BANCO CENTRAL DE RESERVA DEL PERÚ

Forecasting Peruvian Macroeconomic Variables with Bayesian Vector Autorregressions with Time-Varying in the mean

Fernando Pérez Forero*

* Banco Central de Reserva del Perú

DT. N°. 2021-001
Serie de Documentos de Trabajo
Working Paper series
Marzo 2021

Forecasting Peruvian Macroeconomic Variables with Bayesian Vector Autorregressions with Time-Varying in the mean*

Fernando J. Pérez Forero†

March 17, 2021

Abstract

Macroeconomic Forecasting in a changing and uncertain environment over time is a great challenge today. This paper uses a Bayesian VAR with a time-varying mean and stochastic volatility, in order to elaborate forecasts for the Peruvian economy. The model is flexible enough to consider the structural changes that potentially occur in the economy. Forecasts are made mainly for variables such as inflation and GDP growth, although the model might be adapted to include other variables. The empirical model uses information from the survey of macroeconomic expectations as observables linked to the long-term means, following Banbura and van Vlodrop (2018). Results show a good fit, and reaffirm the idea associated with the use of expectations surveys to reduce long-term uncertainty, while the time varying parameters improve the predictive power of the model.

Resumen

Realizar predicciones macroeconómicas en un entorno cambiante en el tiempo e incierto es hoy en día un gran desafío. Este trabajo utiliza un modelo VAR Bayesiano con una media cambiante en el tiempo y volatilidad estocástica, y para así elaborar proyecciones para Perú. Estas propiedades mencionadas le brindan al modelo suficiente flexibilidad para considerar los cambios estructurales que potencialmente se registren en la economía. Las proyecciones se realizan principalmente para variables como inflación y crecimiento del PBI, aunque el modelo es suficientemente flexible como para ser adaptado en el futuro hacia el uso de otras variables. Este ejercicio utiliza información de la encuesta de expectativas macroeconómicas como observables para estimar las medias de largo plazo, siguiendo a Banbura and van Vlodrop (2018). Los resultados muestran un buen ajuste, y reafirman la idea asociada a que el uso de encuestas de expectativas permite reducir la incertidumbre a largo plazo, a la vez que los parámetros cambiantes en el tiempo mejoran el poder predictivo del modelo dinámico utilizado.

JEL Classification: C11, C32, C53, E37, E47

Key words: Density Forecasts, Stochastic Volatility, Time-Varying Parameters

*I would like to thank Carlos Barrera, Carlos Montoro, Marco Vega, and seminar participants at the BCRP for their helpful comments and suggestions. The views expressed are those of the author and do not necessarily reflect those of the Central Reserve Bank of Peru. All remaining errors are mine.

†Deputy Manager of Monetary Policy Design, Central Reserve Bank of Peru (BCRP), Jr. Santa Rosa 441, Lima 1, Perú; Email address: fernando.perez@bcrp.gob.pe

1 Introduction

Macroeconomic Forecasting in a changing and uncertain environment over time is a great challenge today. As a matter of fact, density forecasts became a relevant output for policy makers and forecasters, especially after observing the end of the Great Moderation and the outbreak of the Great Financial Crisis (Clark, 2011). That is, instead of providing simple point-forecasts, in a Bayesian setup it is possible to exploit the estimated uncertainty of the parameter set from the data (and some priors), and to combine it efficiently to produce fan charts.

This paper uses a Bayesian VAR with a time-varying mean and stochastic volatility (Banbura and van Vlodrop, 2018), in order to elaborate forecasts for the Peruvian economy. The main motivation for using this model is the fact that it is flexible enough to consider the structural changes that potentially occur in the economy. Forecasts are made mainly for variables such as inflation and GDP growth, although the model might be adapted to include other variables. The empirical model uses information from the BCRP's survey of macroeconomic expectations as observables linked to the time-varying long-term means.

Our results show a good fit and a considerable predictive power for inflation and output growth and for the last years previous to the Covid-19 pandemic. These results reaffirm the idea associated with the use of expectations surveys to reduce long-term uncertainty, while the time varying parameters improve the predictive power of the model.

Literature Review The existing forecasting models using Bayesian techniques for the Peruvian economy can be found first in Llosa *et al.* (2005) and, the main workhorse of the Central Bank of Peru, i.e. the Quarterly Projection Model (Departamento de Modelos Macroeconómicos, 2009; Winkelried, 2013). The model presented in this paper points towards being an alternative satellite model which, because of its flexibility, could be extended in the future for using different sets of variables and specifications.

Bayesian Vector Auto-regressive models with time varying parameters are nowadays widely used, both for forecasting and for policy experiments, see e.g. Cogley and Sargent (2005), Primiceri (2005) Del Negro and Primiceri (2015), Canova and Gambetti (2009), Canova and

Gambetti (2010), Banbura and van Vlodrop (2018), among others. The story would not be completed if we do not mention the Stochastic Volatility component, which is part of the Markov Chain Monte Carlo (MCMC) applied literature, starting with Jacquier *et al.* (1994), the widely applied approximation of Kim *et al.* (1998), and recent papers, e.g. Carriero *et al.* (2016) among others. There are also plenty of applications of Stochastic Volatility using particle filters, but since the presented model is based on a state-space structure the is conditionally linear and normal, we keep it simple and use the popular simulation smoother based on kalman filter techniques, i.e. see Carter and Kohn (1994) and Durbin and Koopman (2002).

The document is organized as follows: section 2 describes the BVAR model, section 3 describes the estimation procedure, section 4 discusses the main results, section 5 illustrates predictive power of the model, and section 6 concludes.

2 The model

Consider the following Vector Autorregressive (VAR) model (Banbura and van Vlodrop, 2018):

$$y_t - \tau_t = \sum_{k=1}^{p} B_k \left(y_{t-k} - \tau_{t-k} \right) + \varepsilon_t, \qquad \varepsilon_t \sim i.i.d.N\left(0, H_t\right) \tag{1}$$

where y_t is an $M \times 1$ vector of macroeconomic variables, B_k with $k = 1, \ldots, p$ are $M \times M$ matrices of coefficients that satisfy the stability condition, $H_t = A^{-1}\Lambda_t A^{-1\prime}$ is the $M \times M$ covariance matrix, being $\Lambda_t = diag\left(\sigma^2_{H,1,t}, \ldots, \sigma^2_{H,M,t}\right)$ a $M \times M$ diagonal a positive definite matrix with the shock variances in its main diagonal, and A^{-1} a lower triangular matrix with $1s$ in its main diagonal.

The vector τ_t contains the local mean of each variable. In order to undertake a flexible forecasting exercise relative to the standard approach of Litterman (1986) and Waggoner and Zha (1999), we follow Banbura and van Vlodrop (2018) and assume that this vector is time varying according to:

$$\tau_t = \tau_{t-1} + \eta_t, \qquad \eta_t \sim i.i.d.N\left(0, V_t\right) \tag{2}$$

with $V_t = diag\left(\sigma^2_{V,1,t}, \ldots, \sigma^2_{V,M,t}\right)$ as a $M \times M$ diagonal a positive definite matrix with the shock variances in its main diagonal. Under these assumptions, we have that:

$$lim_{h\to\infty} E_t\left[y_{t+h} \mid \tau_t\right] = \tau_t \tag{3}$$

where E_t denotes the expectation conditional on the information set available at time t.

