WORKING PAPER SERIES 1

2017

Michal Andrle, Miroslav Plašil

System Priors for Econometric Time Series

CZECH NATIONAL BANK

WORKING PAPER SERIES

System Priors for Econometric Time Series

Michal Andrle
Miroslav Plašil

1/2017

CNB WORKING PAPER SERIES

Reviewed by: Wouter den Haan (London School of Economics)
Dale Poirier (University of California-Irvine)
Michal Franta (Czech National Bank)

Project Coordinator: Jan Brůha

System Priors for Econometric Time Series

Michal Andrle and Miroslav Plašil *

Abstract

This paper introduces "system priors" into Bayesian analysis of econometric time series and provides a simple and illustrative application. Unlike priors on individual parameters, system priors offer a simple and efficient way of formulating well-defined and economically meaningful priors about model properties that determine the overall behavior of the model. The generality of system priors is illustrated using an AR(2) process with a prior that its dynamics comes mostly from business-cycle frequencies.

Abstrakt

V tomto článku je názorně představeno využití „systémových" apriorních informací v bayesovské analýze ekonometrických časových řad. Na rozdíl od formulace apriorních rozdělení pro jednotlivé parametry přestavuje použití systémových apriorních informací jednoduchý a účinný způsob implementace jasně definovaných a ekonomicky smysluplných představ o vlastnostech modelu a jeho celkovém chování. Obecnost systémových apriorních restrikcí je ilustrována na příkladu autoregresního procesu AR(2) s využitím apriorního přesvědčení, že většina dynamiky procesu je tvořena cyklickým kolísáním o délce hospodářského cyklu.

JEL Codes: C11, C18, C22, C51.

Keywords: Bayesian analysis, system priors, time series.

*Michal Andrle, International Monetary Fund, Research Department (mandrle@imf.org);
Miroslav Plašil, Czech National Bank, Financial Stability Division (miroslav.plasil@cnb.cz).
The authors would like to thank Jan Brůha, Joris de Wind, Wouter den Haan, Michal Franta, Ben Hunt, and Dale Poirier for useful comments and suggestions.

Nontechnical Summary

This paper introduces "system priors" as a tool to incorporate prior knowledge into an economic model. Unlike priors about individual parameters, system priors offer a simple and efficient way of formulating well-defined and economically meaningful priors about high-level model properties. Central banks have historically been users of empirical models in forecasting and policy analysis. It is important that all such models represent the best use of theory and existing institutional knowledge. A top-down approach to incorporating prior knowledge – as embodied by system priors – naturally meets the needs of central banks, since policy makers do not need to have a view on the individual parameters or model details but have prior views about the aggregate model properties, for which the data are often uninformative, especially in an ever-changing economic environment with a paucity of reliable data. With system priors, it is possible to reflect the practitioners' priors in a transparent way.

The paper provides the background theory of system priors, placing an emphasis on the elements and mechanics of the application of system priors. Similar to traditional Bayesian inference, the initial priors on the parameters are updated using the likelihood function. However, they are additionally updated by prior views about the high-level behavior of the model. Essentially, system priors penalize parameter values not conforming to the prior beliefs about the system properties of the model.

The application of system priors is illustrated using the simple but relevant example of an AR(2) process, which, despite its simplicity, can display a wide array of diverse dynamics. In many cases, it might be difficult to discipline its behavior within a traditional Bayesian setup, because it does not provide enough flexibility to fully reproduce researchers' prior knowledge. In our illustration, we incorporate a prior view that the second-order AR process is stationary and a large portion of its variance comes from business-cycle frequencies. Such a prior might be an advantage when an AR(2) process is used for modeling cyclical components of economic variables and researchers need to confine the parameter space to regions they find economically plausible.

The implementation of system priors is based on the business-to-total variance ratio – a univariate spectral characteristic of the model that measures how much variability of the (stationary) process is generated by business-cycle frequencies. The ratio is a non-linear function of both autoregressive parameters, so the system priors result in them having a non-trivial joint prior distribution. In the paper, we consider two technical options for implementing prior beliefs about the cyclical component of the process. First, one may consider a condition that at least 60% of the variance of the process originates from business-cycle frequencies. Second, prior beliefs about the ratio can be expressed using a functional form. Given the range of admissible values for the ratio, a Beta-distributed prior is a feasible option. In our example we use $Be(15,5)$, which places a large portion of its probability mass on processes dominated by business-cycle frequencies. The results show that the system prior considered is fairly informative and leads to a non-normal joint distribution of the parameters. Naturally, it also has an impact on the spectral characteristics of the model as well as on its impulse-response function. The stationarity condition itself does not restrict the process in an economically meaningful way, but the system priors do. While our system prior is not diffuse, it is very transparent, simple to implement, and easy for others to agree or disagree with.

1. Introduction

This paper introduces "system priors," a simple and tractable way of employing economically meaningful a priori judgment about model properties in statistical inference. Using Bayesian methods and formulating priors on individual coefficients is becoming common in many scientific fields, including economics. Analysts provide prior distributions about model parameters based on either theory or other research. However, there are often situations where the a priori beliefs of analysts relate to properties or features of the full model, features that can be highly non-linear functions of individual coefficients – often with no closed-form solution. There are also situations where the individual model parameters are hard to interpret and to elicit priors about, whereas the model system properties are easy to interpret. These are all situations where system priors offer a solution.

