105
DISCUSSION PAPER

Originally published by Ipea in September 2001 as
number 825 of the series Texto para Discussão.

BAYESIAN ANALYSIS OF ECONOMETRIC TIME SERIES MODELS USING HYBRID INTEGRATION RULES

Ajax R. B. Moreira
Dani Gamerman

ipea

105
DISCUSSION PAPER

Originally published by Ipea in September 2001 as
number 825 of the series Texto para Discussão.

Brasília, January 2015

BAYESIAN ANALYSIS OF ECONOMETRIC TIME SERIES MODELS USING HYBRID INTEGRATION RULES

Ajax R. B. Moreira[1]
Dani Gamerman[2]

1. Da Diretoria de Estudos Macroeconômicos do Ipea.
2. Do Instituto de Matemática da Universidade Federal do Rio de Janeiro.

DISCUSSION PAPER

A publication to disseminate the findings of research directly or indirectly conducted by the Institute for Applied Economic Research (Ipea). Due to their relevance, they provide information to specialists and encourage contributions.

Discussion paper / Institute for Applied Economic

Research.- Brasília : Rio de Janeiro : Ipea, 1990-

ISSN 1415-4765

1. Brazil. 2. Economic Aspects. 3. Social Aspects. I. Institute for Applied Economic Research.

CDD 330.908

SUMMARY

RESUMO

Este artigo utiliza procedimentos de inferência bayesiana para estimar modelos econométricos freqüentemente usados. Em particular, os modelos dinâmicos ou de espaço de estado são considerados detalhadamente. Procedimentos de inferência baseiam-se em esquemas de integração híbridos, em que as variáveis de estado são integradas analiticamente, e os hiperparâmetros são integrados utilizando o método de cadeias de Markov de Monte Carlo. As regiões de credibilidade da previsão e das funções de resposta a impulso são também avaliadas. Os procedimentos são ilustrados com dados reais da economia brasileira.

BAYESIAN ANALYSIS OF ECONOMETRIC TIME SERIES MODELS USING HYBRID INTEGRATION RULES

Dani Gamerman and Ajax R. B. Moreira
Instituto de Matemática, Universidade Federal do Rio de Janeiro, Brazil
Instituto de Pesquisa Econômica Aplicada, Brazil
dani@im.ufrj.br

Address for correspondence
Dani Gamerman
Instituto de Matemática
Universidade Federal do Rio de Janeiro
Caixa Postal 68530
21945-970 Rio de Janeiro, RJ
Tel: + 55 21 2562 7911
Fax: + 55 21 2290 1095
email: dani@dme.ufrj.br

Key words: Bayesian; Dynamic; Hyperparameters; Impulse response; Markov chain Monte Carlo; Metropolis-Hastings algorithm; Vector autoregressive models.

ABSTRACT

This paper is concerned with the study of Bayesian inference procedures to commonly used time series models. In particular, the dynamic or state-space models, the time-varying vector autoregressive model and the structural vector autoregressive model are considered in detail. Inference procedures are based on a hybrid integration scheme where state parameters are analytically integrated and hyperparameters are integrated by Markov chain Monte Carlo methods. Credibility regions for forecasts and impulse responses are then derived. The procedures are illustrated in real data sets.

1. INTRODUCTION

One of the most widely used method to estimate parameters in econometric time series models is maximum likelihood (ML). The method is described in detail in the books by Harvey (1) and Hamilton (2). Typically, model parameters are divided into µ, that is usually referred to as state parameter, and Ã referred to as hyperparameter. More explicitly, µ consist of the parameters that can be eliminated by integration or concentration. The method consists on eliminating µ and performing inference on Ã. This procedure is inappropriate for many reasons. First, it does not take into account the uncertainty associated with Ã when making inference about µ, thus leading to overly optimistic con¯dence bounds for functions of µ. Secondly, inference procedures are based on asymptotic approximations that may work very poorly in applications, specially when transformations of the parameters are required. Also, it disregards the possibility of multiple maxima.

From a Bayesian perspective, most calculations required for inference involve integration. In many time series models, this poses a computational problem that cannot be solved analytically. There has been a surge in the literature in recent years using Markov chain Monte Carlo (MCMC) methods in order to solve the computational problem. After the initial work by Carlin et al. (3), other papers by Carter and Kohn (4), Fruhwirth-Schnatter (5) and Chib and Greenberg (6) followed with a computationally improved methodology for normal state space models. Extensions to non-normal observations are presented by Shephard and Pitt (7) and Gamerman (8). MCMC methods were used in vector autoregressive (VAR) models by Kadiyala and Karlsson (9). All relevant features of the likelihood and of the posterior of Ã, including multimodality and asymmetry, can clearly be identi¯ed now. A common feature of these papers is the use of MCMC-based approximations to all model parameters. In doing it, they disregards the analytical tractability of the models with respect to µ and thus make unnecessary approximations for that component of the model.

This paper combines the advantages of analytical integration of µ used for ML estimation with the advantages of MCMC techniques for Ã in a fully Bayesian analysis to all model parameters. Integrations required for inference about Ã are replaced by averages based on MCMC samples. Note that proper account of the uncertainty about Ã is retained here. The integrations required for inference are hybrid as they combine analytic expressions for µ with (MCMC-based) approximations for Ã. Note also that typically the dimension of µ is orders of magnitude larger than the dimension of Ã. For example, in the ¯rst application of this paper the dimension of µ is 971 whereas the dimension of Ã is 7. Approximations over a 7-dimensional space should be more precise than those over a 978-dimensional space that includes the smaller space. Hybrid rules were previously used in econometric models by Lopes, Moreira and Schmidt (10) and by Sims and Zha (11) with resampling based approximations for Ã. Their main di±culty was the speci¯cation of a suitable importance sampling density. This point is discussed by Schmidt, Gamerman and Moreira (12) and they suggest using adaptive resampling procedures. Their approach works well for models with a limited number of hyperparameters but suitable importance distributions becomes increasingly hard to specify. This paper proposes a di®erent class of procedures that is more encompassing and works with any dimension of the hyperparameter space.

The next Section describes the time series models considered in this paper. They are the univariate dynamic linear model (DLM), described from a Bayesian perspective by West and Harrison (13) and from a frequentist perspective by Harvey (1), the time-varying VAR models, described by Kitagawa and Gersch (14), and the structural VAR models, described from a Bayesian perspective by Sims and Zha (11) and from a frequentist perspective by Hamilton (2). They can be encompassed as special cases of common components multivariate DLM's.

The analytic part of the required integrations is described in Section 2. The MCMC methodology used to perform the sampling-based approximations for Ã is described in Section 3. The methods are very °exible and can accommodate for many choices of prior distribution for Ã. A few guidelines on the choices of prior distributions are provided and serve as an example to applications often found in econometric time series. The method is an inferential tool as easily applicable as ML estimation, with replacement of optimization by sampling routines.

Section 4 presents the applications to real data sets to illustrate the methodology. In particular, speci¯c issues of the models are discussed. Special attention is given to transformations of the model parameters and an appropriate account of their uncertainty. In dynamic models, interest centers in forecasts and estimation of some model components. In structural VAR models, interest centers in forecasts and the impulse response function of the identi¯ed shocks, which describes the interpretable dynamic properties of the model. Finally, Section 5 draws some concluding remarks.

Distributional notation is speci¯ed in an Appendix. All distributions mentioned there are easy to draw samples from, have known moments and allow simple evaluation of probability intervals. It will be shown in the sequel that conditional on ψ, the required distributions are given as in the Appendix, and the analysis is analytically tractable. Approximating MCMC methods will only be used for the intractable part of the posterior distribution, namely the posterior marginal of ψ.