In addition to the data vector y_t we also consider survey variables z_t, where it is assumed that they are closely linked to the local mean vector τ_t as follows[1]:

$$z_t = P_\tau \tau_t + g_t, \qquad g_t \sim i.i.d.N\left(0, G_t\right) \tag{4}$$

with $G_t = diag\left(\sigma^2_{G,1,t}, \ldots, \sigma^2_{G,M,t}\right)$ as a $M \times M$ diagonal a positive definite matrix with the shock variances in its main diagonal. P_τ is a selection matrix for the variables in y_t that are associated with their corresponding expectations in z_t. Innovations g_t should not be considered as measurement errors. Instead, since z_t are *proxies* for the long term value of each variable, the mentioned innovations capture the short term deviations from these values, and these might be eventually large. Nevertheless, the inclusion of these additional variables is crucial for disciplining the forecasting exercise, as they play the role of democratic priors.

Finally, in this setup we assume Stochastic Volatility, i.e. that variances are also time varying according to a random walk (Kim *et al.*, 1998):

$$log\left(\sigma^2_{i,m,t}\right) = log\left(\sigma^2_{i,m,t-1}\right) + e_{i,m,t}, \qquad e_{i,m,t} \sim i.i.d.N\left(0, \phi_{i,m}\right) \tag{5}$$

where $m = 1, \ldots, M$ and $i = \{H, V, G\}$. The stochastic volatility component is crucial for mitigating the potential problems of model misspecification. In particular, it is also a well known fact that the use of stochastic volatility improves the model forecasts in a significant way (see e.g. Diebold *et al.* (2017)).

[1]The dimension of vector z_t can be potentially different than vector y_t (see Banbura and van Vlodrop (2018)), but in this paper we simplify this assumption including one survey variable for each component of y_t. Last but not least, the latter specification assumes that both vectors are expressed in the same units.

The complete model can be cast as a state-space system. In the next section we describe the estimation setup, together with the data and priors, etc.

3 Bayesian Estimation

3.1 Data and estimation setup

Our baseline specification for the Bayesian Vector Autorregresive model (BVAR) takes into account four core macroeconomic variables (y_t): i) Year-on-year GDP growth (Y), ii) Year-on-year CPI Inflation (P), iii) Overnight Interbank Rate in soles and in annual terms (R), and iv) Year-on-year PEN Nominal Depreciation (E). The sample exercise considers the entire inflation targeting regime, starting in February 2002 up to December 2020[2], as it is shown in Figure 1.

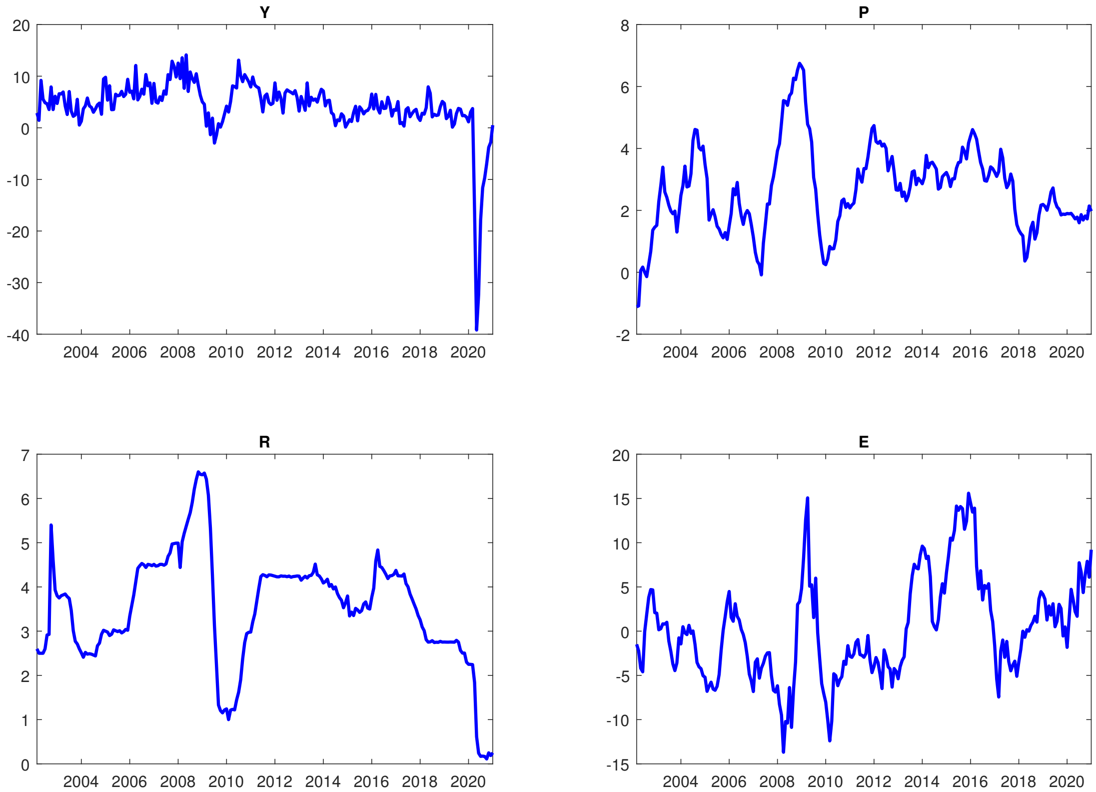


Figure 1: Peruvian Macroeconomic Data (y_t)

[2]We do some *nowcasting* estimations for GDP growth at the end of the sample using the model proposed by Pérez Forero (2018).

In addition, the presented model considers Long-term expectations for the four macroeconomic variables (z_t), which will be useful for estimating the time varying means (τ_t). We employ the BCRP's monthly Macroeconomic Expectations Survey[3] for the cases of GDP growth, Inflation and PEN Depreciation. In all cases we take the furthest expectation on the horizon, which is typically between 18 and 24 months. Fore the case of the expectations associated with the interbank rate, we use the 18-month interest rate of the Central Bank Securities in soles (CDBCRP). What is also valuable for the case of the state space models, is the fact that we can consider data with missing values. As a result, we do not need a balanced panel to undertake the estimation. In general, we might consider a smaller set of expectations variables z_t relative to the number of variables included in the BVAR model (y_t).

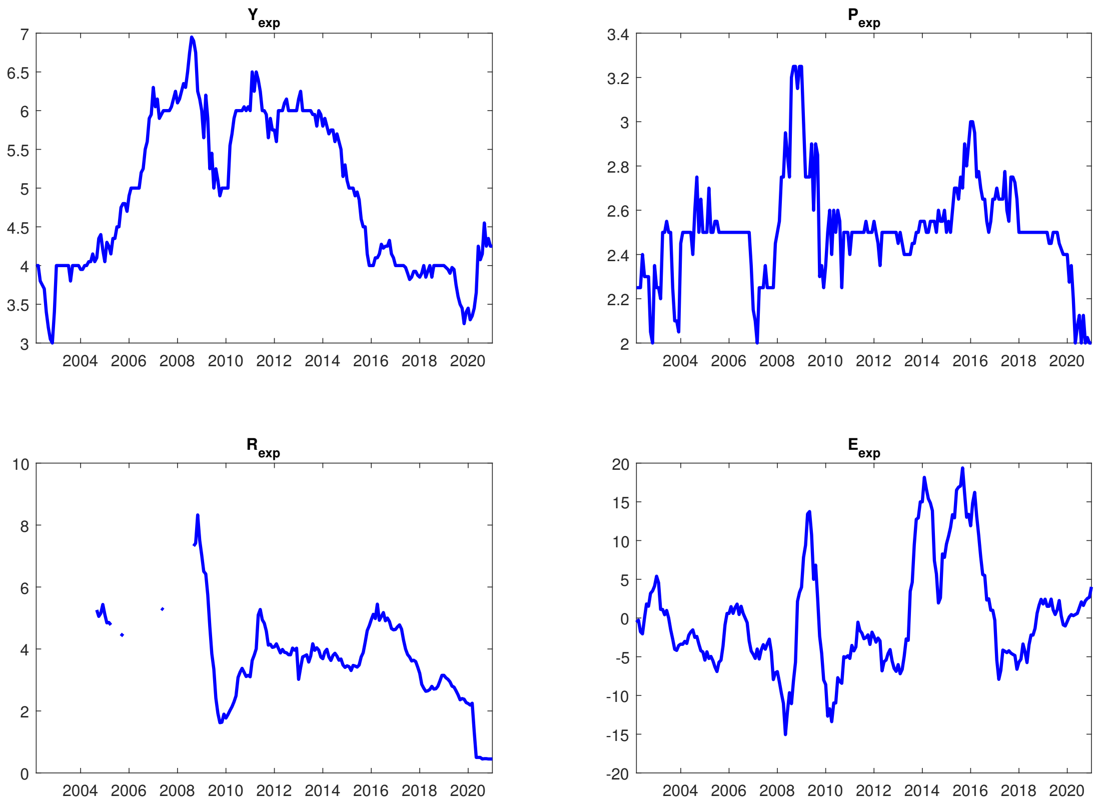


Figure 2: Long Term Expectations (z_t)

