Econometrics with System priors

Michal Andrle, Miroslav Plašil (presenter)

CNB Research Open Day
21 May 2018

Disclaimer

Michal Andrle

The views expressed herein are those of the authors and should not be attributed to the International Monetary Fund, its Executive Board, or its management.

Miroslav Plašil

The views expressed here are those of the authors and do not necessarily reflect the position of the Czech National Bank.

Aims and scope

- To provide more nuanced and more general introduction to system priors (devised by Andrle and Beneš within the DSGE context)
- To demonstrate the generality of principles and its wide range of application
- To illustrate the use of system priors with a simple but practically relevant example
- ... to invite fellow researchers to jump on the bandwagon

What are system priors?

- Economically-meaningful priors about high-level model properties
  - impulse-response functions
  - variance error decompositions
  - frequency-domain properties
  - policy scenarios (sacrifice ratios, delayed MP response)
  - ...anything that can be computed with the model (forecasts...)
- Two layer approach that facilitates formulation of priors on both the parameter and model level
- Complement rather than substitute for traditional Bayesian setup

Why and when one should use system priors?

- In many complex models individual parameters are difficult to interpret.
- Reasonable priors for individual parameters may lead in sum to highly erratic priors about the overall model behavior.
  - Even „non-informative“ priors can be implicitly very informative in a highly undesirable way
  - Prior predictive analysis – which parameter priors “bite”?
- Policy makers only hold firm views about economic behavior.
  - Communication channel between modelers and policy makers

First glance at system priors

- Traditional Bayesian setup

$$p(\theta | Y;M) \propto L(Y|\theta;M) \times p_m(\theta)$$

- System priors setup

$$p(\theta | Y;M) \propto L(Y|\theta;M) \times [\, p_s(h(\theta);M) \times p_m(\theta)]$$

  - $p_m(\theta)$ – priors on individual parameters
  - $p_s(h(\theta);M)$ – system priors „add-on“
  - $[p_s(h(\theta);M) \times p_m(\theta)]$ – composite prior enabling to implement views on elements in both layers

How to understand system priors I

- **(Non-conjugate) dummy observation prior**
  - Instead of inserting dummy observations into the dataset, create a dummy/artificial likelihood for the auxiliary model that summarizes the information in the dummy observations
- $[\, p_s(h(\theta);M) \times \; p_m(\theta)] \equiv likelihood \times prior\ on\ parameters$
  - Structure of the auxiliary model corresponds to the high-level property of interest: $h(\theta;M)$ + error term (set of stochastic restrictions)
- Posterior inference is obtained by updating priors on individual parameters twice:
  - first with artificial likelihood of the auxiliary model (system priors)
  - second with real likelihood based on observed data

How to understand system priors II

- **Penalized likelihood problem**
  - Taking logs of the RHS...

$$p(\theta\,|\,Y;M) \propto L(Y|\theta;M) \times [\,p_s(h(\theta);M) \times p_m(\theta)]$$

  - ... one obtains

$$\log(L(Y|\theta;M)) + \log\left(p_m(\theta)\right) + \log(p_s(h(\theta);M)$$

- Finding the mode of the posterior distribution is a traditional maximum likelihood approach with additional penalties that “regularize” the problem

- Penalty terms are nothing new in econometrics
  - ridge regression
  - lasso
  - many others...

Related literature I

- A desire for priors on model properties in not new, however most of the existing attempts only solve *ad hoc* problems
  - priors only solve specific a problem at hand (e.g. steady-state priors – Villani, 2005; priors on impulse responses – Dwyer, 1998, Kocięcki, 2012; long-run priors – Giannone et al., 2016; priors on frequencies – Planas et al., 2008)
  - priors only take specific form (usually gaussian priors)

- More general approaches
  - *Feature of interest priors*: Hollifield et al. (2003) – this approach is conceptually identical to system priors
  - *Priors on observables*: Jarociński and Marcet (2013)

Related literature II

Comparison of our approach with that of Jarociński and Marcet

- Both approaches can be used to solve similar problems, however they differ in concept (and flexibility & versatility).