2. MODELS

This section presents the univariate dynamic linear models, the time-varying vector autoregressive models and the structural VAR models. These model are applied to real data sets in Section 4. They can all be written as special cases of the common component dynamic linear models. Normal inverse Wishart priors have shown good empirical performance in similar models (Kadiyala and Karlsson (9)) and are used here.

For univariate DLM's and time-varying VAR models, θ consists of the state parameters or time-varying regression coefficients and the hyperparameter ψ consists of the system evolution variance and sometimes also elements of the system transition matrix. For the structural VAR models, θ is the matrix of regression coefficients of the reduced form of the model and the hyperparameter ψ consists of the elements of the matrix of contemporaneous relation of the endogenous variables. In many cases, observational variances can also be integrated out analytically and are included in θ.

2.1. Common component dynamic linear models

This model is used to describe the joint temporal movement of multivariate time series models that share the same explanatory variables. The description of the different component series differs by a different regression coefficent for each series. In addition, it will be assumed that the regression coefficients are subject to the same time evolution. Namely,

$$y_t' = x_t'(\psi)\theta_t + v_t'; \quad v_t \sim N(0;\Sigma);\ t = 1;\ldots;n \tag{1}$$

$$\theta_t = G_t(\psi)\theta_{t-1} + w_t; \quad w_t \sim N(0;W_t(\psi);\Sigma) \tag{2}$$

where $y_t = (y_{t;1};\ldots;y_{t;q})$ is a q-vector of observations, x_t is a p-vector of explanatory variables, $\theta_t = (\theta_{t;1};\ldots;\theta_{t;q})$ is a $p \times q$ parameter matrix, $\theta_{t;j} = (\theta_{t;j1};\ldots;\theta_{t;jp})'$ is the p-vector of regression coefficients for the jth observation, $j = 1;\ldots;q$ and the notation $N(A;B;C)$ for the matricvariate normal distribution is explained in the Appendix. The model is completed with prior $(\theta_0;\Sigma)|\psi;D_0 \sim NIW(m_0;C_0;\gamma_0;S_0)$ where all parameters may depend on ψ and

$\psi | D_0$ with density f_ψ and D_t denote the information at time t.

Given the value of ψ, relevant conditional distributions are given by the Kalman filter and smoother (Harvey (1); Hamilton (2); West and Harrison (13)). In particular, the one-step-ahead predictive distributions of $y_{t+1} | \psi, D_t$ and the smoothed distributions are $\theta_t | \psi, D_n$ are Student t with defining parameters analytically obtained.

Quantities of interest for inference include components of the parameters θ_t and Σ or their functions such as impulse response functions and future observations y_{n+h}. If ϕ denotes a quantity of interest, the densities of interest are given by

$$p(\phi | D_n) = \int p(\phi | \psi, D_n) p(\psi | D_n) d\psi. \tag{3}$$

For many of the quantities above, $p(\phi | \psi, D_n)$ is analytically tractable. In the other cases, it is always possible to sample from it. Also, the marginal posterior of ψ is given by

$$p(\psi | D_n) \propto l(\psi | D_0) f_\psi(\psi | D_0) \tag{4}$$

where f_ψ is the prior density of ψ and $l(\psi | D_0)$ is the likelihood of ψ given by

$$l(\psi | D_0) = p(y_1, \ldots, y_n | \psi, D_0) = \prod_{t=1} p(y_t | \psi, D_{t-1}) \tag{5}$$

and each term in the right hand side of (5) is given by the Student t one-step-ahead predictive distribution.

Analytic integration of (3) will very rarely be possible but it can be approximated by

$$\hat{p}(\phi | D_n) = \frac{1}{N} \sum_{i=1}^{N} p(\phi | \psi^{(i)}, D_n) \tag{6}$$

where the $\psi^{(i)}$'s are a sample from $p(\psi | D_n)$. Unfortunately, the analytic dependance of (5), and consequently of (4), on ψ is far too complicated to allow the use of easy sampling schemes. Efficient resampling schemes become increasingly hard to build with increased dimensionality of ψ (Schmidt, Gamerman and Moreira (12)). In this paper, MCMC methods will be used to draw the values of ψ from (4).

2.2. Special cases

Univariate DLM's are obtained when $q = 1$ and Σ is a scalar. Typically, hyperparameters include unknown elements of the system variance matrix, for example $W_t = \mathrm{diag}(\psi_1, \ldots \psi_p)$.

Examples of dependence of x_t on ψ include the threshold and the smooth transition autoregressive models (see, for example, Chan and Tong (15) and Terasvirta and Anderson (16)). Examples of dependence of G_t on ψ include models for cycles. When cycles are present, the system equation can be written in trigonometric form or in the AR(2) form $\theta_t = \psi_1\theta_{t-1} + \psi_2\theta_{t-2} + w_t$ where $w_t \sim N(0; \sigma^2\psi_3^2)$, $-2 \le \psi_1 \le 2$ and $-1 \le \psi_2 \le 0$. In the notation of equation (2), this means that

$$G_t(\psi) = \begin{pmatrix} \psi_1 & \psi_2 \\ 1 & 0 \end{pmatrix} \text{ and } W_t(\psi) = \begin{pmatrix} \psi_3^2 & 0 \\ 0 & 0 \end{pmatrix}. \qquad (7)$$

The hyperparameter associated with the cycle is $\psi = (\psi_1; \psi_2; \psi_3)$ and the cycle wavelength λ and decay ρ are given by $\lambda = 2\pi/\cos^{-1}(\psi_1/2\rho)$ and $\rho = (-\psi_2)^{1/2}$. When $\psi_2 = -1$, there is no decay and the series exhibits a persistent cycle.

Another special case of (1)-(2) is obtained when x_t consist on previous values of the observation vector, i.e., $x_t' = (y_{t-1}'; \ldots; y_{t-r}')$. In fact, x_t may also include exogenous variables z_t but should necessarily include lagged values of the observed series. This is the time varying (or dynamic) vector autoregressive model discussed in Kitagawa and Gersch (14, ch. 12) and references therein. The can be written as

$$\begin{aligned} y_t &= \theta_t' x_t + v_t; \quad v_t \sim N(0; \Sigma) \\ \theta_t &= \theta_{t-1} + w_t; \ w_t \sim N(0; W_t(\psi); \Sigma) \end{aligned} \qquad (8)$$

and completed with prior $(\theta_0; \Sigma)|\psi; D_0 \sim NIW(m_0; C_0; \nu_0; S_0)$.

When the matrices of regression coefficents in (8) are fixed with time, reduced form VAR models are obtained with observation equation

$$y_t = \theta x_t + v_t; \quad v_t \sim N(0; \Sigma) \qquad (9)$$

and prior $(\theta; \Sigma)|D_0 \sim NIW(m_0; C_0; \nu_0; S_0)$ as before. In this simple case, full analytical analysis can be performed conditional on ψ.

Econometric theories are frequently used to specify contemporaneous as well as lagged relations between variables observed through time. This setup can be accommodated into a form that is similar to (9) but for the presence of restrictions in the form of Σ. Let A_0 be a non-singular matrix such that $\Sigma = A_0^{-1}(A_0^{-1})'$. Equation (9) can be rewritten as

$$A_0 y_t = \theta^* x_t + \epsilon_t; \quad \epsilon_t \sim N(0; I) \qquad (10)$$

where $\theta^* = A_0\theta$ is the new matrix of regression coefficents and $\epsilon_t = A_0 v_t$, $t = 1, \ldots, n$, are the vectors of structural shocks.

The identification of the model is achieved by the introduction of restrictions on the matrix A_0 of contemporaneous relations. The restrictions are suggested by theory and estimation of A_0 helps in assessing the fit of the theory to empirical data. The model can thus be exactly or over identified (Hamilton (2), pg. 332) and (10) is called the structural form of a VAR model. Exact inference cannot be performed for this model. Also, the model can be equivalently defined with diagonal entries of A_0 equal to 1 and shocks having a diagonal variance matrix $*$[1].