[3]See https://www.bcrp.gob.pe/en/estadisticas/encuesta-de-expectativas-macroeconomicas.html

3.2 Prior specification

Consider the complete parameter set of the model $\Theta = \left\{\Lambda^T, V^T, G^T, \phi_H, \phi_V, \phi_G, B, A, \tau^T\right\}$, where the superscript T denotes the full time series of the parameter block. Moreover, B represents the BVAR matrix coefficients, A is the lower triangular matrix with ones in the main diagonal and the covariances as free parameters.

For the VAR coefficients $\beta = vec(B)$ we take an independent normal prior, i.e. a conjugated prior:

$$p(\beta) = N(\mu_B, \lambda_0 \Omega_B) \tag{6}$$

with μ_B as the common mean and λ_0 as the overall tightness parameter. The covariance matrix Ω_B takes the form of the typical Minnesota prior (Litterman, 1986), i.e. $\Omega_B = diag(\omega_{ij,l})$ such that

$$\omega_{ij,l} = \begin{cases} \frac{1}{l^{\lambda_3}} & , i = j \\ \frac{\lambda_1}{l^{\lambda_3}} \left(\frac{\widehat{\sigma}_j^2}{\widehat{\sigma}_i^2}\right) & , i \neq j \\ \lambda_2 & , exogenous \end{cases} \tag{7}$$

where

$$i, j \in \{1, \ldots, M\} \text{ and } l = 1, \ldots, p$$

and $\widehat{\sigma}_j^2$ is the variance of the residuals from an estimated $AR(p)$ model for each variable $j \in \{1, \ldots, M\}$.

In this setup, since the variables included in the model are transformed to be stationary, we set $\mu_B = 0_{dim(\beta)}$. In addition, we set the parameters $\lambda_0 = 0.2$, $\lambda_1 = 0.5$, $\lambda_2 = 1$, $\lambda_3 = 2$. We take the benchmark values of Doan *et al.* (1984) (see also Canova (2007)) except the one for the exogenous component λ_2, which is typically set to $10,000$. We do not consider this value since this was set for a constant intercept, and the context is completely different is this model. In addition, the presented parameter configuration is traditional for US data, and part of the future agenda is to properly estimate these parameters for the Peruvian case[4].

[4] Giannone *et al.* (2015) discuss the estimation of some of these hyper-parameters, and this could a natural extension of this setup.

The prior distributions for the covariances parameters included in matrix A could be specified row by row. That is, since the VAR model is recursive, we can estimate parameters equation by equation assuming that they are independent[5]. Consider $\alpha_i, i = 2, \ldots, M$, where α_i is a column vector with the free parameters of the i^{th} row in A, i.e. $dim\left(\alpha_i\right) = i - 1$. The prior distribution for each column vector in this case is

$$\alpha_i \sim N\left(\mu_{\alpha,i}, \Omega_{\alpha,i}\right) \tag{8}$$

However, following Canova and Pérez Forero (2015)[6] we extract the vector of parameters as follows

$$vec\left(A\right) = S_A \alpha + s_A \tag{9}$$

Thus, it is then possible to specify a prior for the entire vector such that:

$$\alpha \sim N\left(\mu_\alpha, \Omega_\alpha\right) \tag{10}$$

where we assume $\mu_\alpha = 0_{dim(\alpha)}$ and $\Omega_\alpha = 10 \times I_{dim(\alpha)}$.

In the case of the stochastic volatility processes, we need to specify the distribution of the initial point and the prior distributions for the variance parameters ϕ governing the amount of time variation in the process as follows:

$$ln\sigma^2_{H,i,p+1} \sim N\left(0, \upsilon^2_H\right), \phi_{H,i} \sim IG\left(d_{\phi_H} \times \underline{\phi}_{H,i}, d_{\phi_H}\right), i = 1, \ldots, M \tag{11}$$

$$ln\sigma^2_{V,i,p+1} \sim N\left(0, \upsilon^2_V\right), \phi_{V,i} \sim IG\left(d_{\phi_V} \times \underline{\phi}_{V,i}, d_{\phi_V}\right), i = 1, \ldots, M \tag{12}$$

$$ln\sigma^2_{G,i,p+1} \sim N\left(0, \upsilon^2_G\right), \phi_{G,i} \sim IG\left(d_{\phi_G} \times \underline{\phi}_{G,i}, d_{\phi_G}\right), i = 1, \ldots, M \tag{13}$$

Specifically, we set $\nu_H = \nu_V = \nu_G = 100$, i.e. we treat this as a diffuse filter since the prior law of motion of volatility is a random walk. In addition, we set $d_{\phi_H = d_{\phi_V} = d_{\phi_G} = 10}$ and

[5] See Cogley and Sargent (2005) and Primiceri (2005) and Banbura and van Vlodrop (2018).

[6] See also Amisano and Giannini (1997)

$\underline{\phi}_{H,i} = \underline{\phi}_{V,i} = \underline{\phi}_{G,i} = 0.1$ for $i = 1, \ldots, M$. A similar parametrization can be found in Carriero *et al.* (2016).

Finally, for the case of the time-varying mean τ_t, we specify the prior for the initial point in a similar way:

$$\tau_{p+1} \sim N(\mu_\tau, \Omega_\tau) \tag{14}$$

In this case, since this is also a diffuse filter, we set $\mu_\tau = 0_{dim(\tau)}$ and $\Omega_\tau = 100 \times I_{dim(\tau)}$.