Examples of system priors include features of the impulse-response or frequency-response function of the model, such as the response of the economy to a permanent disinflation shock, the maximum persistence of the response of inflation to a demand shock, the sign or magnitude of shock contributions in selected historical periods, or the signs of the responses to shocks of interest – for instance, a non-negative response of inflation after a positive demand shock. This paper explains how to implement such priors in practice.

Even when economically meaningful priors about individual coefficients are easy to formulate, system priors can still be very helpful. This is because individually sensible marginal independent priors on coefficients may induce unintended consequences for model properties when the joint prior distribution does not reflect the relationships between the parameters. Abstracting from the existing mutual dependence of parameters may cause the implied prior on selected properties of the model to become implausible. This often goes unnoticed unless a prior-predictive analysis is carried out. System priors allow analysts to employ the prior view about the relevant system properties directly and may complement or replace the marginal independent priors on the coefficients.

System priors were proposed by Andrle and Benes (2013) as a flexible tool for incorporating economically meaningful prior knowledge into Dynamic Stochastic General Equilibrium (DSGE) models. However, the range of potential applications is far more general. This paper introduces the use of system priors for time-series auto-regressive models and illustrates their use with a simple but relevant example. We also provide a more nuanced motivation for system priors and their implementation, as the exposition in Andrle and Benes (2013) may be less accessible to an audience unfamiliar with the DSGE literature.

To illustrate the main principles, we assume a stationary second-order auto-regressive – AR(2) – process and incorporate a belief that most of its variance comes from business-cycle frequencies. Such a prior is useful when an AR(2) process is used for modeling cyclical components of economic variables and researchers need to confine the parameter space to regions they find economically plausible. This setting is common in many structural time-series models, where an AR(2) process is used to model variables in "gap" form (see, for example, Watson, 1986, Clark, 1987, and Kuttner, 1994 for canonical examples). We illustrate how a system prior about the frequency distribution of the variance (spectrum) creates the joint prior distribution of both coefficients and facilitates the estimation.

Importantly, system priors are easy to implement and the integration of system priors into existing standard Bayesian computations is straightforward. In some sense, the procedure is related to that for "dummy observation" priors (Theil and Goldberger, 1961) and uses the Bayes formula to solve the inversion problem (going from properties to coefficients). From the non-Bayesian point of view, system priors can simply be interpreted as another penalty in the criterion function, along with the likelihood and marginal prior distribution penalties. The formal discussion below will make the computational implementation clear.

2. System Priors

The estimation of models with system priors closely follows the general principles of Bayesian inference. The difference rests in the form of the prior distribution formulation. To demonstrate this, let us start with a traditional Bayesian setup: we assume that the joint prior beliefs about a $(k \times 1)$ vector of individual parameters, θ, of a model M are expressed using independent marginal probability distributions, i.e., as: $p_m(\theta)= p_m(\theta_1)\times\ldots\times p_m(\theta_k)$. Other setups of priors are possible with no loss of generality. We further assume that given the observed data, Y, it is possible to evaluate the likelihood function of the model, $L(Y\,|\,\theta;M)$ for different parameter values. Applying Bayes' law, it is well known that the posterior distribution of the parameters is proportional to the product of the likelihood and the prior distribution:

$$p(\theta\,|\,Y;M) \propto L(Y\,|\,\theta;M) \times p_m(\theta). \tag{1}$$

Now let us incorporate a priori views about the model's system properties. To proceed, let us define a property of interest, r, that a prior view will be formulated about. The property, r, is a function of the individual parameters θ given the model M: $r = h(\theta;M)$. We assume that the feature can be evaluated for different parameter values. Such a function may describe the impulse-response function characteristics or frequency-domain properties of the model, for instance. As in the case of the individual parameters, the prior beliefs about the values of feature r can be summarized by a feasible functional form, specifically by a probability distribution. We will call it the "system prior" and denote it as $p_s(r;h,M) \equiv p_s(h(\theta);M)$. Putting together the effects of the marginal prior, the system prior, and the likelihood function, the posterior distribution of the parameters emerges as

$$p(\theta\,|\,Y;M) \propto L(Y\,|\,\theta;M) \;\times [\,p_s(h(\theta);M) \;\times\; p_m(\theta)\,]. \tag{2}$$

The form of the posterior kernel in (2) is intuitive, read from the right to left. For a given value of parameter θ, the posterior distribution is based on a two-step updating process. In the *first step* the marginal prior, p_m, is updated with the system priors, p_s, resulting in a composite prior distribution. As the system priors operate on functions of parameters, the composite prior $p_c = [\,p_s(h(\theta);M) \times p_m(\theta)]$ implies restrictions on individual coefficients but generally not in a unique or invertible way. In the *second step*, the composite prior beliefs are updated with the information contained in the data using a likelihood function of the model. Estimation with system priors thus represents a two-layer approach where beliefs about the parameters are complemented with beliefs about the model properties. With diffuse marginal priors, the system prior will dominate the composite prior used for estimation.