This section can be summarized in table 1 below.

TABLE I. Special cases of the common component DLM

model	model element q	Σ	G_t	W_t
univariate DLM	1	scalar	any	any
time-varying VAR	integer	any	I	any
structural VAR	integer	$A_0^{-1}(A_0^{-1})'$	I	0

2.3. Inference for structural VAR models

Although the model is just a special case of (1)-(2), it is useful to highlight a few aspects of the inference. The model can be written in terms of θ or θ^*. Opting for the first parametrization, gives the likelihood function

$$l(\theta; A_0|D_0) = \prod_{t=1} p(y_t|\theta; A_0; D_{t-1}) = (2\pi)^{-nq/2}|A_0|^n \exp\left\{-\frac{1}{2}Q(\theta; A_0)\right\}$$

where

$$Q(\theta; A_0) = \mathrm{tr}\left[\sum_{t=1} (y_t - \theta x_t)(y_t - \theta x_t)' A_0' A_0\right]$$

For the prior distribution, we will assume a conditional conjugate form $\theta|\Sigma; D_0 \sim N(m_0; C_0; \Sigma)$, where m_0 and C_0 may depend on ψ and $\psi|D_0$ with density f_ψ. This prior also implies a

[1] In this case, the contemporaneous matrix is denoted by A_0^* and $\Sigma = (A_0^*)^{-1}*[(A_0^*)^{-1}]'$.

normal prior for $\theta^*|\Sigma$. If m_0 and C_0 do not depend on ψ, the hyperparameter is reduced to the unknown elements of A_0 (and we will denote this by $\psi = A_0$) and the prior density for A_0 will be denoted by f_A. The non-informative prior is the limiting case $C_0^{-1} \to 0$.

The posterior distribution is obtained by Bayes theorem as $\theta|\psi, D_n \sim N(m_1, C_1, \Sigma)$ where $m_1 = C_1\left(C_0^{-1}m_0 + \sum_{t=1}^n x_t y_t'\right)$ and $C_1^{-1} = C_0^{-1} + \sum_{t=1}^n x_t x_t'$. Once again, inference conditional on hyperparameters is analytically tractable for the model with state parameters θ or θ^*. In any case, the state parameter can be integrated out, leading to $p(\psi|D_n) \propto |A_0|^n|C_0|^{-q/2}|C_1|^{q/2} f_\psi(\psi)\exp\{-\mathrm{tr}(n_1S_1A_0'A_0)/2\}$ where $n_1S_1 = (m_1 - m_0)'C_0^{-1}(m_1 - m_0) + \sum_{t=1}^n (y_t - m_1x_t)'(y_t - m_1x_t)$. When m_0 and C_0 do not depend on ψ, the marginal posterior distribution of the hyperparameter A_0 is

$$p(A_0|D_n) \propto |A_0|^n \exp\left\{-\frac{1}{2}\mathrm{tr}\,[n_1S_1A_0'A_0]\right\} f_A(A_0) \qquad (11)$$

which is not analytically tractable[2].

In the case of a non-informative prior for θ, $m_1 \to \hat{\theta} = (\sum_t x_t x_t')^{-1}\sum_t x_t y_t'$, $C_1 \to (\sum_t x_t x_t')^{-1}$ and (11) simplifies by replacement of $n_1S_1A_0A_0'$ by $Q(\hat{\theta}; A_0)$. This is the expression for the marginal posterior distribution obtained by Sims and Zha (11) but for the prior density f_A, which remains unspecified here. They make inference by simulation using a resampling technique (Rubin (17)). The basis for sampling is provided by a normal distribution obtained by a second order Taylor expansion of $\log p(A_0|D_n)$ about its mode.

Other quantities of interest also depending on A_0 are the measures of the propagation of the shocks ϵ_t over the observed variables. These measures are usually called impulse response functions and are obtained as the coefficients of $A^{-1}(L)$ where $A(L)$ is the autoregressive polynomial derived from (10) and L is the lag operator.

The fully Bayesian approach considers implicitly the uncertainty about Σ and therefore the distribution of the impulse response functions takes into account the uncertainty about all model parameters. This methodology can be naturally extended to structural VEC models where restrictions on A_0 are derived from the specification of suitable forms for the long-run impulse response functions as in King et al. (18) or Mellander, Vredin and Warne (19).

[2] If the $*$ parametrization is used and independent $IG(\alpha_i; \beta_i)$ distribution are set a priori for the λ_i's then they are still independent a posteriori given A_0^* with $IG(\alpha_i^*; \beta_i^*)$ distributions where $\alpha_i^* = \alpha_i + n/2$, $\beta_i^* = \beta_i + d_{ii}/2$ and $D = (d_{ij}) = A_0^*n_1S_1(A_0^*)'$ and $p(A_0^*|D_n) \propto |A_0^*|^n \prod_{i=1}^q \beta_i^{*-\alpha_i^*} f_A(A_0^*)$.

3. MCMC METHODOLOGY

We tackle the intractable marginal posterior of the hyperparameter by drawing samples from it. As already shown, the form of the density $p(\psi|D_n)$ in the models considered in this paper does not allow for simple, reliable sampling schemes. Sims and Zha (11) reported large variation in the resampling weights for structural VAR models. Schmidt, Gamerman and Moreira (12) encountered similar problems in the context of dynamic models.

The more sophisticated MCMC methods are particularly suitable in those situations. They provide a sample of the posterior of interest by embeding it in a Markov chain as an equilibrium distribution and simulating a trajectory from the chain until equilibrium is reached. An introductory account of MCMC is provided in Gamerman (20). Further advanced reading and areas of application are given in Gilks, Richardson and Spiegelhalter (21).

In this paper, chains formed by Metropolis-Hastings algorithms are used. This means that chain moves are made in two steps: a proposal transition and an acceptance/rejection of the move proposed. Moves can be proposed according to transition kernels in random walk forms $q(\psi^{(old)}, \psi^{(new)}) = g(\psi^{(new)} - \psi^{(old)})$ where g is the density of a distribution symmetric around zero, e.g. $N(0, C)$. These moves can also be made componentwise according to univariate $N(0, c)$ distributions. The acceptance probability for random walk proposals is given by $\alpha(\psi^{(old)}, \psi^{(new)}) = \min\{1, p(\psi^{(new)}|D_n)/p(\psi^{(old)}|D_n)\}$. In theory, these proposals are not acceptable for parameters with limited variation such as system variances and should be adapted with the corresponding truncation probabilities. In our applications, the range of likely values for the hyperparameters lead to effective proposals with truncation probabilities negligibly small. If, however, any of the parameters is close to the boundary then truncation probabilities cannot be discarded and the acceptance probability will no longer have a simple form. In these cases, other proposals suggested below can just as easily be used. The value of the random walk variance is crucial in ensuring that chain moves are reasonably paced towards equilibrium. Therefore, whenever a value for the variance is leading to large (small) acceptance rates it is automatically reduced (increased) to allow for suitable values of the acceptance rate and hence faster convergence.

There are other possibilities for the proposal kernel. One can consider likelihood-based or even prior-likelihood-based normal forms for the proposal (Gamerman (22)). One possibility is a normal distribution centered on the posterior mode and with precision matrix given by the observed posterior information matrix. Evaluation of the posterior mode would require a maximization algorithm that typically also provides the information matrix. Another attractive alternative for moves on variances and variance matrices is to use a generalized form of random walk where the proposal is centered around the previous chain value but Gamma and Wishart distributions are used instead of normal forms. For these cases, the acceptance probability requires more computational effort than for the random walk above.