- Both approaches need to solve the inverse problem:

- Jarociński and Marcet
  - Priors on high-level features -> Priors on observables -> Fredholm equation/fixed point solution -> implied priors on individual parameters -> bayesian update (likelihood) -> posterior distribution

- System priors
  - Priors on individual parameters -> bayesian update (artificial likelihood) -> bayesian update (likelihood) -> posterior distribution

Illustrative example

- Stationary AR(2) process with additional belief that most of its variance is generated by business-cycle frequencies
  - AR(2) is a very simple case, but the process can exhibit non-trivial dynamics
- We use the example only as an illustration, however it can be quite useful for empirical work
  - output gaps are frequently modelled as the AR(2) process: (see e.g. Watson, 1986, Clark, 1987, Kuttner, 1994, Planas et al., 2008, Jarociński and Lenza, 2016 and many others)
  - the same goes for inflation gaps (Clark and Doh, 2014)
  - ...or unemployment gaps (Chan et al., 2016)

Illustrative example

- Full (but still simple) model may look something like this...

$$
\begin{aligned}
(\pi_t - \tau_t^{\pi}) &= \rho_t^{\pi} \left(\pi_{t-1} - \tau_{t-1}^{\pi}\right) + \lambda_t \left(u_t - \tau_t^{u}\right) + \varepsilon_t^{\pi} \\
(u_t - \tau_t^{u}) &= \rho_1^{u} \left(u_{t-1} - \tau_{t-1}^{u}\right) + \rho_2^{u} \left(u_{t-2} - \tau_{t-2}^{u}\right) + \varepsilon_t^{u} \\
\tau_t^{\pi} &= \tau_{t-1}^{\pi} + \varepsilon_t^{\tau\pi} \\
\tau_t^{u} &= \tau_{t-1}^{u} + \varepsilon_t^{\tau u} \\
\rho_t^{\pi} &= \rho_{t-1}^{\pi} + \varepsilon_t^{\rho\pi} \\
\lambda_t &= \lambda_{t-1} + \varepsilon_t^{\lambda}
\end{aligned}
$$

Source: A Bounded Model of Time Variation in Trend Inflation, NAIRU and the Phillips Curve
Chan, Koop and Potter (2016)

Illustrative example

If the AR(2) is used to capture some "business-cycle gap" variable, what are the options to consider?

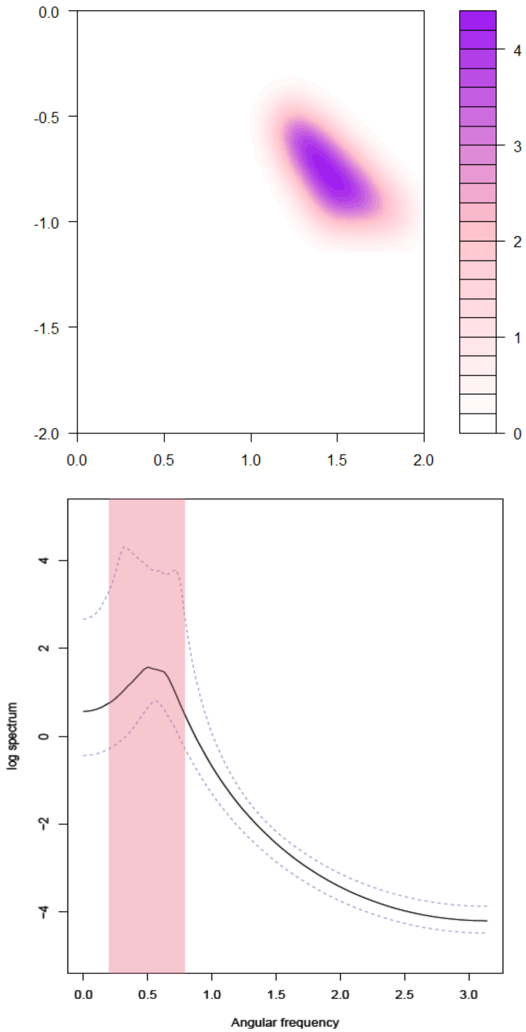
- basic Gaussian option: Chan and Grant(2017); earlier versions of Chan et al. (2015)
- model reparametrization: Planas et al. (2008)
- more refined Gaussian option: Chan et al. (2015); Grant and Chan (2017); Lenza and Jarociński (2016)
- System priors based on the business-to-total-variance ratio
  - at least 60% of variance comes from business cycle frequencies
  - the ratio follows some distribution [Be(15,5) is used in the paper]