Although the composite prior emerges as the product of the marginal and system priors, these should not be treated as independent entities. System priors simply imply additional stochastic restrictions on the parameter values. To help with the intuition, it is useful to think of system priors as an artificial likelihood function[1] summarizing the information contained in the artificial data on r, which are put into an auxiliary probabilistic model with a structure corresponding to function $r = h(\theta; M)$. In other words, system priors can be interpreted as measuring how likely the values of individual parameters are given the "as-if observed" outcomes of r. As such, they penalize parameter values not conforming to the prior beliefs about the system properties of the model. Combining the prior distribution of the individual parameters with both the artificial and the conventional likelihood function results in the posterior distribution of the parameters expressed in (2).[2]

Essentially, estimation with system priors just involves applying the Bayes inversion formula twice: first with the artificial likelihood function to obtain the composite prior and second with the conventional likelihood function of the underlying model to obtain the posterior distribution of the model parameters.

Although we rely on a Bayesian interpretation of system priors in this paper, the Bayesian paradigm is not necessary for system priors to be used. From a non-Bayesian perspective, the criterion function (2) is simply a penalized likelihood problem with two types of penalties. The Bayesian and frequentist approaches are often closely connected. The equivalence between the literature on shrinkage in statistics (ridge regression, the lasso estimator) and the specific form of the priors in Bayesian analysis is a good example of such dual interpretation.[3] If feasible, the criterion function (2) can be optimized numerically with respect to θ to find the posterior mode. The inverse Hessian matrix evaluated at the posterior mode can then serve directly for (non-Bayesian) inference or as an important ingredient of Markov Chain Monte Carlo (MCMC) procedures.

To analyze the implications of the composite joint prior distribution in greater detail, computations analogous to posterior sampling in (2) are needed, with evaluation of the conventional likelihood function switched off. Such analysis and an associated prior predictive analysis of the model's properties are highly recommended to check if the formulation of the priors leads to desired or plausible properties of the model (see Geweke, 2010).

In practical applications, the advantage of system priors is that existing Bayesian computations and computer code can stay almost unchanged when system priors are used. The only difference is that for a particular j-th draw of the parameter vector, θ_j, three instead of two components need

[1] This brings it close to the idea of "dummy observation" priors. As pointed out in Sims (2005), "The prior takes the form of the likelihood function for the dummy observations." However, the system priors are not fully equivalent to traditional dummy-observation priors. For example, system priors do not assume conjugacy.
[2] Regardless of whether or not Jacobian terms are involved, the resulting prior distribution of the aggregate model properties is key to understanding all the consequences of the prior specification used, especially in non-trivial models.
[3] For instance, the ridge regression can be recast as a Bayesian problem with Gaussian priors.

to be evaluated – with the system prior component adding to the overhead.[4] Given the computational progress made in the last decade and expected in the years to come, there is no need for the system priors to have closed-form solutions or conjugate forms in most setups.

3. Relationship to the Literature

The idea to use a prior on the properties of an economic model is not new. The use of priors reflecting system properties of the model is arguably the most frequent in the area of VAR models, perhaps due to the lower interpretability of VAR reduced-form coefficients. To name just a few prominent examples, the concept of system-like priors is echoed in the dummy-observation priors used for "shrinkage" (Doan, et al., 1986), the priors on the VAR steady-state by Villani (2009), the long-run behavior priors by Giannone et al. (2016), and in the DSGE-based conjugate priors of Del Negro and Schorfheide (2004), for instance. Useful as these approaches are, they do not generalize to a broad class of models or prior beliefs and are only constructed for a particular problem at hand – often requiring strong assumptions to derive conjugate solutions. A fully general and flexible solution is thus still needed. In the domain of DSGE models, for example, we are not aware of any related work on the use of priors about the overall model behavior, except for Andrle and Benes (2013).

An approach related to system priors but based on an operationally distinct concept is the "priors about observables" by Jarociński and Marcet (2016).[5] The authors start with a formulation of priors about the observed variables (i.e., about the marginal data density) and solve an inverse problem to derive the implied priors on individual parameters. To solve the inverse problem given by the Fredholm equation of the first kind, the authors propose an approximate conjugate algorithm based on a complex iterative fixed-point formulation.

In contrast to their approach, the idea behind system priors is to re-cast the solution of the inverse problem into a Bayesian setting as another step of the Bayesian update. Note that until the 1950s the term "inverse probability" was used for what today is called the posterior distribution. Unlike the procedure by Jarociński and Marcet, system priors start directly with some – usually marginal independent – priors on the individual coefficients and use the system prior to update this information. Using coefficient priors and system priors jointly gives analysts the opportunity to formalize a wide variety of prior views. In VAR models, for instance, system priors can play a decisive role, as the interpretation of individual coefficients is often difficult, but in the case of more theory-based, structural, models (DSGE models and semi-structural models) researchers often do have meaningful prior information on interpretable coefficients alongside the system priors. Solving the inversion problem through the Bayesian update is also very convenient computationally, since standard Bayesian computations need little modification with the use of the Bayes formula.