To avoid chains getting trapped in local modes, it is sometimes recommended to have a few chains starting from (preferably overdispersed) initial points. Chains which appear to show convergence towards local, insignificant modes, as measured by their posterior density values, are discarded. In the applications it was not difficult to specify a likely region for the hyperparameters which allows easy specification of reasonable initial points near or at the border of the region.

Once convergence is ascertained according to some of the many methods available (Gelman and Rubin (23); Geweke (24)), values from the chains form an approximate sample $\psi^{(1)}, \ldots, \psi^{(N)}$ from (4) or (11). Inference about the hyperparameter is then based on these values. If interest centers on the state parameters θ, some of their transformations $h(\theta)$ or future values of the series then their marginal distribution can be approximated according to (6). If the transformations are more complicated as is the case for impulse responses, then inference for them can also be based on samples $h(\theta^{(j)})$'s obtained after drawing the $\theta^{(j)}$'s from $\hat{p}$ in (6).

The full MCMC algorithm can be summarized as follows

1. start the chain with $\psi^{(1)}$, evaluate the posterior density at $\psi^{(1)}$ and set $j = 1$;

2. propose a new value $\psi^{(new)}$ according to $\psi^{(new)} \sim N(\psi^{(j)}; C)$ either in block or componentwise and evaluate the posterior density at $\psi^{(new)}$;

3. take $\psi^{(j+1)} = \psi^{(new)}$ with probability $\alpha(\psi^{(j)}, \psi^{(new)})$, otherwise take $\psi^{(j+1)} = \psi^{(j)}$;

4. check convergence according to the criteria above and if the chain is assumed to have converged go to the next step only for further N iterations;

5. set $j \to j + 1$ and return to (2).

4. APPLICATIONS

Three different applications were made. The first one to the series of the Brazilian industrial production index, the last ones to a multivariate study of government policy making. Vague prior distributions were used for the hyperparameters according to the relevant information available. They were taken as independent and truncated normal for the system standard deviations and uniform for the other parameters. Truncation was caused by natural restriction (on variances, for example) and also by meaningful interpretation on the parameter in question.

In all the applications, 4 chains were run from initial overdispersed values. These values were chosen from prior knowledge of the likely regions for the hyperparameter. They were taken as the endpoints of intervals for components that were uniformly distributed and the 0.01 and 0.99 quantiles for components that were normally distributed. Chains were run until convergence was diagnosed according to the Gelman and Rubin (23) and Geweke (24) diagnostics applied to the posterior density. After convergence, samples of size 2000 were stored for inference from all the chains. The pace of the moves between iterations was set by increasing and/or decreasing the proposal random walk variance in such a way as to have the acceptance rates between 30% and 60%. These proposal have been extensively used in the literature with good results reported. This was confirmed in our simulations. Therefore, alternative forms outlined above were not used in the applications.

4.1. Univariate dynamic linear model

The industrial production index is the main indicator of economic activity in Brazil and is measured monthly by the official Brazilian Institute of Statistics and Geography (IBGE). In this application, the data goes from Jan/1981 to Jun/2001. The model chosen is a univariate DLM with observation equation

$$y_t = \mu_t + s_t + c_t + \alpha_t x_t + v_t\,; \quad v_t \sim N(0; \sigma^2)$$

where μ_t, s_t and c_t respectively denote the trend, seasonal and cycle components and x_t is the only explanatory variable with the number of working days in month t. The trend is subject to a local linear growth

$$\begin{aligned} \mu_t &= \mu_{t-1} + \beta_t + w_{1t}\,; \quad w_{1t} \sim N(0; \sigma^2\psi_3^2) \\ \beta_t &= \beta_{t-1} + w_{2t}\,; \quad w_{2t} \sim N(0; \sigma^2\psi_4^2) \end{aligned}$$

The seasonal component reflects the production pattern and follows a permutation evolution governed by a zero-sum constraint on effects and driven by a $(p-1)$-dimensional disturbance vector w_{St}. The cycle component reflects short range fluctuations and its evolution was described by (7). The regression coefficient follows a locally constant random walk with disturbance w_{Rt}. The seasonal, cycle and regression evolutions disturbances have respective standard deviations $\sigma\psi_5$, $\sigma\psi_6$ and $\sigma\psi_7$.

The model can be phrased in terms of (1)-(2) with $\theta_t = (\mu_t, \beta_t, s_t, \ldots, s_{t-10}, c_t, c_{t-1}, \alpha_t)'$, $G_t(\psi) = \mathrm{diag}(G_T, G_S, G_C(\psi), 1)$ and $W_t(\psi) = \mathrm{diag}(\psi_3^2, \psi_4^2, \psi_5^2 I_{11}, \psi_6^2, 0, \psi_7^2)$ and $\Sigma = \sigma^2$ where $G_T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $G_S = \begin{pmatrix} 0_{10} & I_{10} \\ -1 & -1 \cdot 1_{10}' \end{pmatrix}$, G_C is as given by (7) and 0_m (1_m) is a m-dimensional vector of 0's (1's). The hyperparameter is therefore $\psi = (\psi_1, \psi_2, \psi_3, \psi_4, \psi_5, \psi_6, \psi_7)$.

The prior for the cycle parameters was taken as uniform over $[0.5, 2] \times [-1, 0]$. The system standard deviations had prior means $2, 1, 0.5, 1$ and 2 and standard deviations $0.66, 0.33, 0.15, 0.33$ and 0.66, before truncation. The respective variances of the componentwise random walk proposals were 0.2, 0.2, 1.5, 0.2, 0.2, 1.0 and 0.5.

The summary of the posterior inference is provided in Tables 2 and 3 and Figures 1 and 2 below. The dynamic movement of the growth and the seasonal component are rather small and perhaps, these components could be taken as constant over time. Table 3 contains predictions at the end of the series for 1 up to 6 months ahead . Note that the largest (negative) correlation is again that between ψ_1 and ψ_2. The other correlations are reasonably small. Therefore, we do not expect to have convergence difficulties for single move chains in such situations. Also, as expected, the lengths of the predictive intervals increase with the prediction horizon.

TABLE II. Summary of estimation for application 1 - Brazilian industrial production index data

	Mode	Mean	S.D.	95% limits
$\tilde{A}_1$	1.44	1.41	0.10	(1.17, 1.60)
$\tilde{A}_2$	-0.69	-0.61	0.10	(-0.83, -0.38)
$\tilde{A}_3$	1.25	1.10	0.47	(0.17, 2.05)
$\tilde{A}_4$	0.01	0.10	0.07	(0.00, 0.28)
$\tilde{A}_5$	0.11	0.13	0.04	(0.06, 0.23)
$\tilde{A}_6$	0.68	1.02	0.27	(0.58, 1.60)
$\tilde{A}_7$	0.09	0.18	0.08	(0.04, 0.36)

TABLE III. Summary of prediction for application 1 - Brazilian industrial production index data

	Mean	S.D.
y_{n+1}	131.2	3.3
y_{n+2}	134.2	4.2
y_{n+3}	132.3	5.2
y_{n+4}	136.9	6.0
y_{n+5}	131.6	6.6
y_{n+6}	119.6	7.1

The correlation matrix of $\tilde{A}$ is

$$\begin{pmatrix} 1 & & & & & & \\ ¡0{:}87 & 1 & & & & & \\ 0{:}10 & ¡0{:}22 & 1 & & & & \\ ¡0{:}05 & ¡0{:}02 & ¡0{:}41 & 1 & & & \\ 0{:}15 & ¡0{:}16 & 0{:}35 & 0{:}00 & 1 & & \\ ¡0{:}48 & 0{:}57 & ¡0{:}10 & 0{:}15 & 0{:}29 & 1 & \\ 0{:}04 & ¡0{:}05 & 0{:}18 & 0{:}10 & 0{:}40 & 0{:}34 & 1 \end{pmatrix}$$

It is interesting to note the large variation between values of the system standard deviations with some system volatilities exhibiting large values while other system components show virtually no time volatility. Also, the trend volatility is larger than the observational volatility. The results again con¯rm the large correlation between cycle parameters as the only relevant correlation between hyperparameters. The predictions however are very well behaved and their predictive distributions can very well be approximated by normal forms, in this case. Examination of the fitted residuals shows a significant but small autocorrelation of order 1 (the 95% confidence interval is [0.17, 0.33]) but no further relevant autocorrelation structure. We have therefore considered the model to be reasonably adequate.