$$ratio = \int_a^b S_y(w)\mathrm{d}w \Big/ \int S_y(w)\mathrm{d}w,$$

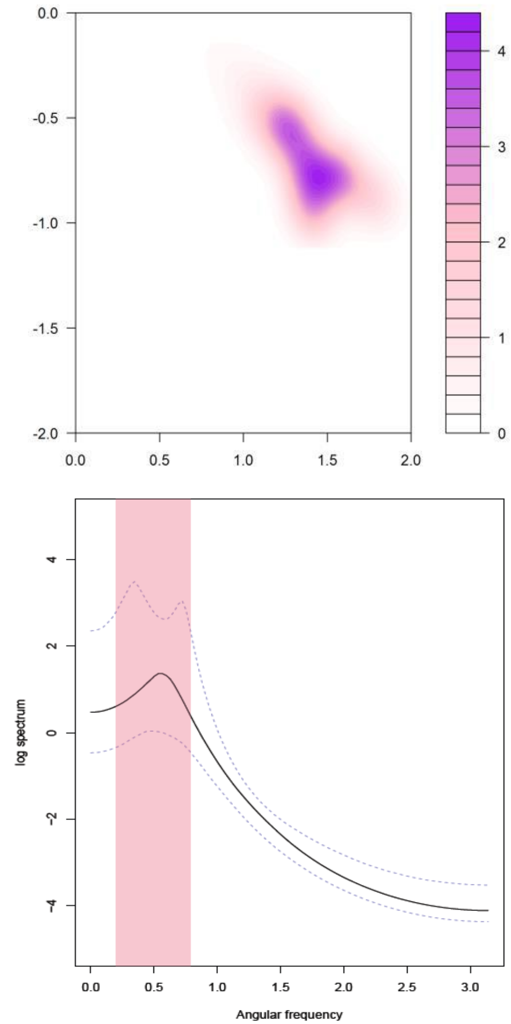
$S_y(w)$ – spectral density of the process

a,b – limits for business cycle frequencies

System priors



At least 60 %



Be(15,5)

Computational Aspects

- Integration of system priors into existing Bayesian toolkit is straightforward.
- General-purpose samplers can be used:
  - Standard Metropolis-Hastings algorithm (low-dimensional problems)
  - Sequential Monte Carlo samplers (Herbst and Schorfheide, 2014)
  - Dynamic striated Metropolis-Hastings algorithm (Waggoner, Wu and Zha, 2016)
  - Homotopy optimization methods (for finding a mode of posterior distribution)

Software implementation and current use of system priors

- SW
  - IRIS [Matlab, Octave]
  - Dynare [Matlab, Octave]
  - Yada [Matlab]
  - R Project – Dynamic Striated Metropolis-Hastings algorithm, SMC
- Applications
  - IMF
  - CNB, Financial Stability Dept. (a forecasting model for residential property prices – “Trend-Cycle VAR”)
  - ECB: NAWM – The New-Area Wide Model of the Euro Area