We consider the system priors approach to be general, transparent, and simpler to implement. In the most common setup, when the dimensionality of the coefficient vector is larger than the

[4] All three components, however, can be evaluated independently and thus in parallel for a particular draw of the parameter vector, or resources can be re-used in multiple components, as both the likelihood function and the system priors make use of a model solution for a new vector of parameters.

[5] We would like to thank Wouter den Haan for the suggestion to contrast the two approaches to the "inverse problem."

dimensionality of the system prior, the iterative fixed-point algorithm by Jarociński and Marcet that solves for the prior distribution of the coefficients inevitably faces the challenge of an ill-posed inversion problem where multiple solutions arise and some way of selecting one solution is needed.[6] System priors make this step very explicit, using the Bayesian update step all the way through the problem. Further, system priors are not limited to priors about observables and generalize to any property of the model that researchers have a view about.

4. System Priors for an AR(2) Process (Example)

After the theoretical exposition of system priors, let us proceed with an illustration using an AR(2) process. Second-order autoregressive models are sometimes used for modeling the cyclical components of output, the unemployment rate, and other economic variables (see, for example, Planas et al., 2008, Clark and Doh, 2014, Chan et al., 2016, and Berger et al., 2016 for some of the latest applications).[7] Having the ability to elicit economically meaningful priors about the process is thus relevant for empirical work. To keep the illustration simple, we focus only on the process itself and do not consider the estimation of a fully specified semi-structural model. Let us consider a zero-mean AR(2) process:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0,\sigma_\varepsilon^2). \tag{3}$$

What would be reasonable priors for the two auto-regressive coefficients $\phi_1$and ϕ_2? A common point of departure would be to start with normally distributed independent marginal priors for the individual coefficients, that is, $\phi_1 \sim N(0,\sigma_{\phi_1}^2)$ and $\phi_2 \sim N(0,\sigma_{\phi_2}^2)$. However, this hardly sounds right if the prior is supposed to convey relevant, economically meaningful information. When the coefficients can vary independently and the joint distribution is spread out, it implies a wide array of model dynamics, including wild oscillations or unstable non-stationary impulse-response functions. Researchers have been aware of this issue for a long time and have been striving to come up with better ways of formulating priors, even in the particular, and simple, example of the second-order autoregressive process (see, for example, Planas et al., 2008).[8]

In the case of the AR(2) process, a polar-form specification of the cycle was proposed as one of the solutions, as it helps incorporate prior beliefs about cyclical behavior more efficiently. In the polar-form specification, the coefficients are analytically re-parameterized such that the priors are imposed on the periodicity and the amplitude of the cycle.[9] However, such re-parameterization

[6] The need to select one of multiple solutions is mentioned in Jarociński and Marcet (2016) but not discussed as a general principle.

[7] There are other common specifications of cyclical components in the literature, for instance the trigonometric form in Harvey and Trimbur (2003), which in its univariate form corresponds to a restricted ARMA(2,1) process.

[8] Planas et al. (2008) have recently commented on the problem with a business cycle modeled as an AR(2) process [p. 19]: "Indeed, assuming a normal prior distribution on parameters ϕ_1, ϕ_2, we found it difficult to reproduce our prior knowledge by tuning the mean and the covariance matrix of the autoregressive parameters… In some cases, the implied distribution for the periodicity and amplitude can be counterintuitive…Putting the prior on the AR coefficients [*in the traditional way*] is probably inadequate for cyclical analysis."

[9] The polar-form specification is as follows: $(1-2A\cos(2\pi/\tau)L + A^2L^2)y(t)=\varepsilon(t)$, where L is the backshift operator, A is the amplitude, and τ is the periodicity. The amplitude is given by $A=\sqrt{-\phi_2}$ and the periodicity by $\tau = 2\pi/\mathrm{acos}\{\phi_1/2A\}$.

may still be too vague for other types of a priori views about the business cycle. Further, analytical re-parameterization is not generally feasible except for very simple models. Luckily, there is absolutely no need for it. System priors usually will not have closed form solutions. Not having a closed-form solution may add some computing overhead but does not affect the general principles.

In our application example, we incorporate a prior view that the second-order AR process is stationary and a large part of its variance comes from business-cycle frequencies (i.e., frequencies of 8-32 periods in a quarterly model). Such a prior is economically appealing if the AR(2) process is supposed to model the cyclical behavior of the economy.

The process in (3) is stationary only if $\phi_1 + \phi_2 < 1$, $\phi_2 - \phi_1 < 1$ and $|\phi_2| < 1$. These assumptions restrict the parameter space but they do not restrict the oscillatory properties of the model in a sensible way, as illustrated below. More disciplined behavior of the model can be achieved through a prior assumption about the spectral characteristics of the process. The spectral density of the model can be interpreted as a distribution of the variance across frequencies and is thus a natural starting point for formulating a system prior in this case.