3.3 Gibbs Sampling

Given the priors and the likelihood function of the model, we proceed to estimate the parameter set Θ. First, we need to recall the Bayes' Theorem in order to identify the posterior distribution, so that:

$$p(\Theta \mid Y) \propto p(Y \mid \Theta)\, p(\Theta) \tag{15}$$

Given the parameter set $\Theta = \left\{\Lambda^T, V^T, G^T, \phi_H, \phi_V, \phi_G, B, A, \tau^T, s^T\right\}$, where $s^T = \left\{s_H^T, s_V^T, s_G^T\right\}$ is the set of indicators associated with the mixture of normals approximation when simulation Stochastic Volatility (Kim *et al.*, 1998), and we put this block at the end of the simulation in line with the correction proposed by Del Negro and Primiceri (2015). We denote Θ/χ as the parameter vector Θ excluding the element χ. We set $k = 1$ and consider K as the total number of draws and $Y = \left\{y^T, z^T\right\}$ as the full data set. The simulation algorithm is as follows:

1. Draw $p\left(\Lambda_i^T \mid \Theta/\Lambda_i^T, y^T, z^T\right)$, $i = 1, \ldots, M$: Stochastic Volatility (Kim *et al.*, 1998)
2. Draw $p\left(V_i^T \mid \Theta/V_i^T, y^T, z^T\right)$: $i = 1, \ldots, M$: Stochastic Volatility (Kim *et al.*, 1998)
3. Draw $p\left(G_i^T \mid \Theta/G_i^T, y^T, z^T\right)$: $i = 1, \ldots, M_Z$: Stochastic Volatility (Kim *et al.*, 1998)
4. Draw $p\left(\phi_{H,i} \mid \Theta/\phi_{H,i}, y^T, z^T\right)$: Inverse-Gamma simulation, $i = 1, \ldots, M$
5. Draw $p\left(\phi_{V,i} \mid \Theta/\phi_{V,i}, y^T, z^T\right)$: Inverse-Gamma simulation, $i = 1, \ldots, M$
6. Draw $p\left(\phi_{G,i} \mid \Theta/\phi_{G,i}, y^T, z^T\right)$: Inverse-Gamma simulation, $i = 1, \ldots, M$

7. Draw $p\left(A \mid \Theta/A, y^T, z^T\right)$: Linear Regression, Normal Distribution

8. Draw $p\left(B \mid \Theta/B, y^T, z^T\right)$: Linear Regression, Normal Distribution

9. Draw $p\left(\tau^T \mid \Theta/\tau^T, y^T, z^T\right)$: State-Space simulation (Carter and Kohn, 1994)

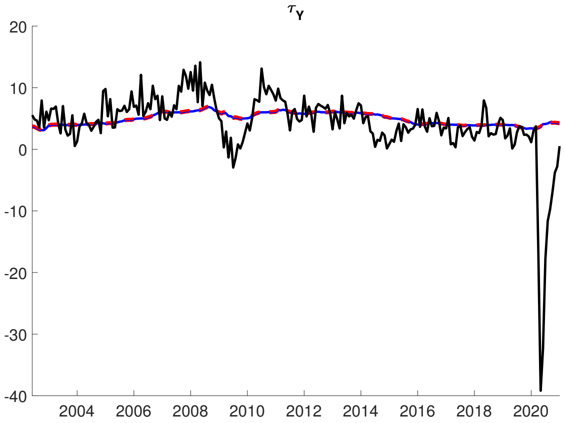
10. Draw $p\left(s^T \mid \Theta/s^T, y^T, z^T\right)$: Discrete distribution (Kim *et al.*, 1998)

11. If $k < K$, set $k = k + 1$ and back to step 1.

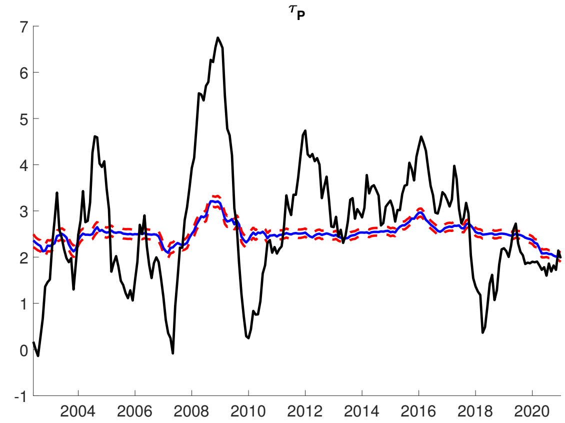
See details for each block in Appendix A. We run the Gibbs sampler for $K = 50,000$ and discard the first $25,000$ draws in order to minimize the effect of initial values. In order to reduce the serial correlation across draws, we set a thinning factor of 10. As a result, we have $2,500$ draws for conducting inference. After simulation and convergence, we are ready to present the results of the empirical exercise in the next section.

4 Results

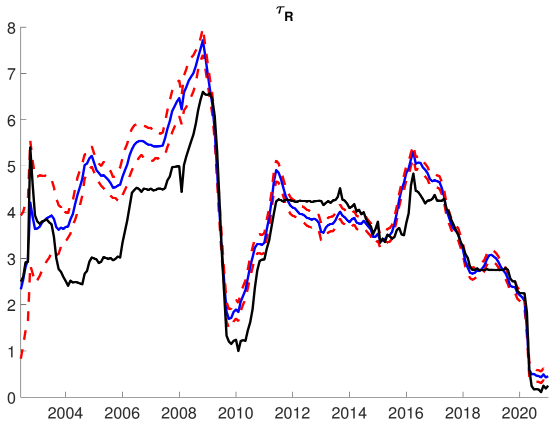
The estimated time-varying means are depicted in Figure 3. The estimated time-varying means are of course less volatile than the observed macroeconomic data, and the credible set of these latent variables exhibit a good symptom of precision, although the priors described in subsection 3.2 are standard and not very restrictive.



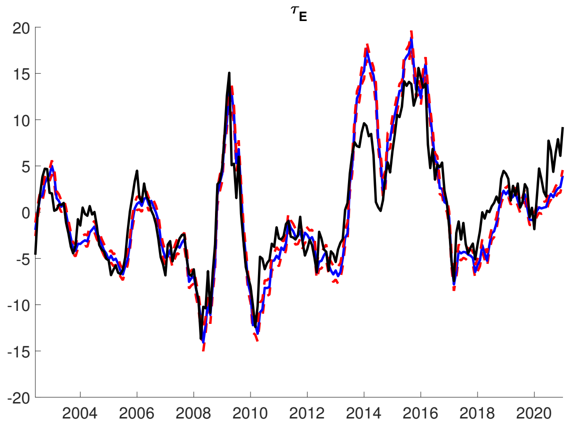
(a) Year-on-year GDP



(b) Year-on-year CPI Inflation



(c) Interbank Rate



(d) PEN Nominal Depreciation

Figure 3: Estimated Time varying means τ_t

Given the stability assumption for the BVAR model, at each point in time these values could be considered as a long term mean, i.e. the reference value for the long term forecasts. In other words, if we consider a long horizon, forecasts distribution should converge to a value close to τ_t. One of the main limitations of the constant coefficients BVAR forecasting model, although it is disciplined through the use of shrinkage priors (Litterman, 1986), is the fact that it is forced to converge to the sample mean[7]. On the other extreme, full time-varying parameters BVAR models might be limited by the fact that they are computationally costly (because of

[7] In this context, in order to improve the forecasting performance, the naive solution is to consider rolling window, or to select the "appropriate" sub-sample. None of these alternatives sound particularly robust to changes in data, if you are willing to use your model every month or quarter.

the use intensive use of the kalman filter), and they could crash easily if we impose stationarity. Therefore, they are not suitable for including a large set of variables (even 6 or more). As a consequence, the estimated model is in the middle of these two extremes, and because of its flexibility it could be easily extended without worrying about the mentioned issues.

Given the estimated parameters (see appendix B for the rest of the blocks in Θ), in the next section we present a simple evaluation of the forecasting power of the model.

5 Predictive Power

In order to test the predictive power of the presented model, we perform a forecasting exercise for a medium term horizon, i.e. two years or $h = 24$. It is important to remark that the two-year horizon is the one that is relevant for the BCRP in terms of its monetary policy design. In addition, the BCRP survey of expectations also has the same maximum horizon.

