/g_{xx}$
- 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 $\bar{x} = (1 - g_x)/g_{xx}$
- 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 $\bar{x} = (1 - g_x)/g_{xx}$
- 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 $\bar{x} = (1 - g_x)/g_{xx}$
- 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

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

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

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

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

Analytical derivatives

Analytical derivatives

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

Analytical derivatives

- We 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

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

Analytical derivatives

- We 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

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

- The same holds for the Jacobian

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

Analytical derivatives

- We 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

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

- The same holds for the Jacobian

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

- And the Hessian

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

Analytical derivatives

- Now it is straightforward (but very tedious) to derive the analytical derivatives of first-order, second-order and third-order solution matrices wrt parameters
- Once we have that, the pruned state-space can be used to compute analytical derivatives of first four moments or corresponding polyspectra
- On paper Kronecker products provide closed-form expressions
- Computationally generalized Sylvester equations are more efficient

Identification Criteria: Time Domain

Proposition Time Domain
Consider the pruned state-space of a nonlinear DSGE model. Let $q \leq T$ and assume that

$$\overline{m}(\theta, q) := \begin{pmatrix} \mu_y' & m_2(\theta, q)' & m_3(\theta, q)' & m_4(\theta, q)' \end{pmatrix}'$$

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 four cumulants (or moments) of y_t, if and only if

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

has a full column rank equal to the number of parameters for $q \leq T$.

Identification Criteria: Frequency Domain

Proposition Frequency Domain

Consider the pruned state-space of a nonlinear DSGE model. Assume that the power spectrum, bispectrum and trispectrum are continuous in $\omega \in [-\pi; \pi]$ and continuous and differentiable in $\theta \in \Theta$. Let

$$\begin{aligned}\overline{G}(\theta) = \mathrm{d}\left(\mu_y(\theta)\right)' \mathrm{d}\left(\mu_y(\theta)\right) + \int_{-\pi}^{\pi} \mathrm{d}\left(\mathcal{S}_{2,y}(\omega_1;\theta)\right)^* \mathrm{d}\left(\mathcal{S}_{2,y}(\omega_1;\theta)\right) d\omega_1 \\ + \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mathrm{d}\left(\mathcal{S}_{3,y}(\omega_1,\omega_2;\theta)\right)^* \mathrm{d}\left(\mathcal{S}_{3,y}(\omega_1,\omega_2;\theta)\right) d\omega_1 d\omega_2 \\ + \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mathrm{d}\left(\mathcal{S}_{4,y}(\omega_1,\omega_2,\omega_3;\theta)\right)^* \mathrm{d}\left(\mathcal{S}_{4,y}(\omega_1,\omega_2,\omega_3;\theta)\right) d\omega_1 d\omega_2 d\omega_3\end{aligned}$$

and $\theta_0 \in \Theta$ be a regular point. Furthermore, assume there is an open neighborhood of θ_0 in which $\overline{G}(\theta_0)$ has a constant rank. Then, θ is locally identifiable at a point θ_0 from the mean, power spectrum, bispectrum and trispectrum of y_t, if and only if $\overline{G}(\theta_0)$ is nonsingular, i.e. its rank is equal to the number of parameters.

Strength of Identification: Bayesian Learning Rate Indicator

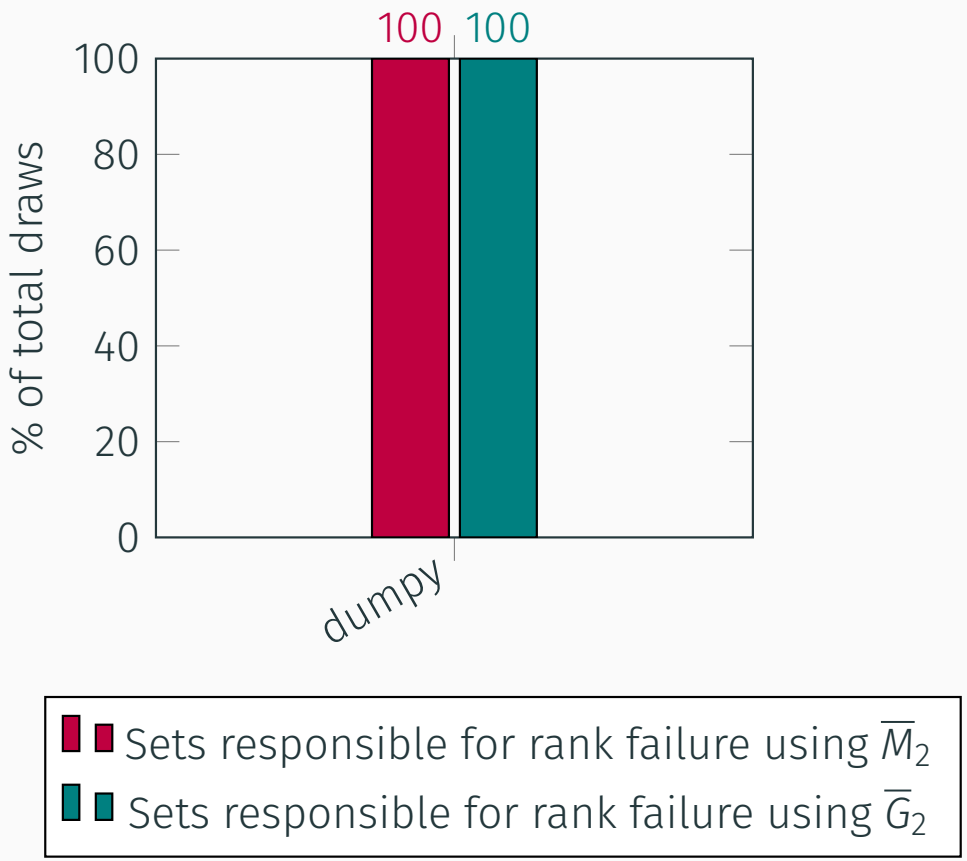
Koop, Pesaran and Smith (2013)'s approach

- 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* using a nonlinear Kalman filter or a particle filter to evaluate the likelihood

Identification analysis of An and Schorfheide (2007)

Figure 1: Non-identified sets



Notes: Identification results for 100 draws from the prior domain using analytical derivatives with robust tolerance level, $T = 30$ and $N = 10\,000$. Sets by brute-force method.

Concluding Remarks

Concluding Remarks

- identifiability is a model property that depends on functional specifications and choice of observables
- how to ”solve” lack of identification:
  - change your model (slightly), e.g. add shocks (preference, investment-specific technology), capital utilization, monetary policy rule, utility function, habit formation,...
  - nonlinear or non-Gaussian approach might enrich identifiability and model dynamics
- my own toolbox works for the framework of Schmitt-Grohé and Uribe (2004) and Andreasen, Fernández-Villaverde and Rubio-Ramírez (2018)
- extension of DYNARE's `identification(order=2|3,pruning)` is still work-in-progress

Appendix

Global Identification

- Global identification is, however, more difficult to verify than local identification.
- A very recent approach proposed by Qu and Tkachenko (2017) is to search for the solution of the Kullback-Leibler discrepancy

$$\Delta_{KL}(\theta|\theta_0) = -\int \log\left(\frac{p(Y|\theta)}{p(Y|\theta_0)}\right) p(Y|\theta_0)dY = 0$$

  by using a frequency domain transformation.
- If θ_0 is the unique solution, then the DSGE model is globally identified.
- However, finding the solution to this objective function is computationally challenging and only demonstrated for a stylized small scale linear DSGE model.