4.2. Time-varying VAR

The high volatility of the Brazilian economy led to large and persistent inflation rates. The rates forced private and public agents to adopt defensive mechanisms such as price indexation and the Government to design successive stabilization plans to interfere with the mechanism of price corrections.

The form that monetary policies have been implemented in Brazil suggests that the most parsimonious list of variables capable of describing the price dynamics is given by $y_t = (P_t, E_t, I_t)$ where P_t is the consumer price index (INPC) series, E_t is the series of exchange rates Real/USD and I_t is 1 + the (SELIC) rate of public bonds (all measured in logs). The effect of nominal shocks over price can be described by a VAR model with lag $r = 3$ that dynamically relates these variables. The VAR coefficients implicitly describe the correction mechanisms.

The choice of the lag was based on various considerations. Model selection criteria such as BIC when applied to the static version of this model pointed to $r = 3$. Also, it allows for complex roots for the AR polynomial and can be used to model dampened cycles. Finally, parsimony leads to the choice of the smallest possible number of meaningful lags.

The economic agents are expected to adapt their price adjustments mechanisms to the varying conditions, in particular, to the alterations in the inflation rate and to the stabilization plans. This may imply fluctuation of the VAR coefficients through time. These fluctuations are expected to be higher when a stabilization plan takes place.

In this application, we have used monthly data collected in Brazil from Apr/73 to Dec/98. The model used is given by (8) with two regimes of time variation. One of them takes place at the times $T = \{Jun/1980, Jun/1986, Jul/1990, Oct/1994\}$. Given the autoregressive structure of the model, we have discarded data information of the r months preceding the times in T. They correspond to r months after an actual plan explictly attempting to alter the indexation mechanisms was applied. The other regime is applied to all other times. The hyperparameter is $\psi = (\psi_1, \psi_2)'$ and the values of $W_t(\psi)$ are $\psi_1 I$, if $t \in T$ and $\psi_2 I$, if t 2 T . The hyperparameter have prior mean (0:05; 0:5) and variance diag(0:001; 0:1). The respective variances of the componentwise random walk proposals were 0.02 and 0.3.

Inference about Ã illustrates the degree of adaptation su®ered by the economy at periods of instability and at periods of introduction of economic plans. It can be summarized in ¯gures 3 and 4 and table 4 below.

TABLE IV. Summary of inference for application 2 - time-varying VAR

	Mode	Mean	S.D.	95% limits	Correlation	
$\tilde{A}_1$	0.000	0.003	0.002	(0.000, 0.008)	1	
$\tilde{A}_2$	0.017	0.192	0.158	(0.000, 0.528)	0.398	1

The system variance at structural changes is substantially higher than the variance at other times, re°ecting the anticipated impact of the stabilization plans. There is great uncertainty about $\tilde{A}_1$ because of the scarcity of data points and these ¯gures should be interpreted with care. The hyperparameters are not strongly correlated. The modal integrated log-posterior of the time-varying VAR is 7.28 points higher than the integrated static VAR log-posterior and 2.33 points higher than the VAR varying only in T (i.e., $\tilde{A}_1 = 0$) indicating the better ¯t of the fully time-varying VAR.

4.3. Structural VAR

Structural VAR models have been used to analyse the e®ects of changes in economic policies. There is an extensive list of papers analysing the e®ects of public policies that use structural VAR models. Relevant work in the area of monetary policy is given by Cristiano, Eichenbaum and Evans (25), Sims (26) and Sims and Zha (27). Relevant references in the area of open economy are Eichenbaum and Evans (28), King et al. (18) and Mellander, Vredin and Warne (19).

The use of this methodology is a consequence of the endogeneity of the relevant variables and, therefore, of the impossibility of direct measurement of the exogenous components of economic policy changes. The model residuals are uncorrelated, exogenous shocks that under certain conditions can be interpreted as the impact of unobserved policy changes.

The Brazilian economy had experienced high in°ation rates and successive stabilization plans that promoted instability in the nominal variables up to July 1994, as shown in the previous application. Fiorencio and Moreira (29) show that the economy moved smoothly towards a stable phase after the last plan. Therefore, only data after July 1994 and a static version of the model is considered here. In structural VAR models, the matrix A_0 has p^2 entries of which only $\binom{p}{2}$ are linearly independent and $p(p+1)/2$ restrictions must be imposed.

The standard estimation procedure is maximum likellihood but the distribution of estimators is difficult to obtain and the relevant significance of parameters cannot be ascertained. The restrictions imposed here are:

1. the exchange rate E is not contemporaneously affected by $(I; P)$
2. the interest rate I is affected by E and P, but the P coefficient is restricted to -1 to represent the effect of E on the real interest rate $I - P$
3. the price index P is affected by $(E; P)$
4. the diagonal elements are all 1

With these restrictions, the matrix of contemporaneous effects is given by

$$A_0^* = \begin{pmatrix} 1 & \psi_1 & \psi_2 \\ 0 & 1 & 0 \\ -1 & \psi_3 & 1 \end{pmatrix};$$

with a 3-dimensional hyperparameter ψ. The model for $y_t = (P_t; E_t; I_t)$ becomes $A_0 y_t = \bar{y}_t + \epsilon_t$ where $\bar{y}_t = A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3}$ and $\epsilon_t = (\epsilon_{P;t}; \epsilon_{E;t}; \epsilon_{I;t})' \sim N(0; \Lambda)$. The structural equations are individually given by

$$\begin{aligned} P_t &= -\psi_1 E_t - \psi_2 I_t + \bar{P}_t + \epsilon_{P;t} \\ E_t &= \bar{E}_t + \epsilon_{E;t} \\ I_t &= P_t - \psi_3 E_t + \bar{I}_t + \epsilon_{I;t} \end{aligned}$$

where $(\bar{P}_t; \bar{E}_t; \bar{I}_t)' = \bar{y}_t$.

This is a simultaneous system of equations where economic theory suggests that shocks on interest rates have a dampening effect on price over time and that a shock on the exchange rate has a positive effect on prices. It is particularly relevant to measure the magnitude of that effect. It is also expected that a price innovation will cause a rise on interest rates. The reasoning above suggests that $\tilde{A}_2 > 0$, $\tilde{A}_1 < 0$ and $\tilde{A}_3 < 0$. All these parameters are short-run elasticities and were given independent N(0,1) priors. The variances of the componentwise random walk proposals were all taken as 0.2.

In this application, we have used the same data of the previous section. Inference about the hyperparameter is summarized on Table 5 and Figures 5 and 6. They con¯rm empirically that the parameters lie mostly in the regions suggested by the theoretical reasoning above, even though this was not imposed a priori. All components of the hyperparameter are signi¯cantly di®erent from zero. The e®ect of an exogenous increase of the interest rates, namely restrictive monetary policies, is only partially incorporated into prices. The exchange rate is not as a®ected by restrictive monetary policies as expected. Another interest in these types of study is the determination of the impulse response functions. These are depicted on Figure 7 along with their respective credibility bands.