We proceed as follows: First, we need to recall the state space structure of the model, in particular equation (1)

$$y_t - \tau_t = \sum_{k=1}^{p} B_k \left(y_{t-k} - \tau_{t-k}\right) + \varepsilon_t, \qquad \varepsilon_t \sim i.i.d.N\left(0, H_t\right)$$

together with the transition equation (2)

$$\tau_t = \tau_{t-1} + \eta_t, \qquad \eta_t \sim i.i.d.N\left(0, V_t\right)$$

and the volatilities in equation (5).

$$log\left(\sigma^2_{i,m,t}\right) = log\left(\sigma^2_{i,m,t-1}\right) + e_{i,m,t}, \qquad e_{i,m,t} \sim i.i.d.N\left(0, \phi_{i,m}\right)$$

where $m = 1, \ldots, M$ and $i = \{H, V, G\}$. See the complete state-space form in Appendix A. The forecasting exercise starts at point $t = T$. Then, repeat the followings steps for $t = T + 1, \ldots, T + h$ for a given draw of Θ:

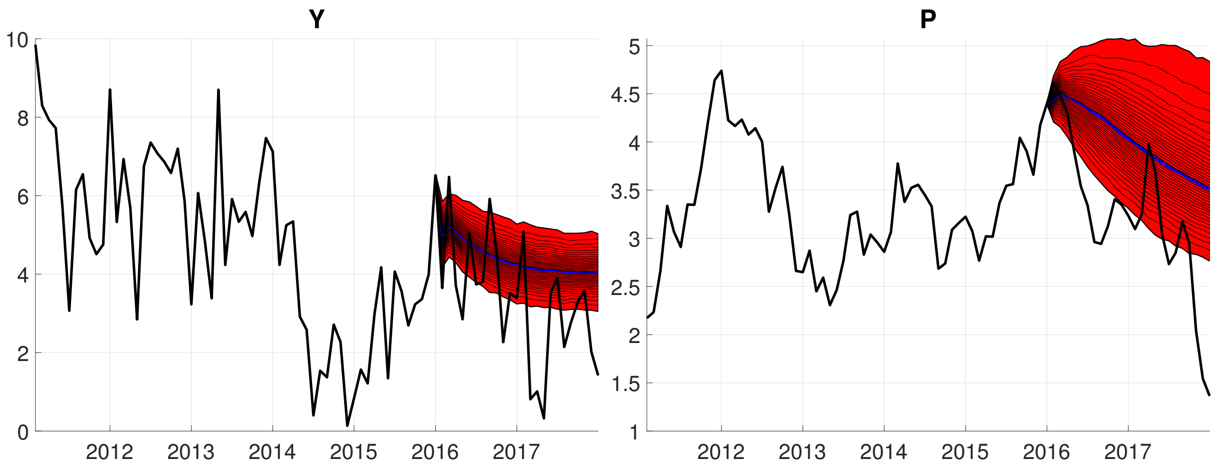
1. Draw $e_{i,m,t} \sim i.i.d.N\left(0, \phi_{i,m}\right)$ for $m = 1, \ldots, M$ and $i = \{H, V, G\}$.

2. Forecast $log\left(\sigma^2_{i,m,t+1}\right)$ using equation (5) for $m = 1, \ldots, M$ and $i = \{H, V, G\}$.

3. Draw $\varepsilon_t \sim i.i.d.N\left(0, H_t\right)$ and $\eta_t \sim i.i.d.N\left(0, V_t\right)$

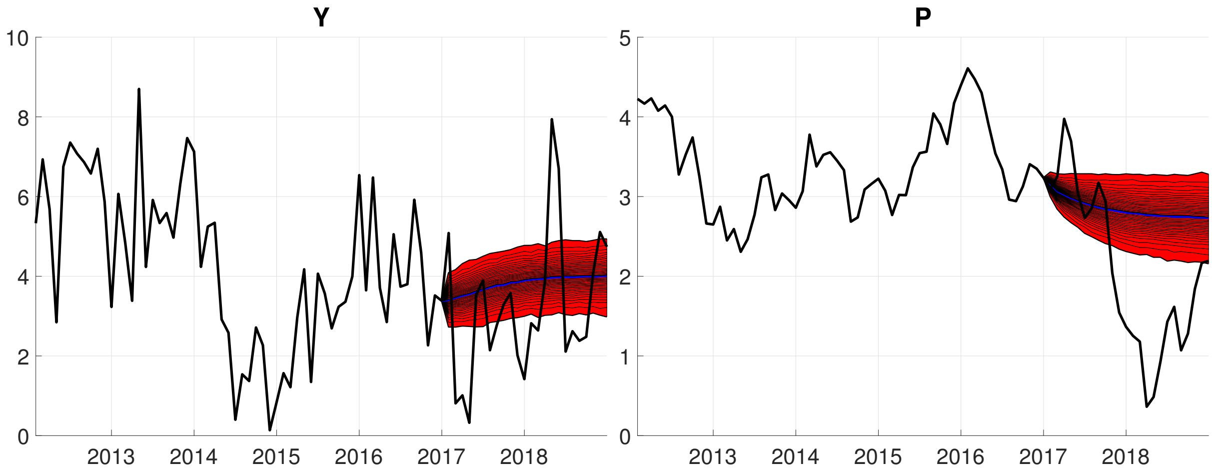
4. Forecast τ_{t+1} and y_{t+1} using the system (39)-(40).

In addition, order assess the uncertainty in an accurate way, we repeat the algorithm $l = 1, \ldots, L$ times and take averages in each point in time. As a result, we have the fancharts for selected dates depicted in figures 4, 5, 6, 7 and 8 together with the ex-post observed values for comparison purposes.



(a) Year-on-year GDP fanchart (b) Year-on-year CPI Inflation fanchart

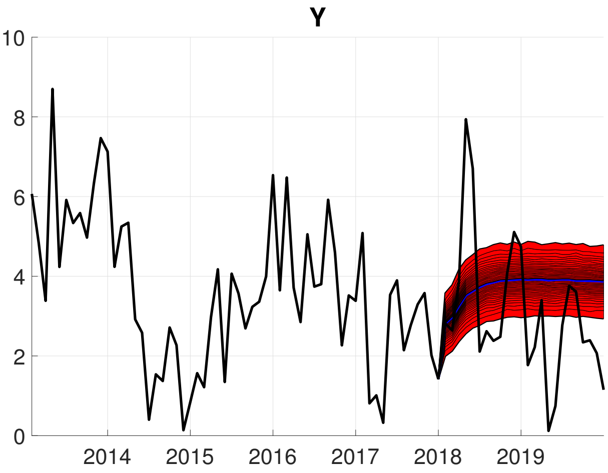
Figure 4: Forecast starting in Dec-2015



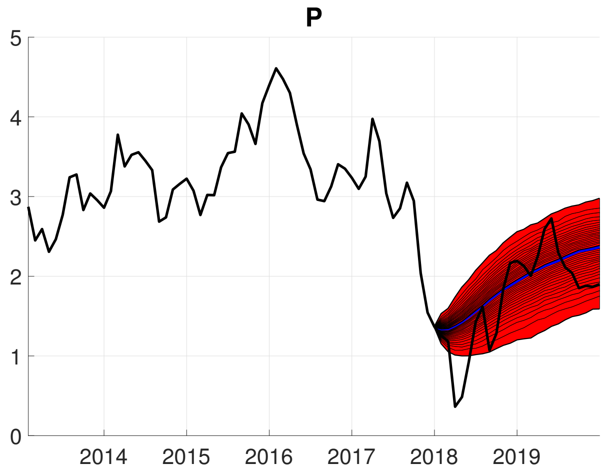
(a) Year-on-year GDP fanchart (b) Year-on-year CPI Inflation fanchart

Figure 5: Forecast starting in Dec-2016

The forecasting model performs particularly well for the case of inflation, and the for GDP growth. It is important to clarify that this is not a nowcasting or a short term horizon forecast model. Instead, given the fact that we are using long-term expectations, and that we have estimated the time-varying long term means, we should consider this model at least as a medium run one. The presented fancharts consider a horizon of two years ($h = 24$) using monthly data. Although the horizon is quite large, density forecasts look quite stable and converge to the local mean τ_t.