Conclusions

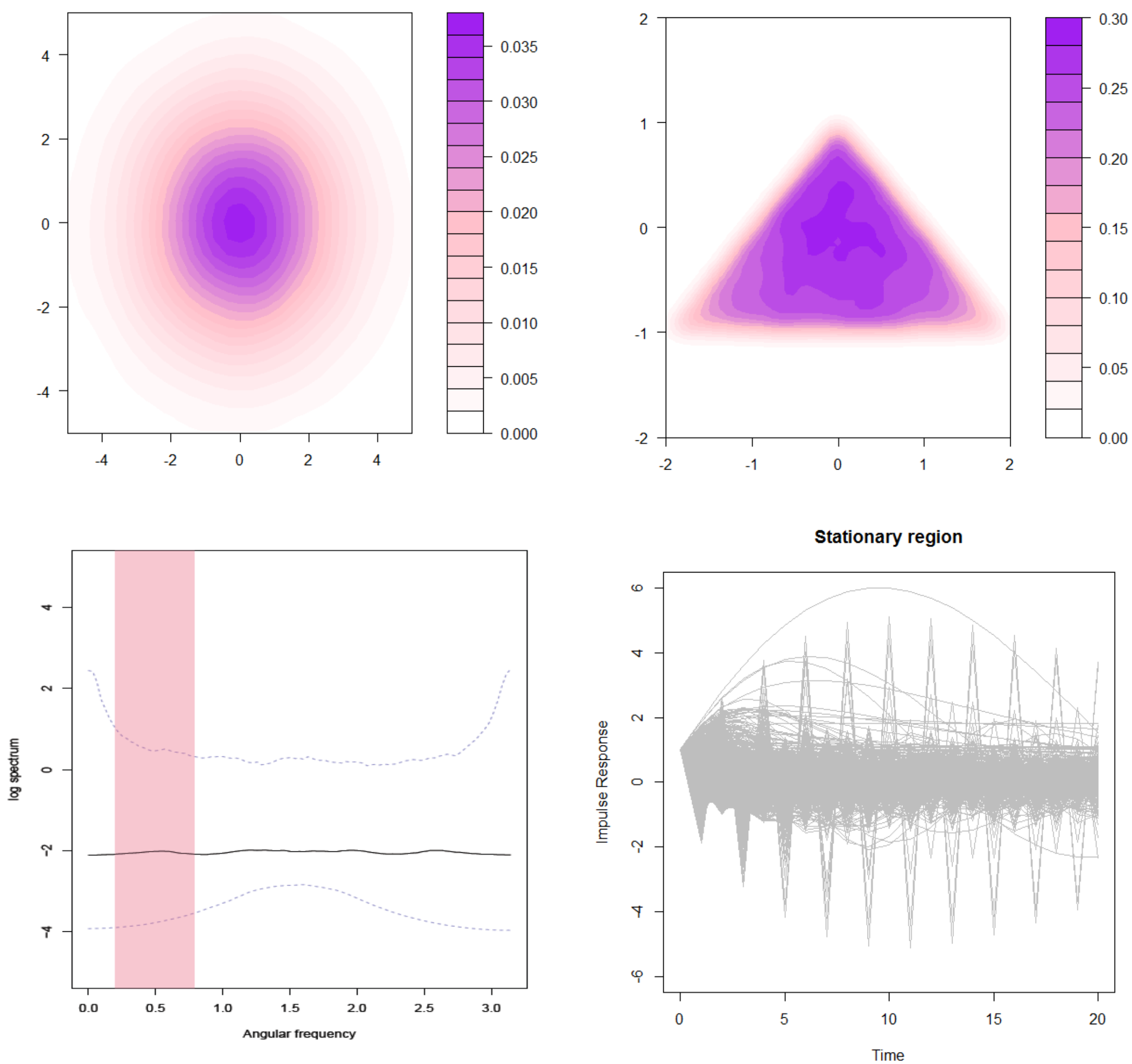
- System priors represent a flexible way of incorporating economically meaningful information.
- They are very general and can be easily implemented within existing Bayesian toolkit.
- The paper places emphasis on the elements and mechanics of system priors' application.
- Implementation of system priors was illustrated using second-order autoregressive process and constraints on stationarity and frequency-domain properties.

Thank you for your attention

- Q & A section
  - Questions and comments are more than welcome!