TABLE V. Summary of inference for application 3 - structural VAR

	Mode	Mean	S.D.	95% limits	Correlation		
$\tilde{A}_1$	-0.236	-0.256	0.119	(-0.612, 0.058)	1		
$\tilde{A}_2$	0.822	0.914	0.239	(0.377, 1.650)	-0.513	1	
$\tilde{A}_3$	-0.319	-0.313	0.116	(-0.635, 0.023)	-0.001	0.000	1

4.4. Comments

The main purpose of this section was to highlight the °exibility of the methodology in three di®erent classes of models that share a common structure. A number of issues relating to DLM require further clari¯cation. One referee raised an interesting question about the suitability of the use of trended regressors. Discussion of these and many other properties of DLM can be found in West and Harrison (13). Use of MCMC to these models is also described in the above book and simulation studies were carried out by Carter and Kohn (4) and Fruhwirth-Schnatter (5).

In terms of the applications themselves, further studies are called for in terms of model comparison and model determination. The ¯rst application suggests that a constant trend growth is a plausible alternative with $\tilde{A}_4 = 0$. The second application seems to indicate constancy of VAR coe±cients when no stabilization plan was being introduced. In the third application, models leading to anticipated forms for the impulse response should also be entartained. These considerations should lead to a fuller comparison of alternative models than the one made here. This is an important component of an econometric analysis but goes beyond the illustrative purpose of the present paper.

In terms of VAR models, it has been common practice to check for unit roots and cointegration. The unrestricted VAR form is preferred here because unit roots test have low power (Campbell and Perron (30)), specially in the case of a short time span. Fiorencio and Moreira (29) discuss this point further suggesting that the unit root approach may not be appropriate in this context.

5. CONCLUDING REMARKS

This paper considers the problem of inference in dynamic models by a sampling based approach for hyperparameters and analytic integration for state parameters. It was shown that this strategy is particularly useful in commonly used state-space models and structural VAR models. It is to be expected that the same methodology could be applied to other models used in Economics, Finance and other fields of Science. Many authors have shown the importance of GARCH models and spatial autocorrelation models. Their mathematical form is not entirely different from the ones used in this paper and we can conjecture at this point that it will be possible to adapt them for application of the methodology presented here.

This methodology is easy to use, flexible and in fact has been implemented for experimental use in PRV for Windows (Moreira (31)). It allows specification of different form of priors for the hyperparameter, including (but not exhausted by) the conditional conjugate forms used in conjunction with the Gibbs sampling methodology. It detects secondary modes by monitoring the integrated posteriors (4) and (11). It also adapts for appropriate pace of the movements of the parallel chains. The software is freely available and can be downloaded from the site `http://www.ipea.gov.br`.

APPENDIX: Normal inverse wishart distributions

A multivariate normal distribution with vector mean μ and covariance matrix Σ is denoted by $N(\mu; \Sigma)$. A $p \times q$ matrix X is said have a matricvariate normal distribution with vector mean μ, left covariance matrix V and right covariance matrix Σ, denoted by $N(\mu; V; \Sigma)$, if $vec(X) \sim N(\mu; V \otimes \Sigma)$, where $vec(X)$ denotes the column vectorization of X.

A $p \times p$ matrix X is said to have a Wishart distribution with ν degrees of freedom and scale matrix S, denoted $W(\nu; S)$, if its density is given by

$$k|X|^{(\nu - p - 1)/2} \exp\left\{ -\frac{1}{2} \mathrm{tr}(\nu S^{-1} X) \right\};$$

if X is a positive definite matrix, and 0, otherwise where k is a normalizing constant. If $X \sim W(\nu; S)$ then X^{-1} has an inverse Wishart distribution with ν degrees of freedom and scale matrix S, denoted $IW(\nu; S)$. When $p = 1$ and X becomes a scalar quantity, (inverse) Wishart distributions are called (inverse) Gamma.

If $Y|\Sigma \sim N(\mu; V; \Sigma)$ and $\Sigma \sim IW(\nu; S)$ then the pair $(Y; \Sigma)$ is said to have Normal-inverse Wishart distribution with parameters $\mu; V; \nu$ and S, denoted by $NIW(\mu; V; \nu; S)$. As a consequence, the marginal distribution of Y is a matricvariate Student t distribution with ν degrees of freedom, location parameter μ, right scale parameter V and left scale parameter S, denoted by $t_\nu(\mu; V; S)$. This means that $vec(Y) \sim t_\nu(\mu; V \otimes S)$. If Y is a vector then V is scalar and the matricvariate results above become multivariate results.

ACKNOWLEDGEMENTS

The research of the authors was supported by a grant from Ministério de Ciência e Tecnologia, Brazil (PRONEX). The second author also benefitted from grants from FAPERJ, Brazil, IPEA and CNPq, Brazil. The authors gratefully acknowledge the computational help provided by Elaine A. Simone and Ingreed Valdez, insightful comments by the referees and comments on an earlier draft by Helio Migon, Hedibert Lopes and Eduardo Gutierrez-Peña.

BIBLIOGRAPHY

(1) Harvey, A. C. Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge, 1989.

(2) Hamilton, J. Time Series Analysis, Princeton University Press, New Jersey, 1994.

(3) Carlin, P. B.; Polson, N. G.; and Stoffer, D. S. A Monte Carlo approach to nonnormal and nonlinear state-space modeling. Journal of the American Statistical Association, 1992, 87, 493-500.

(4) Carter, C. K.; Kohn, R. On Gibbs sampling for state space models. Biometrika, 1994, 81, 541-53.

(5) Fruhwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. Journal of Time Series Analysis, 1994, 15, 183-202.

(6) Chib, S.; Greenberg, E. Hierarchical analysis of SUR models with extensions to correlated serial errors and time-varying parameter models. Journal of Econometrics, 1995, 68, 339-360.

(7) Shephard, N.; Pitt, M. K. Likelihood analysis of non-Gaussian measurement time series. Biometrika, 1997, 84, 653-667.

(8) Gamerman, D. Markov chain Monte Carlo for dynamic generalized linear models. Biometrika, 1998, 85, 215-227.

(9) Kadiyala, K. R.; Karlsson, S. Numerical methods for estimation and inference in Bayesian VAR-models. Journal of Applied Econometrics, 1997, 12, 99-132.

(10) Lopes, H. F.; Moreira, A. R. B.; Schmidt, A. M. Hyperparameter estimation in forecast models. Computational Statistics & Data Analysis, 1999, 29, 387-410.

(11) Sims, C. A.; Zha, T. Error bands for impulse responses. Cowles Foundation discussion paper no. 1085, 1995.

(12) Schmidt, A. M.; Gamerman, D.; Moreira, A. R. B. An adaptive resampling scheme for cycle estimation. Journal of Applied Statistics, 1999, 26, 619-641.

(13) West, M.; Harrison, P. J. Bayesian Forecasting and Dynamics Models, 2nd ed., Springer-Verlag, New York, 1997.

(14) Kitagawa, G.; Gersch, W. Smoothness Priors Analysis of Time Series, Lecture Notes in Statistics 116. Springer-Verlag, New York, 1996.

(15) Chan, K. S.; Tong, H. On estimating thresholds in autoregressive models. Journal of Time Series Analysis, 1986, 7, 179-190.

(16) Terasvirta, T.; Anderson, H. M. Characterizing nonlinearities in business cycles using smooth transition autoregressive models. Journal of Applied Econometrics, 1992, 7, S119-S132.

(17) Rubin, D. R. Using the SIR algorithm to simulate posterior distributions (with discus-

sion), in Bayesian Statistics 3 (eds J. M. Bernardo et al.), Oxford University Press, Oxford, pp. 395-402, 1988.

(18) King, R. G.; Plosser, C. I.; Stock, J. H.; Watson, M. H. Stochastic trends and economic °uctuation. The American Economic Review, 1981, 81, 819-840.

(19) Mellander, E.; Vredin, A.; Warne, A. Stochastic trends and economic °uctuation in a small open economy. Journal of Applied Econometrics, 1992, 7, 269-394.

(20) Gamerman, D. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, Chapman & Hall, London, 1997.