The spectral density of y_t, denoted $S_y(w)$, can be computed as follows:

$$S_y(w) = \frac{\sigma_\varepsilon^2}{2\pi\left[1+\phi_1^2+\phi_2^2+2\phi_1(\phi_2-1)\cos(w)-2\phi_2\cos(2w)\right]}, \qquad (4)$$

where $w \in [0,\pi)$ is the angular frequency. A brief inspection shows that the spectral density is a non-linear function of both auto-regressive parameters. The variance of the error term, σ_ε^2, determines the level of the spectrum but not its shape.[10] Therefore, any prior exploiting a spectral restriction would result in a non-trivial joint prior distribution for the individual regression parameters. To introduce the system prior outlined above, we define the total variance of the process y_t as the integral of the spectrum (4) over the full frequency range and the business-cycle variance as the integral limited to the range of business-cycle frequencies (a,b). Specifying the business-to-total variance ratio as:

$$r = \int_a^b S_y(w)\mathrm{d}w \Big/ \int S_y(w)\mathrm{d}w, \qquad (5)$$

results in a statistic that is univariate and has clear units and interpretation. The ratio (5) can only take values within the interval [0,1]. A change in the shock variance, σ_ε^2, shifts the spectrum up or down but never affects the ratio. As such, our system prior is completely uninformative about the coefficient σ_ε^2 and so does not update its marginal coefficient prior.[11] In general, system priors are not equally informative about all coefficients.

[10] Although for the AR(2) process the spectrum can be expressed in a closed form, nothing would change if the closed form was not available.

[11] Since the system priors do not modify the initial prior on the error-term variance, we leave it unspecified here. In practical applications, one can naturally use the traditional inverse gamma specification or any other distribution that would meet researchers' needs. The choice of the prior distribution for the error-term variance has no impact on the results.

Now let us present two examples of implementing a system prior that reflects prior beliefs about the cyclical component of output. First, one may consider a condition that at least 60% of the variance of y_t originates from business-cycle frequencies. Second, prior beliefs about the ratio can be expressed using a functional form. Given the range of admissible values for *r*, Beta distribution is a feasible option. In our example we use $r \sim Be(15,5)$, which places a large portion of the probability mass of the variance of y_t as coming from business-cycle frequencies. Other hyper-parameter settings are possible and the values used here are only illustrative.

All the results below are based on simulation techniques with the conventional likelihood function omitted to learn only about the composite prior. Where it is required that a minimum of 60% of the variability comes from business-cycle frequencies, we employ rejection sampling with normally distributed marginal priors used as the proposal distribution. In the latter case with the Beta distribution, our results are based on sequential Monte Carlo sampling (SMC; see, for example Herbst and Schorfheide, 2014). This is an alternative to the traditional Metropolis-Hasting random walk algorithm, which in our simple case might also perform well. For simple models the two algorithms should provide almost identical results, but sequential Monte Carlo sampling may be preferred if complex models (containing dozens of parameters) are estimated.[12]

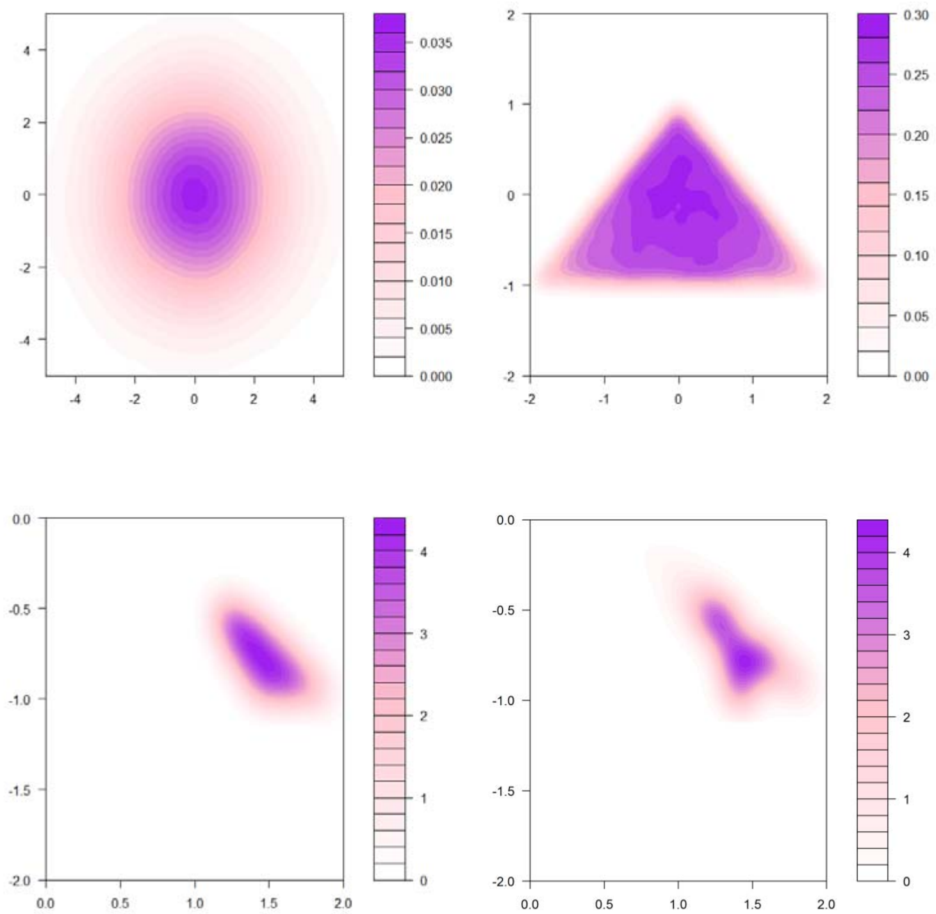
Fig. 1 shows the combinations of parameters that correspond to normally distributed marginal priors, $\phi_1, \phi_2 \sim N(0,2)$ in the upper-left panel and the combinations that conform to the stationarity restriction in the upper-right panel. The bottom panels illustrate the combinations consistent with the requirement for sufficient variance of y_t coming from business-cycle frequencies. The system prior used is fairly informative and leads to a non-normal joint distribution of parameters. The two computational ways of implementing system priors reflect similar prior beliefs and lead to similar results, as they should. System priors impose few restrictions on the actual technical design of the prior – all that matters is the meaningfulness of the prior to the analysts and their audience.[13]