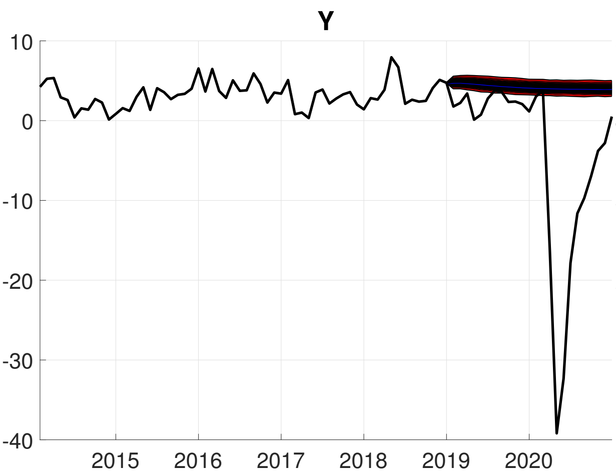
(a) Year-on-year GDP fanchart



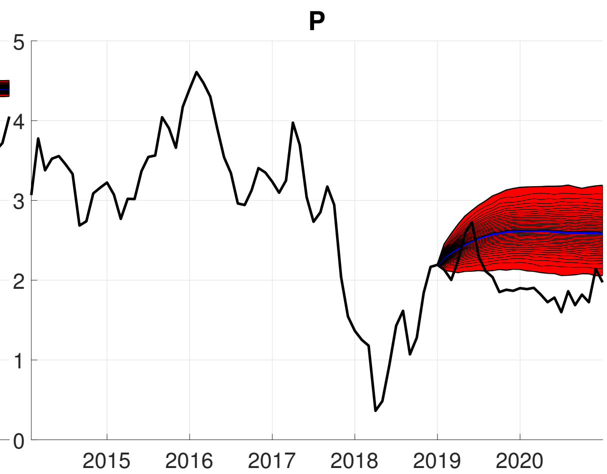
(b) Year-on-year CPI Inflation fanchart

Figure 6: Forecast starting in Dec-2017

Finally, given the extremely large shock of the Covid-19 pandemic, and the given the GDP growth and inflation expectations at the end of 2019, it was simply impossible to forecast economic activity conditional on the information set available at that point in time. Nevertheless, since inflation and its expectations remained stable, the forecast at the end of the year was very close the center of the target range (2%).

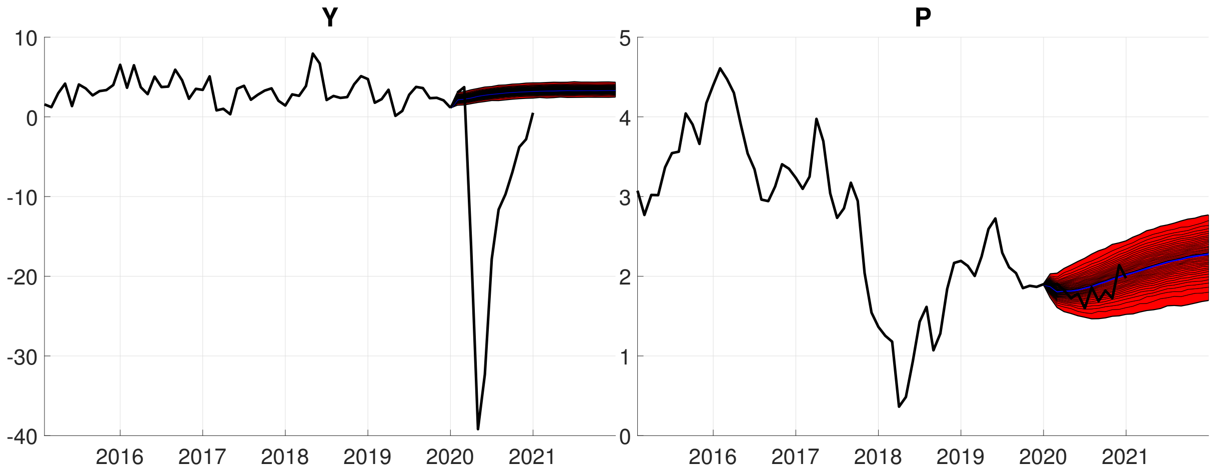


(a) Year-on-year GDP fanchart



(b) Year-on-year CPI Inflation fanchart

Figure 7: Forecast starting in Dec-2018



(a) Year-on-year GDP fanchart (b) Year-on-year CPI Inflation fanchart

Figure 8: Forecast starting in Dec-2019

6 Concluding Remarks

This paper uses a Bayesian VAR with a time-varying mean and stochastic volatility (Banbura and van Vlodrop, 2018), in order to elaborate forecasts for the Peruvian economy. Density Forecasts are made mainly for variables such as inflation and GDP growth, and they provide a useful approach to quantify the uncertainty regarding both parameters and model specification. These outputs are particularly relevant for Central Banks and other policy makers. Moreover, the empirical model uses information from the BCRP's survey of macroeconomic expectations as observables linked to the time-varying long-term means.

Our density forecast results show a good fit and a considerable predictive power for inflation and output growth and for the last years, and they reaffirm the idea associated with the use of expectations surveys to reduce long-term uncertainty, while the time varying parameters and stochastic volatility improve the predictive power of the model, since they are a safe strategy against model misspecification.

The future agenda considers the adaptation of the model to include other relevant variables for policy makers, such as money aggregates, different financial variables with an informative component, and also using different sources of Surveys of Expectations. We expect to continue expanding the set of satellite models in order to boost our forecasting capacity.

A Gibbs Sampling details

1. **Block 1**: Sampling from $p\left(\Lambda_i^T \mid \Theta/\Lambda_i^T, y^T, z^T\right)$, $i = 1, \ldots, M$

   Given the parameter values of the model B and τ_t, compute the innovations ε_t such that

   $$\varepsilon_t = (y_t - \tau_t) - \sum_{k=1}^{p} B_k \left(y_{t-k} - \tau_{t-k}\right), t = p+1, \ldots, T \tag{16}$$

   Recall that $\varepsilon_t \sim i.i.d.N\left(0, H_t\right)$, with $H_t = A^{-1}\Lambda_t A^{-1\prime}$ and $\Lambda_t = diag\left(\sigma^2_{H,1,t}, \ldots, \sigma^2_{H,M,t}\right)$. Therefore, the standardized innovations are $\tilde{\varepsilon}_t = A\varepsilon_t$. In order to sample volatilities Λ_t we proceed for each $i = 1, \ldots, M$ as follows:

   $$ln\left(\tilde{\varepsilon}^2_{i,t} + 0.001\right) = ln\left(\sigma^2_{H,i,t}\right) + u_{H,i,t} \tag{17}$$

   where $u_{H,i,t} \sim log\left(\chi^2\right)$ is approximated through a mixture of 7 normal distributions:

   $$f\left(u_{H,i,t}\right) \approx \sum_{j=1}^{7} q_j f_N\left(u_{H,i,t} \mid m_j - 1.2704, \upsilon_j^2\right) \tag{18}$$

   In addition, we have the transition equation

   $$ln\left(\sigma^2_{H,i,t}\right) = ln\left(\sigma^2_{H,i,t-1}\right) + e_{H,i,t}, \qquad e_{H,i,t} \sim i.i.d.N\left(0, \phi_{H,i}\right) \tag{19}$$

   As a result, equations (17) – (19) form a state-space system, so that we can simulate $log\left(\sigma^2_{H,i}\right)^T$ following Kim *et al.* (1998), i.e. using the algorithm of Carter and Kohn (1994) conditional on the discrete variable $s^T_{H,i}$ and given the prior (11).