(21) Gilks, W. R.; Richardson, S.; Spiegelhalter, D.J. (eds.) Markov Chain Monte Carlo in Practice, Chapman & Hall, London, 1996.

(22) Gamerman, D. Sampling from the posterior distribution in generalized linear mixed models. Statistics and Computing, 1997, 7, 57-68.

(23) Gelman, A.; Rubin, D. R. Inference from iterative simulation using multiple sequences (with discussion). Statistical Science, 1992, 7, 457-511.

(24) Geweke, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments (with discussion), in Bayesian Statistics 4 (eds. J. M. Bernardo et al.), 169-193, University Press: Oxford, 1992.

(25) Cristiano, L. M.; Eichenbaum, M.; Evans, C. The e®ect of monetary policy shocks: evidence from the °ow of funds. Review of Economics and Statistics, 1994, 78, 16-34.

(26) Sims, C. A. Are forecasting models usable for policy analysis. Quarterly Review of the Federal Reserve Bank of Minneapolis, 1986, 2-16.

(27) Sims, C. A.; Zha, T. Bayesian methods for dynamic multivariate models, Technical Report, Yale University, 1996.

(28) Einchenbaum, E.; Evans, C. Some empirical evidence on the e®ects of monetary policy shocks on exchange rates. Quarterly Journal of Economics, 1993, 975-1009.

(29) Fiorencio, A.; Moreira, A. R. B. Measuring the stability of the price system. Economic Modelling, 2001, 18, 381-397.

(30) Campbell, J. Y.; Perron, P. Pitfalls and oportunities : what macroeconomist should know about unit roots, NBER working papers, 1991.

(31) Moreira, A. R. B. PRV for Windows manual. Technical Report, IPEA, Brazil, 1997.