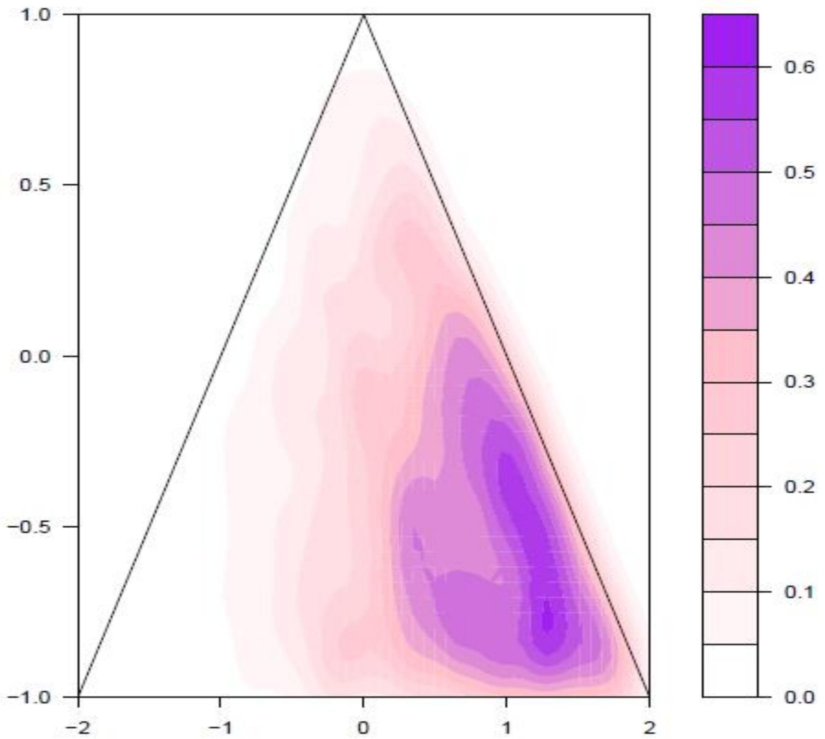
Contact: mandrle@imf.org
miroslav.plasil@cnb.cz