Knowing just the combinations and the full joint prior distribution of the individual parameters that satisfy the constraints is far from enough to evaluate the role of priors. The key knowledge is the understanding of how these priors translate into the behavior of the model in as many aspects as relevant. The analyst should investigate whether the chosen priors have any unintended consequences. For this purpose, the prior-predictive distribution of the model's properties must be analyzed. In our case, the prior-implied distribution of the impulse-response function and the spectral characteristics are natural candidates for closer inspection.

[12] The R code for the examples presented is available upon request.

[13] The computer-code implementation of system priors differs from standard Bayesian analysis in that the prior restrictions are not off-the-shelf functions and users are expected to specify their own. Once a clear interface is established and documented, users only pass their function or function object with clear inputs and outputs to the system.

Figure 1: Parameter Regions for Different Priors

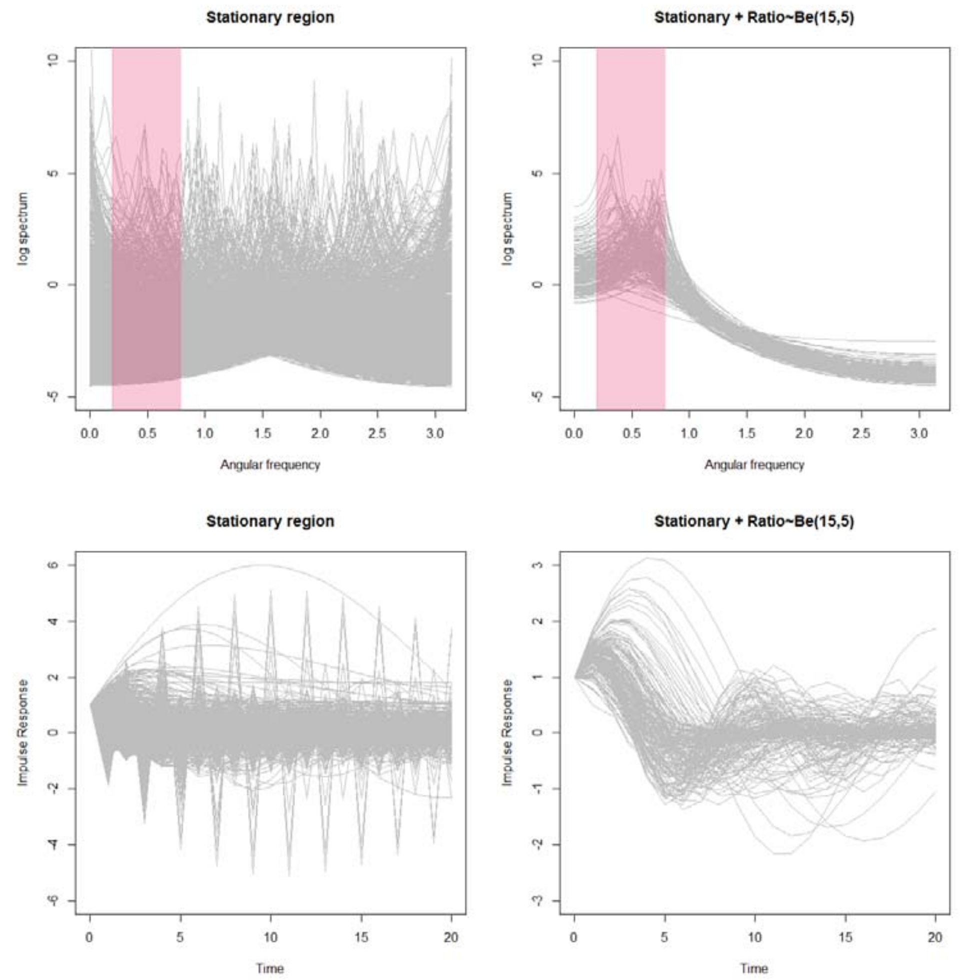


Note: Kernel estimates of the joint prior density. Upper left panel: normally distributed independent marginal priors for ϕ_1 and ϕ_2; upper right: identical priors restricted to the stationarity region; bottom left: stationarity + at least 60% of the variability comes from business-cycle frequencies; bottom right: stationarity + the share of business-cycle frequencies given by *Be*(15,5).