Figure 1

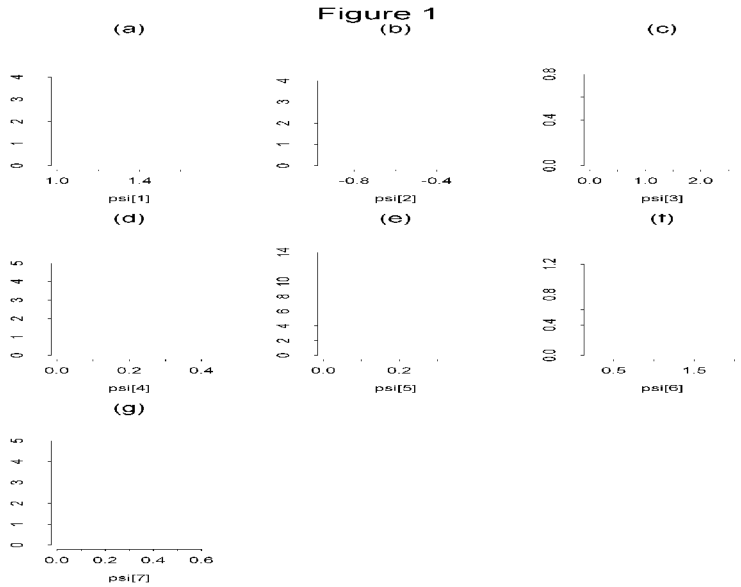


Figure 2

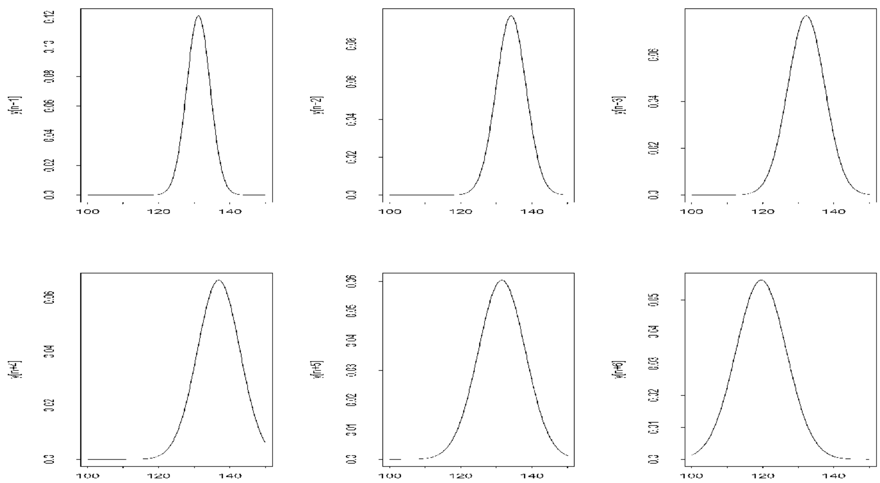


Figure 3

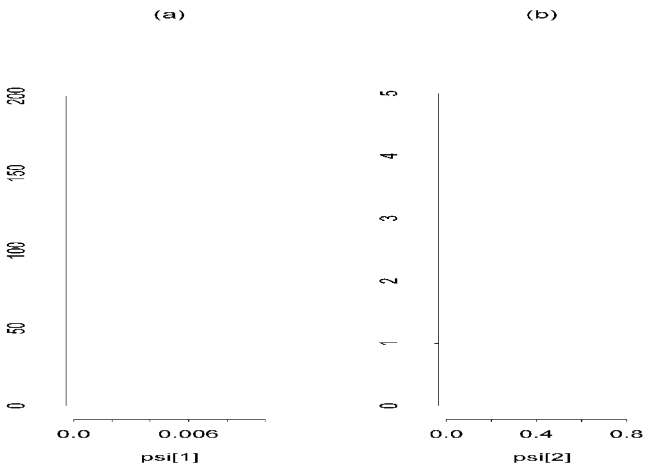


Figure 4

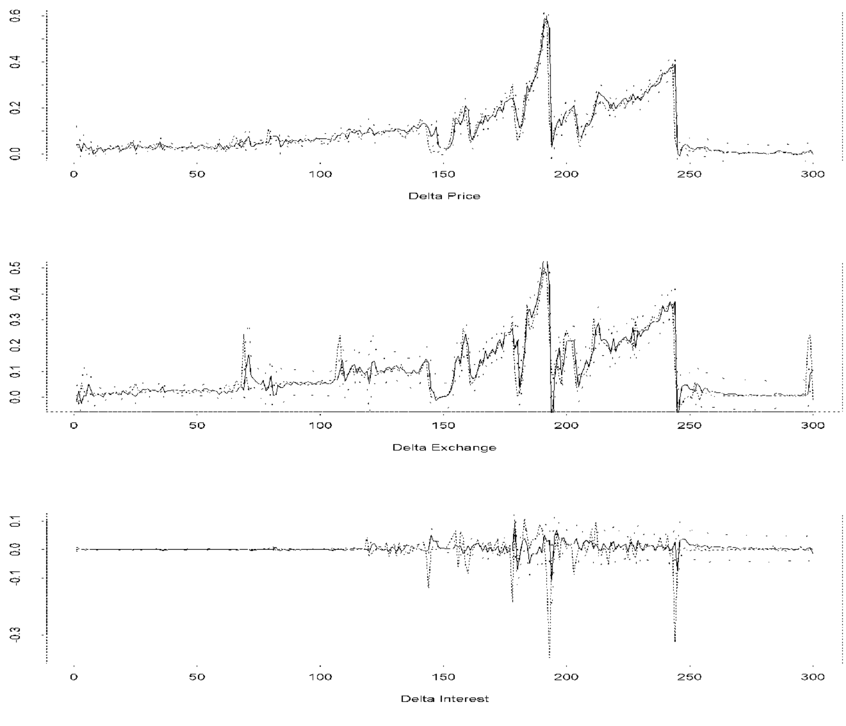


Figure 5

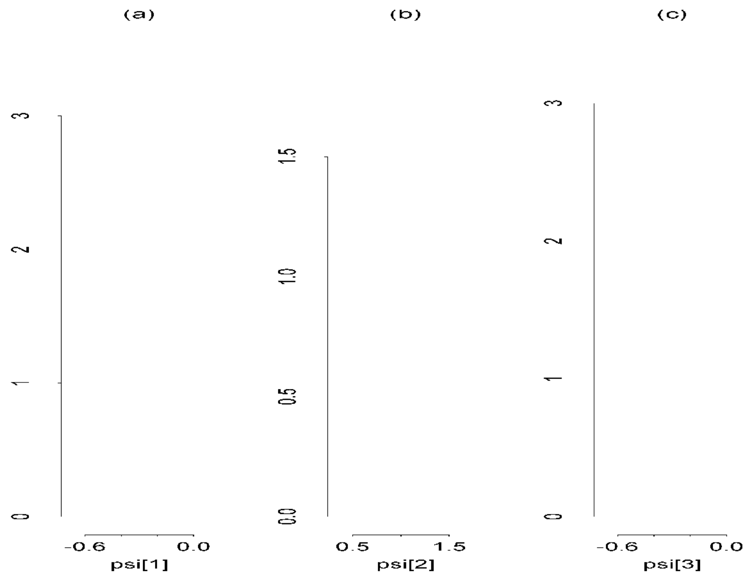


Figure 6

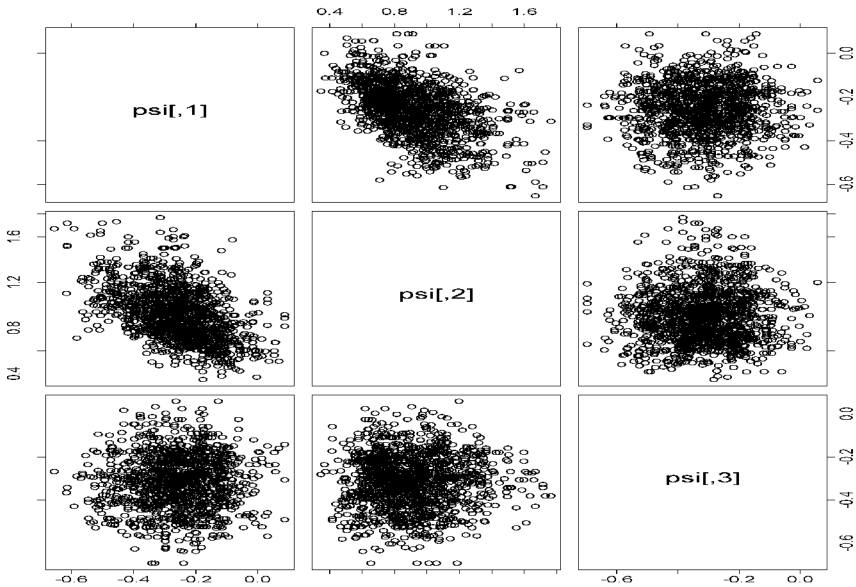
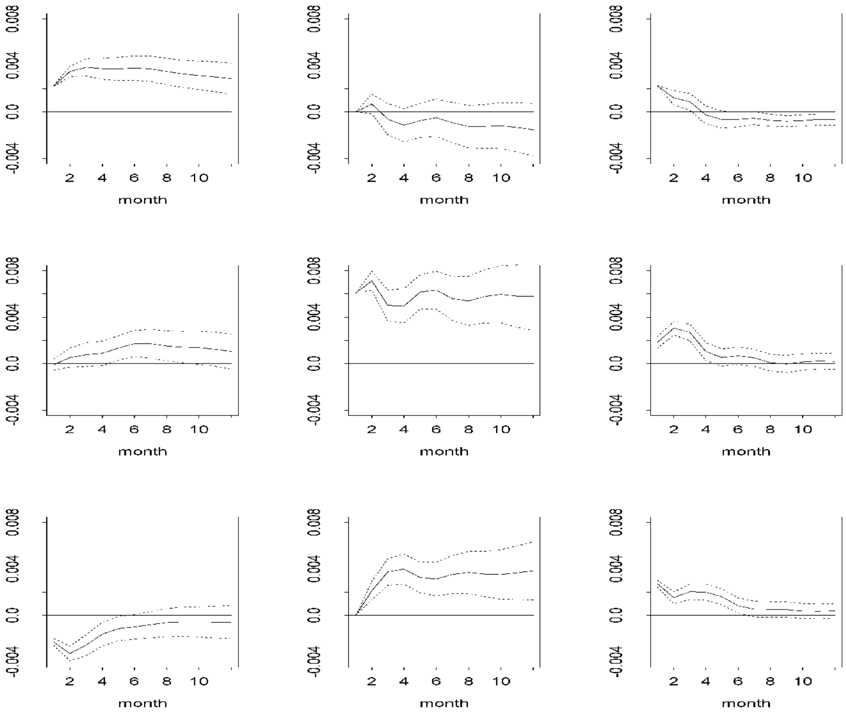


Figure 7



Caption for the ¯gures

Figure 1. Summary of posterior inference for application 1 - histograms of the components of the hyperparameter: a. $\tilde{A}_1$; b. $\tilde{A}_2$; c. $\tilde{A}_3$; d. $\tilde{A}_4$; e. $\tilde{A}_5$; f. $\tilde{A}_6$; g. $\tilde{A}_7$.

Figure 2. Summary of posterior inference for application 2 - 1 up to 6 step ahead predictive distributions.

Figure 3. Summary of posterior inference for application 2 - histograms of the components of the hyperparameter: a. $\tilde{A}_1$; b. $\tilde{A}_2$.

Figure 4. Summary of posterior inference for application 2 - ¯tted mean responses (full line) with two s. d. limits (dashed lies) and data (dots) on $(P; E; I)$. The data is represented in terms of ¯rst order di®erences for visual clarity.

Figure 5. Summary of posterior inference for application 3 - histograms of the 3 components of the hyperparameter: a. $\tilde{A}_1$; b. $\tilde{A}_2$; c. $\tilde{A}_3$.

Figure 6. Summary of posterior inference for application 3 - pairwise plots of the 3 components of the hyperparameter.

Figure 7. Estimates of the impulse response functions: top row, responses due to shock on prices; middle row, responses due to shock on exchange rate policy; bottom row, responses due to shock on monetary policy; 1st. column: responses on prices; 2nd. column: responses on exchange rates; 3rd. column: responses on interest rates. Point estimates are depicted in full lines and one standard deviation limits are provided in dashed lines.

Ipea – Institute for Applied Economic Research

PUBLISHING DEPARTMENT

Coordination
Cláudio Passos de Oliveira

Supervision
Everson da Silva Moura
Reginaldo da Silva Domingos

Typesetting
Bernar José Vieira
Cristiano Ferreira de Araújo
Daniella Silva Nogueira
Danilo Leite de Macedo Tavares
Diego André Souza Santos
Jeovah Herculano Szervinsk Junior
Leonardo Hideki Higa

Cover design
Luís Cláudio Cardoso da Silva

Graphic design
Renato Rodrigues Buenos

The manuscripts in languages other than Portuguese published herein have not been proofread.

Ipea Bookstore

SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo
70076-900 – Brasília – DF
Brazil
Tel.: + 55 (61) 3315 5336
E-mail: livraria@ipea.gov.br

Composed in Adobe Garamond 11/13.2 (text)
Frutiger 47 (headings, graphs and tables)
Brasília – DF – Brazil

Ipea's mission
Enhance public policies that are essential to Brazilian development by producing and disseminating knowledge and by advising the state in its strategic decisions.

ISSN 1415-4765

ipea Institute for Applied Economic Research

Secretariat of Strategic Affairs

BRAZILIAN GOVERNMENT
BRASIL