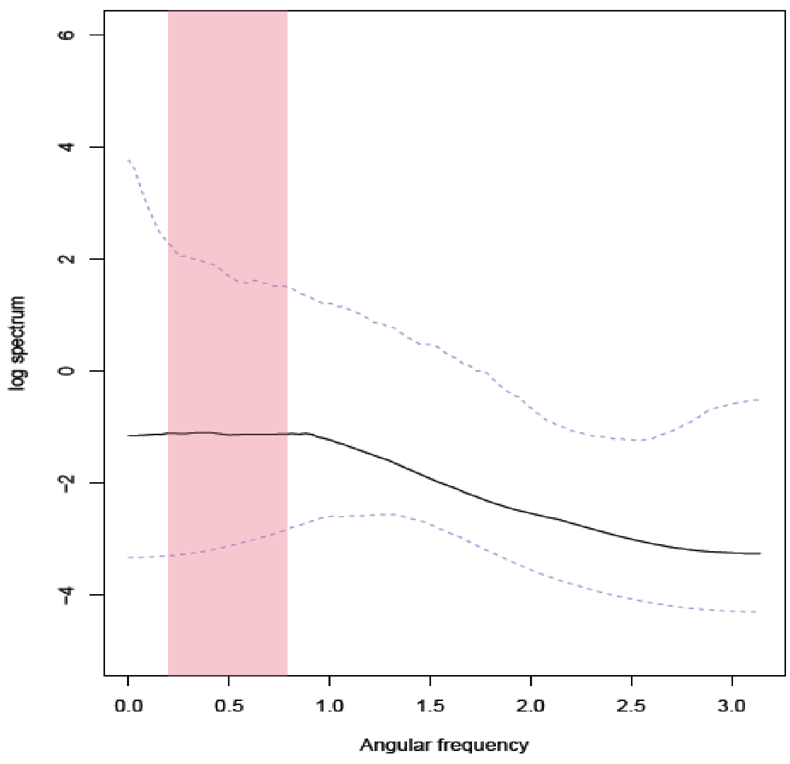
Back-up slides: Basic Gaussian option



Back-up slides: Refined Gaussian option I

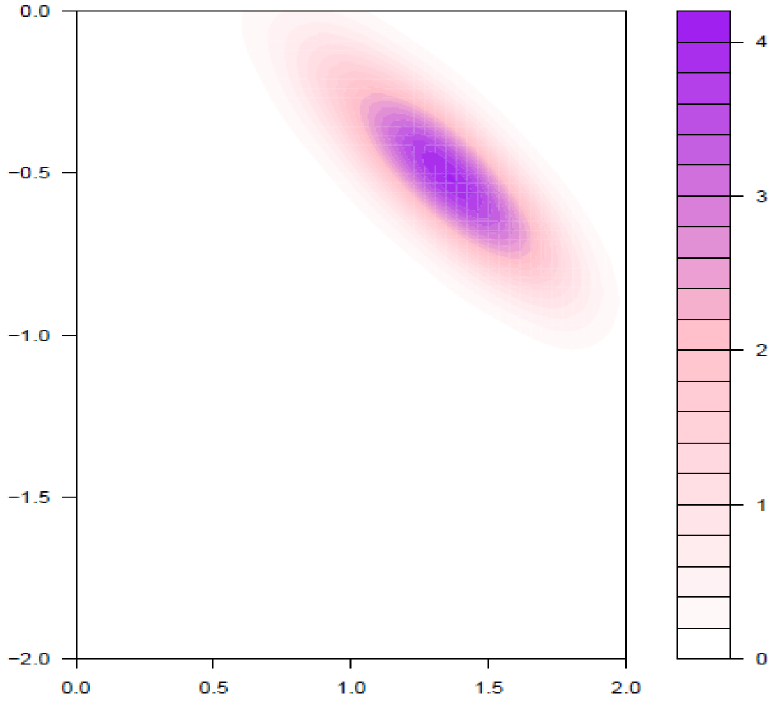
Grant and Chan (2017): $N\left(\begin{pmatrix}1.3\\-0.7\end{pmatrix}, I(2)\right)$