2. **Block 2**: Sampling from $p\left(V_i^T \mid \Theta/V_i^T, y^T, z^T\right)$: $i = 1, \ldots, M$

   Given the parameter values of the model τ_t, compute the innovations η_t such that

   $$\eta_t = \tau_t - \tau_{t-1} \tag{20}$$

Recall that $\eta_t \sim i.i.d.N\left(0, V_t\right)$, with $V_t = diag\left(\sigma^2_{V,1,t}, \ldots, \sigma^2_{V,M,t}\right)$. In order to sample volatilities V_t we proceed for each $i = 1, \ldots, M$ as follows:

$$ln\left(\eta^2_{i,t} + 0.001\right) = ln\left(\sigma^2_{V,i,t}\right) + u_{V,i,t} \tag{21}$$

where $u_{V,i,t} \sim log\left(\chi^2\right)$ is approximated through a mixture of 7 normal distributions:

$$f\left(u_{V,i,t}\right) \approx \sum_{j=1}^{7} q_j f_N\left(u_{V,i,t} \mid m_j - 1.2704, \upsilon^2_j\right) \tag{22}$$

In addition, we have the transition equation

$$ln\left(\sigma^2_{V,i,t}\right) = ln\left(\sigma^2_{V,i,t-1}\right) + e_{V,i,t}, \qquad e_{V,i,t} \sim i.i.d.N\left(0, \phi_{V,i}\right) \tag{23}$$

As a result, equations (21) − (23) form a state-space system, so that we can simulate $log\left(\sigma^2_{V,i}\right)^T$ following Kim *et al.* (1998), i.e. using the algorithm of Carter and Kohn (1994) conditional on the discrete variable $s^T_{V,i}$ and given the prior (12).

3. **Block 3**: Sampling from $p\left(G^T_i \mid \Theta/G^T_i, y^T, z^T\right)$: $i = 1, \ldots, M_Z$

Given the parameter values of the model τ_t, compute the innovations g_t such that

$$g_t = z_t - P_Z \tau_t \tag{24}$$

Recall that $g_t \sim i.i.d.N\left(0, G_t\right)$, with $G_t = diag\left(\sigma^2_{G,1,t}, \ldots, \sigma^2_{G,M,t}\right)$. In order to sample volatilities G_t we proceed for each $i = 1, \ldots, M_Z$ as follows:

$$ln\left(g^2_{i,t} + 0.001\right) = ln\left(\sigma^2_{G,i,t}\right) + u_{G,i,t} \tag{25}$$

where $u_{G,i,t} \sim log\left(\chi^2\right)$ is approximated through a mixture of 7 normal distributions:

$$f\left(u_{G,i,t}\right) \approx \sum_{j=1}^{7} q_j f_N\left(u_{G,i,t} \mid m_j - 1.2704, \upsilon^2_j\right) \tag{26}$$

In addition, we have the transition equation

$$ln\left(\sigma^2_{G,i,t}\right) = ln\left(\sigma^2_{G,i,t-1}\right) + e_{G,i,t}, \qquad e_{G,i,t} \sim i.i.d.N\left(0, \phi_{G,i}\right) \tag{27}$$

As a result, equations (25) – (27) form a state-space system, so that we can simulate $log\left(\sigma^2_{G,i}\right)^T$ following Kim *et al.* (1998), i.e. using the algorithm of Carter and Kohn (1994) conditional on the discrete variable $s^T_{G,i}$ and given the prior (13).

4. **Block 4**: Sampling from $p\left(\phi_{H,i} \mid \Theta/\phi_{H,i}, y^T, z^T\right)$: $i = 1, \ldots, M$

   Variance parameters are simulated using an Inverse-Gamma distribution. Given the prior $\phi_{H,i} \sim IG\left(d_{\phi_H} \times \underline{\phi}_{H,i}, d_{\phi_H}\right)$, the posterior distribution is:

$$p\left(\phi_{H,i} \mid \Theta/\phi_{H,i}, y^T, z^T\right) = IG\left(d_{\phi_H} \times \underline{\phi}_{H,i} + \sum_{t=p+2}^{T} u^2_{H,i,t}, d_{\phi_H} + T - p - 1\right) \tag{28}$$

5. **Block 5**: Sampling from $p\left(\phi_{V,i} \mid \Theta/\phi_{V,i}, y^T, z^T\right)$: Inverse-Gamma simulation, $i = 1, \ldots, M$

   Variance parameters are simulated using an Inverse-Gamma distribution. Given the prior $\phi_{V,i} \sim IG\left(d_{\phi_V} \times \underline{\phi}_{V,i}, d_{\phi_V}\right)$, the posterior distribution is:

$$p\left(\phi_{V,i} \mid \Theta/\phi_{V,i}, y^T, z^T\right) = IG\left(d_{\phi_V} \times \underline{\phi}_{V,i} + \sum_{t=p+3}^{T} u^2_{V,i,t}, d_{\phi_V} + T - p - 2\right) \tag{29}$$

6. **Block 6**: Draw $p\left(\phi_{G,i} \mid \Theta/\phi_{G,i}, y^T, z^T\right)$: Inverse-Gamma simulation, $i = 1, \ldots, M_Z$ Variance parameters are simulated using an Inverse-Gamma distribution. Given the prior $\phi_{G,i} \sim IG\left(d_{\phi_G} \times \underline{\phi}_{G,i}, d_{\phi_G}\right)$, the posterior distribution is:

$$p\left(\phi_{G,i} \mid \Theta/\phi_{G,i}, y^T, z^T\right) = IG\left(d_{\phi_G} \times \underline{\phi}_{G,i} + \sum_{t=t_{z,i}+1}^{T} u^2_{G,i,t}, d_{\phi_V} + T - t_{z,i}\right) \tag{30}$$

7. **Block 7**: Sampling from $p\left(A \mid \Theta/A, y^T, z^T\right)$: Linear Regression, Normal Distribution Consider the BVAR residual terms $\varepsilon_t \sim i.i.d.N\left(0, H_t\right)$. Recall that $H_t = A^{-1}\Lambda_t A^{-1\prime}$

and $\Lambda_t = diag\left(\sigma^2_{H,1,t}, \ldots, \sigma^2_{H,M,t}\right)$. Therefore, the standardized innovations are $A\varepsilon_t = \tilde{\varepsilon}_t$. Recall also that $vec\,(A) = S_A\alpha + s_A$, so that (Amisano and Giannini, 1997; Canova and Pérez Forero, 2015):

$$vec\,(A\varepsilon_t) = \left(\varepsilon_t' \otimes I\right)(S_A\alpha + s_A) \tag{31}$$

As a consequence we can define $\tilde{e}_t = (\varepsilon_t' \otimes I)\, s_A$ and $\tilde{x}_t = -\,(\varepsilon_t' \otimes I)\, S_A$ such that we have the following linear-normal regression model:

$$\tilde{e}_t = \tilde{x}_t\alpha + \tilde{\varepsilon}_t \tag{32}$$

Given the prior $\alpha \sim N\,(\mu_\alpha, \Omega_\alpha)$ we sample the posterior

$$p\left(\alpha \mid \Theta/\alpha, y^T\right) = N\left(\overline{\alpha}, \overline{V}_\alpha\right) \tag{33}$$

with

$$\overline{V}_\alpha = \left(\Omega_\alpha^{-1} + \sum_{t=p+1}^{T} \tilde{x}_t'\Lambda_t^{-1}\tilde{x}_t\right)^{-1} \tag{34}$$

$$\overline{\alpha} = \overline{V}_\alpha\left(\Omega_\alpha^{-1}\mu_\alpha + \sum_{t=p+1}^{T} \tilde{x}_t'\Lambda_t^{-1}\tilde{e}_t\right) \tag{35}$$