Fig. 2 depicts the spectral densities and impulse-response functions for parameters in regions complying with the requirement of stationarity and sufficient variance coming from business-cycle frequencies. It is apparent that the stationarity condition itself does not restrict the process in an economically meaningful way, while the system priors do. Our system prior is not diffuse, it is fairly informative. However, it is also very transparent, simple to implement, and easy for others to agree or disagree with, should they wish to do so.

We could have specified other meaningful priors, for example directly in terms of the impulse-response function of the model. The scope of system priors is wide. System priors are a flexible tool which easily extends to any other type of econometric and statistical model, including state-space models (Andrle and Benes, 2013) and Bayesian VARs (Andrle and Plašil, 2017). Recall also that non-Bayesian analysis can embrace the penalized loss-function approach to inference as well.

Figure 2: Model Properties for Admissible Regions



Note: The business-cycle frequencies are denoted by the shaded region.

5. Conclusion

Building on Andrle and Benes (2013), we provided a more detailed discussion of system priors and an illustrative example, placing an emphasis on the elements and mechanics of the application of system priors. System priors take on board views about the high-level features of models, not necessarily just individual coefficients. As such, they provide a more refined way of incorporating prior information on complex functions of parameters, such as impulse-response or frequency-response functions, in many types of empirical models.

The specification and implementation of system priors was illustrated using a second-order autoregressive process, which, despite its simplicity, can display non-trivial dynamics. Gaussian independent priors on the autoregressive coefficients do not restrain the model dynamics in a meaningful way when it comes to the cyclical properties of the process. The polar re-parameterization suggested in the literature is a specific modification offering only a modest improvement. However, we illustrated that imposing a restriction that a large part of the model's dynamics comes from business-cycle frequencies allows the parameters to be estimated only in a region with plausible cyclical dynamics of the impulse-response function. Other economically relevant priors could have been chosen due to the generality of system priors.

We believe that system priors are a useful approach to eliciting priors – possibly hierarchical – about model characteristics as long as these are computable functions of the underlying coefficients. System priors allow researchers to work with informative and economically meaningful priors in econometric and structural economic models, be it state-space models, Bayesian vector auto-regressions, or others. Importantly, system priors are easy to incorporate into existing Bayesian toolkits with little computational overhead.

References

ANDRLE, M. AND J. BENES (2013): “System Priors: Formulating Priors about DSGE Models’ System Properties.” Working Paper WP/13/257, International Monetary Fund, Washington, D.C.

ANDRLE, M. AND M. PLAŠIL (2017): “System Priors for VARs.” *Mimeo*.

BERGER, T., G. EVERAERT, AND H. VIERKE (2016): “Testing for Time Variation in an Unobserved Components Model for the U.S. Economy.” *Journal of Economic Dynamics and Control*, Vol. 69, pp. 179–208.

CHAN, J. C. C., G. KOOP, AND S. M. POTTER (2016): “A Bounded Model of Time Variation in Trend Inflation, Nairu and the Phillips Curve.” *Journal of Applied Econometrics*, Vol. 31(3), pp. 551–565.

CLARK, P. (1987): “The Cyclical Comovement of U.S. Economic Activity.” *Quarterly Journal of Economics*, Vol. 102, pp. 797–814.

CLARK, T. E. AND T. DOH (2014): “Evaluating Alternative Models of Trend Inflation.” *International Journal of Forecasting*, Vol. 30(3), pp. 426–448.

DEL NEGRO, M. AND F. SCHORFHEIDE (2004): “Priors from General Equilibrium Models for VARS.” *International Economic Review*, Vol. 45(2), pp. 643–673.

DOAN, T., R. LITTERMAN, AND C. SIMS (1986): “Forecasting and Conditional Projection Using Realistic Prior Distributions.” Federal Reserve Bank of Minneapolis, Staff Report 93, July 1986.

GEWEKE, J. (2010): “Complete and Incomplete Econometric Models.” Princeton University Press.

GIANNONE, D., M. LENZA, AND G. PRIMICERI (2016): “Priors for the Long Run.” Centre for Economic Policy Research, London.

HARVEY, A. C. AND S. J. KOOPMAN (1997): “Multivariate Structural Time Series Models.” In: Heij, C., et al. (eds.), System Dynamics in Economic and Financial Models, Wiley, Chichester.

HARVEY, A. C. AND T. M. TRIMBUR (2003): “General Model-Based Filters for Extracting Cycles and Trends in Economic Time Series.” *Review of Economics and Statistics*, Vol. 85(2), pp. 244–255.

HERBST, E. AND F. SCHORFHEIDE (2014): “Sequential Monte Carlo Sampling for DSGE Models.” *Journal of Applied Econometrics*, Vol. 29(7), pp. 1073–1098.

JAROCIŃSKI, M. AND A. MARCET (2016): “Priors about Observables in Vector Autoregressions.” https://www.frbatlanta.org/~/media/Documents/research/seminars/2016/jarocinski-031516.pdf, *mimeo*.

KUTTNER, K. (1994): “Estimating Potential Output as a Latent Variable.” *Journal of Business and Economic Statistics*, Vol. 12(3), pp. 361–368.