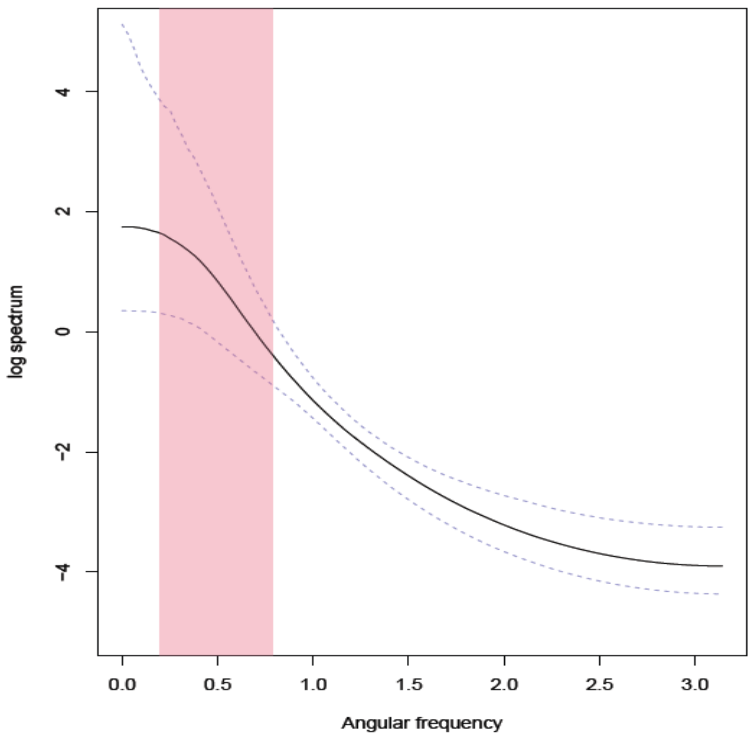




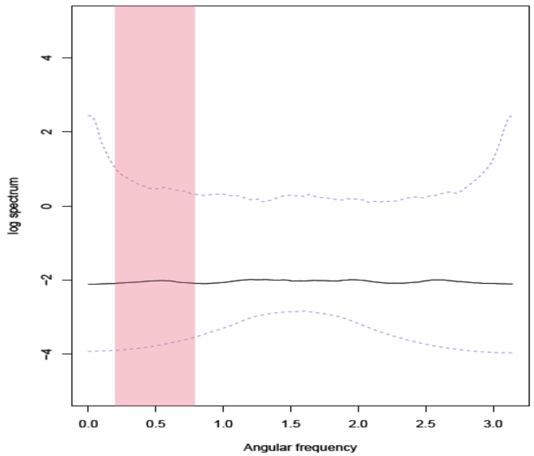
Back-up slides: refined Gaussian option II

Lenza and Jarociński (2016): $N\left(\begin{pmatrix}1.352\\-0.508\end{pmatrix},\begin{bmatrix}0.0806 & -0.0578\\-0.0578 & 0.0464\end{bmatrix}\right)$

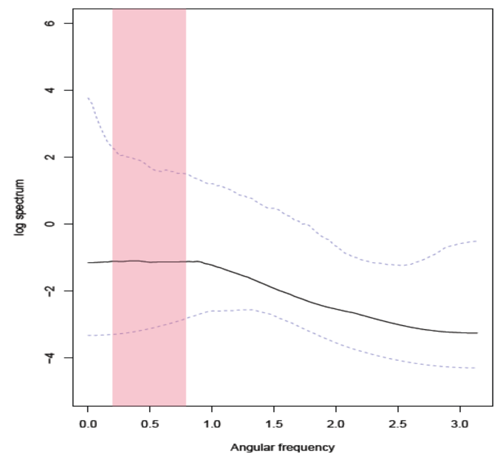




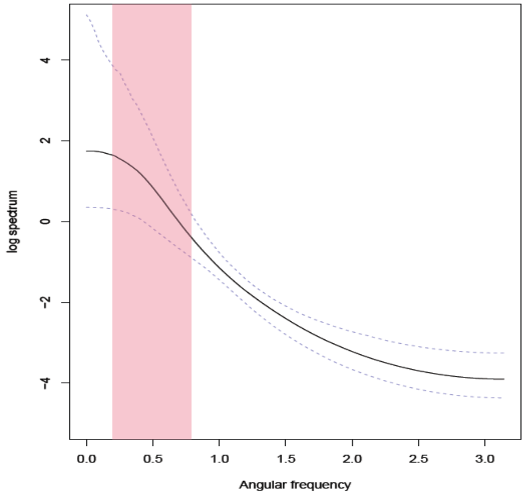
Back-up slides 1



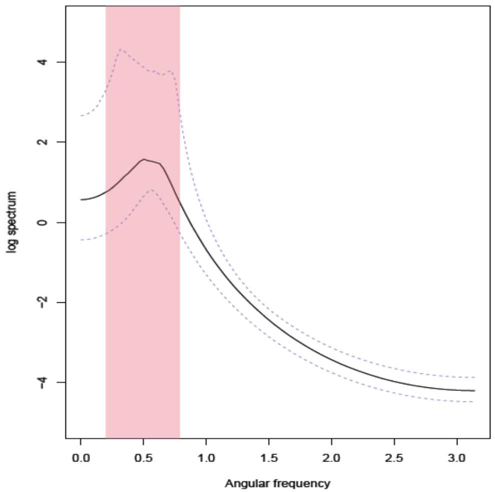
Stationary AR(2)



Chan et al. (2015)

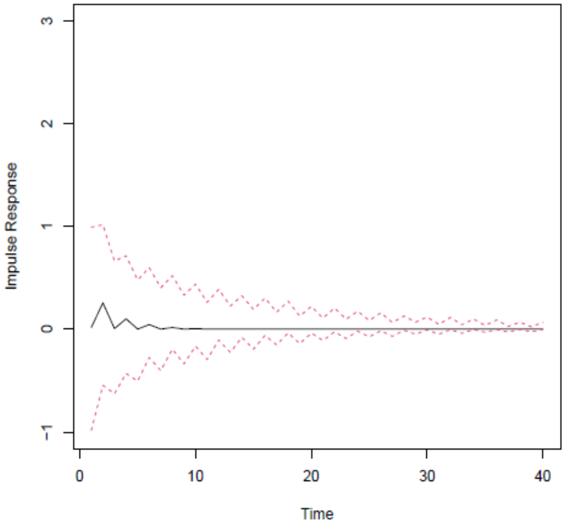


Lenza and Jarociński (2016)

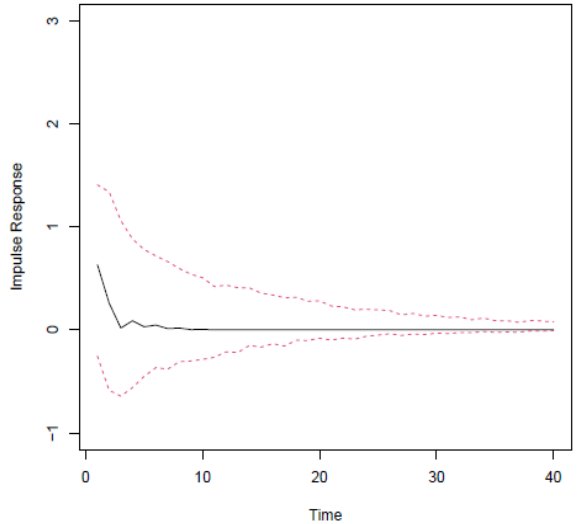


System priors, at least 60%

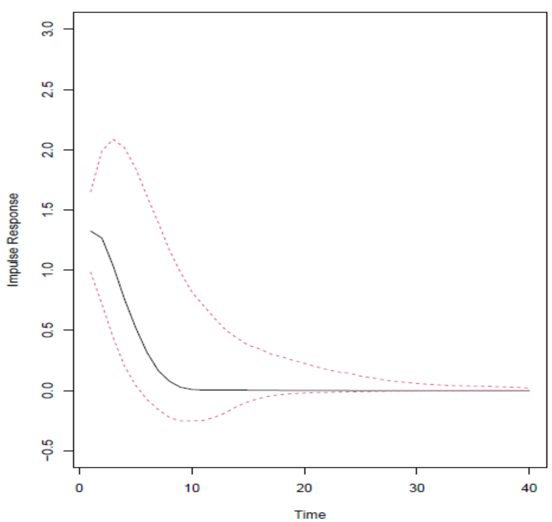
Back-up slides 2: Impulse response functions



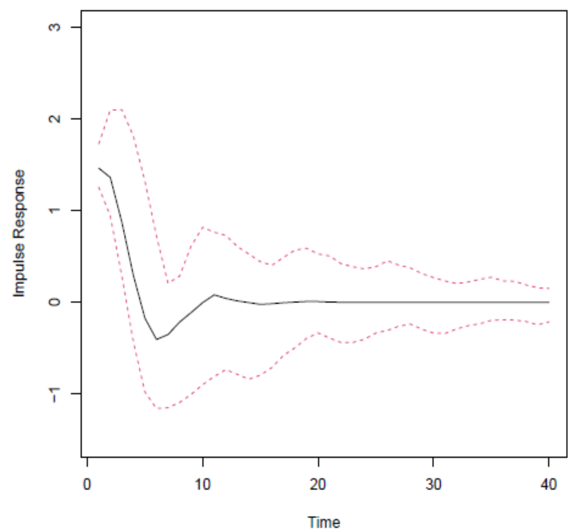
Stationary AR(2)



Chan et al. (2015)



Lenza and Jarociński (2016)



System priors, at least 60%

Back-up slides 2: System priors – alternatives