PLANAS, C., A. ROSSI, AND G. FIORENTINI (2008): “Bayesian Analysis of the Output Gap.” *Journal of Business & Economic Statistics*, Vol. 26(1), pp. 18–32.

SIMS, C. (2005): “Conjugate Dummy Observations for VARs.” Princeton University, *mimeo*.

THEIL, H. AND A. GOLDBERGER (1961): "On Pure and Mixed Statistical Estimation in Economics." *International Economic Review*, Vol. 2(1), pp. 65–78.

VILLANI, M. (2009): "Steady State Priors for Vector Autoregressions." *Journal of Applied Econometrics*, Vol. 24, pp. 630–650.

WATSON, M. (1986): "Univariate Detrending Methods with Stochastic Trends." *Journal of Monetary Economics*, Vol. 18, pp. 49–75.

Appendix A: Pseudo Code for the Posterior Kernel

The following is a simplified pseudo-code for implementing the computations to evaluate formula (2) in the main body of the text, restated here for convenience:

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

The function evaluating all three components of $p(\theta \mid Y; M)$ takes as inputs the vector of coefficients, θ, to evaluate the criterion function for the model (either already solved for θ or to be solved for θ) and the observed data required for evaluation of the log-likelihood and possibly also for evaluation of the system priors.

A crucial input is the user-defined function, `logsprior_user_fun`, which can evaluate the system priors for a given coefficient vector, θ. The function handle, or function object,[14] needs to follow a pre-specified application programming interface (API) to be used with a general toolbox.

The evaluation of the function can be efficient if one solves the model with a new vector of coefficients only once or evaluates all three components in parallel.

The switches allow one to switch between Bayesian estimation with system priors, Bayesian estimation without system priors, maximum likelihood estimation with no explicit priors, and investigation of the compound prior by switching off the likelihood component. Although the "do_xx" switches are not shown explicitly as inputs, they are included in the function (or function object).

```
[crit]  = function(theta, Model, Data, logsprior_user_fun, … )

BEGIN

      /* Evaluate the marginal priors:  p_m(θ)]. */

      IF (do_mprior == TRUE)

            Log_mprior = evalMarginalPriors(theta, hyperParameters);

      ELSE

            Log_mprior = 0;

      END

      /* Evaluate the SYSTEM priors:  p_s(h(θ);M)*/

      IF (do_sprior == TRUE)

            Log_sprior = call(@logsprior_user_fun(theta, Model, Data);

      ELSE

            Log_sprior = 0;
```

[14] For an explanation of function objects in multiple programming languages, see https://en.wikipedia.org/wiki/Function_object. A function object is an object that can be called like a function, yet can do more, such as "remember" a lot of data or its previous state.

```
    END

    /* Evaluate the likelihood or other criterion function: L(Y|θ;M) */

    IF (do_loglik == TRUE)

          Log_lik  = evalLoglikelihood(theta, Data, Model);

    ELSE

          Log_lik  = 0;

    END

    /* Assemble and return the posterior value */

    crit = Log_lik + Log_sprior + Log_mprior

END
```

CNB WORKING PAPER SERIES (SINCE 2016)

1/2017	Michal Andrle Miroslav Plašil	*System priors for econometric time series*
12/2016	Kamil Galuščák Ivan Sutóris	*Margins of trade: Czech firms before, during and after the Crisis*
11/2016	Oxana Babecká Kucharčuková Jan Brůha	*Nowcasting the Czech trade balance*
10/2016	Alexis Derviz	*Credit constraints and creditless recoveries: An unsteady state approach*
9/2016	Jan Babecký Michal Franta Jakub Ryšánek	*Effects of fiscal policy in the DSGE-VAR framework: The case of the Czech Republic*
8/2016	Tomáš Havránek Anna Sokolova	*Do consumers really follow a rule of thumb? Three thousand estimates from 130 studies say "probably not"*
7/2016	Volha Audzei	*Confidence cycles and liquidity hoarding*
6/2016	Simona Malovaná Jan Frait	*Monetary policy and macroprudential policy: Rivals or teammates?*
5/2016	Michal Franta	*Iterated multi-step forecasting with model coefficients changing across iterations*
4/2016	Luboš Komárek Kristyna Ters Jörg Urban	*Intraday dynamics of Euro Area sovereign credit risk contagion*
3/2016	Michal Andrle Jan Brůha Serhat Solmaz	*On the sources of business cycles: Implications for DSGE models*
2/2016	Aleš Bulíř Jan Vlček	*Monetary transmission: Are emerging market and low-income countries different?*
1/2016	Tomáš Havránek Roman Horváth Ayaz Zeynalov	*Natural resources and economic growth: A meta-analysis*

CNB RESEARCH AND POLICY NOTES (SINCE 2016)

CNB ECONOMIC RESEARCH BULLETIN (SINCE 2016)

May 2017	*Trade and external relations*
November 2016	*Financial cycles, macroprudential and monetary policies*
April 2016	*Topics in labour markets*

Czech National Bank
Economic Research Department
Na Příkopě 28, 115 03 Praha 1
Czech Republic
phone: +420 2 244 12 321
fax: +420 2 244 14 278
http://www.cnb.cz
e-mail: research@cnb.cz
ISSN 1803-7070