8. **Block 8**: Sampling from $p\left(B \mid \Theta/B, y^T, z^T\right)$: Linear Regression, Normal Distribution Given the BVAR model in (1), let $\beta = vec\,(B)$ and given the prior (6), then the posterior distribution of β is Normal according to(Koop and Korobilis, 2010):

$$p\left(\beta \mid \Theta/\beta, y^T\right) = N\left(\overline{\beta}, \overline{V}\right) \tag{36}$$

with

$$\overline{V} = \left(\lambda_0^{-1}\Omega_B^{-1} + \sum_{t=p+1}^{T} X_t'H_t^{-1}X_t\right)^{-1} \tag{37}$$

$$\overline{\beta} = \overline{V}\left(\lambda_0^{-1}\Omega_B^{-1}\mu_B + \sum_{t=p+1}^{T} X_t'H_t^{-1}\,(y_t - \tau_t)\right) \tag{38}$$

where $X_t = x_t' \otimes I_M$ and $x_t = \left[(y_{t-1} - \tau_{t-1})', \ldots, (y_{t-p} - \tau_{t-p})'\right]'$.

9. **Block 9**: Sampling from $p\left(\tau^T \mid \Theta/\tau^T, y^T, z^T\right)$: State-Space simulation (Carter and Kohn, 1994)

   We cast the BVAR model and observables in a linear state space model, such that

$$\begin{bmatrix} z_t \\ y_t \end{bmatrix} = \begin{bmatrix} P_\tau & 0 & \ldots & 0 \\ I & I & \ldots & 0 \end{bmatrix} \begin{bmatrix} \tau_t \\ y_t - \tau_t \\ \ldots \\ y_{t-p+1} - \tau_{t-p+1} \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} g_t \tag{39}$$

$$\begin{bmatrix} \tau_t \\ y_t - \tau_t \\ \ldots \\ y_{t-p+1} - \tau_{t-p+1} \end{bmatrix} = \begin{bmatrix} I & 0 & \ldots & 0 & 0 \\ 0 & B_1 & \ldots & B_{p-1} & B_p \\ 0 & I & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & I & 0 \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ y_{t-1} - \tau_{t-1} \\ \ldots \\ y_{t-p} - \tau_{t-p} \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & I \\ 0 & 0 \\ \ldots & \ldots \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix} \tag{40}$$

   where all the innovations are orthogonal and normally distributed with time varying variances, i.e. $g_t \sim i.i.d.N\left(0, G_t\right)$, $\eta_t \sim i.i.d.N\left(0, V_t\right)$ and $\varepsilon_t \sim i.i.d.N\left(0, H_t\right)$. Given the prior in (14) and the fact that the system is linear and normal, we sample the posterior distribution of τ^T using the Kalman Filter and following Carter and Kohn (1994).

10. **Block 10**: Sampling from $p\left(s^T \mid \Theta/s^T, y^T, z^T\right)$: Discrete distribution (Kim *et al.*, 1998)

    Conditional on the values of the other parameters, the terms $\tilde{\varepsilon}_{i,t}^2$, $\eta_{i,t}^2$ and $g_{i,t}^2$ are observable, so that we can sample the states, i.e. the elements of $s^T = \left\{s_H^T, s_V^T, s_G^T\right\}$ independently from the following discrete distributions:

$$p\left(s_{H,i,t} = j \mid \tilde{\varepsilon}_{i,t}^2, \sigma_{H,i,t}^2\right) \propto f_N\left(ln\left(\tilde{\varepsilon}_{i,t}^2 + 0.001\right) \mid ln\sigma_{H,i,t}^2 + m_j - 1.2704, \upsilon_j^2\right) \tag{41}$$

$$p\left(s_{V,i,t}=j\mid\eta_{i,t}^{2},\sigma_{V,i,t}^{2}\right)\propto f_{N}\left(ln\left(\eta_{i,t}^{2}+0.001\right)\mid ln\sigma_{V,i,t}^{2}+m_{j}-1.2704,\upsilon_{j}^{2}\right)\tag{42}$$

$$p\left(s_{G,i,t}=j\mid g_{i,t}^{2},\sigma_{H,i,t}^{2}\right)\propto f_{N}\left(ln\left(g_{i,t}^{2}+0.001\right)\mid ln\sigma_{G,i,t}^{2}+m_{j}-1.2704,\upsilon_{j}^{2}\right)\tag{43}$$

where $j\in\{1,\ldots,7\}$ is the index for the state of the mixture of normals, $f_N\left(x\mid\mu,\sigma^2\right)$ is referred to the normal density function with mean μ, variance σ^2 and evaluated at point x. The values for means, variances and weigths for the mixture of normals can be found in the following table.

s	$P\left(s=j\right)$	m_j	υ_j^2
1	0.00730	−10.12999	5.79596
2	0.10556	−3.97281	2.61369
3	0.00002	−8.56686	5.17950
4	0.04395	2.77786	0.16735
5	0.34001	0.61942	0.64009
6	0.24566	1.79518	0.34023
7	0.25750	−1.08819	1.26261

Table 1: $log\chi^2$ Distribution Approximation (Kim *et al.*, 1998)

B The posterior distribution of hyper-parameters

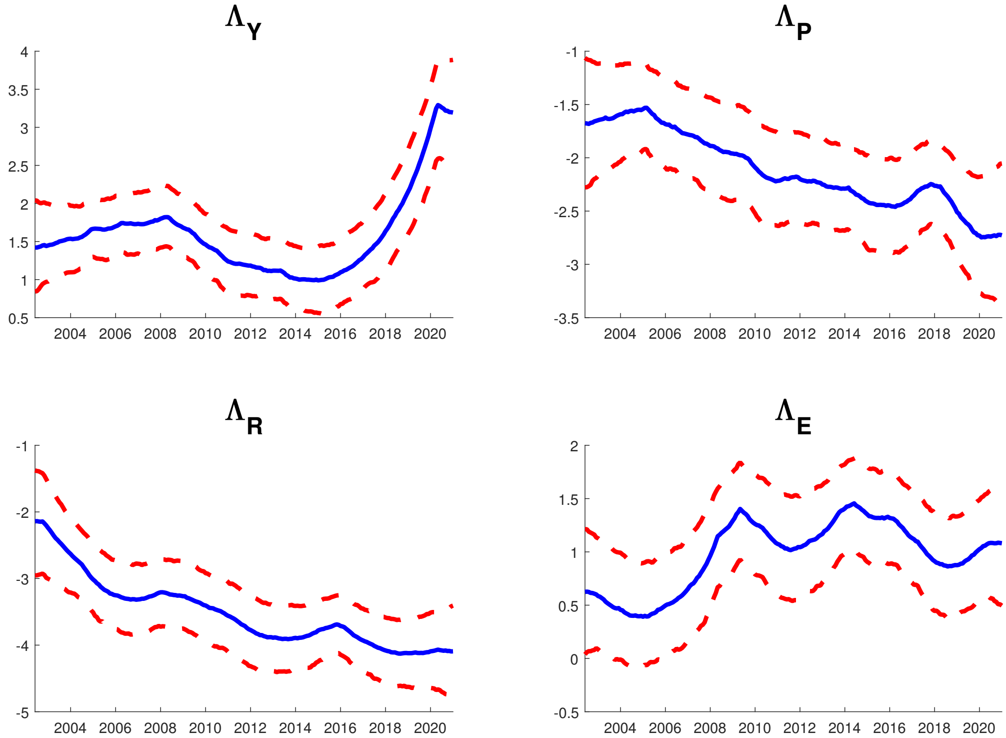


Figure 9: Log-Volatility Λ

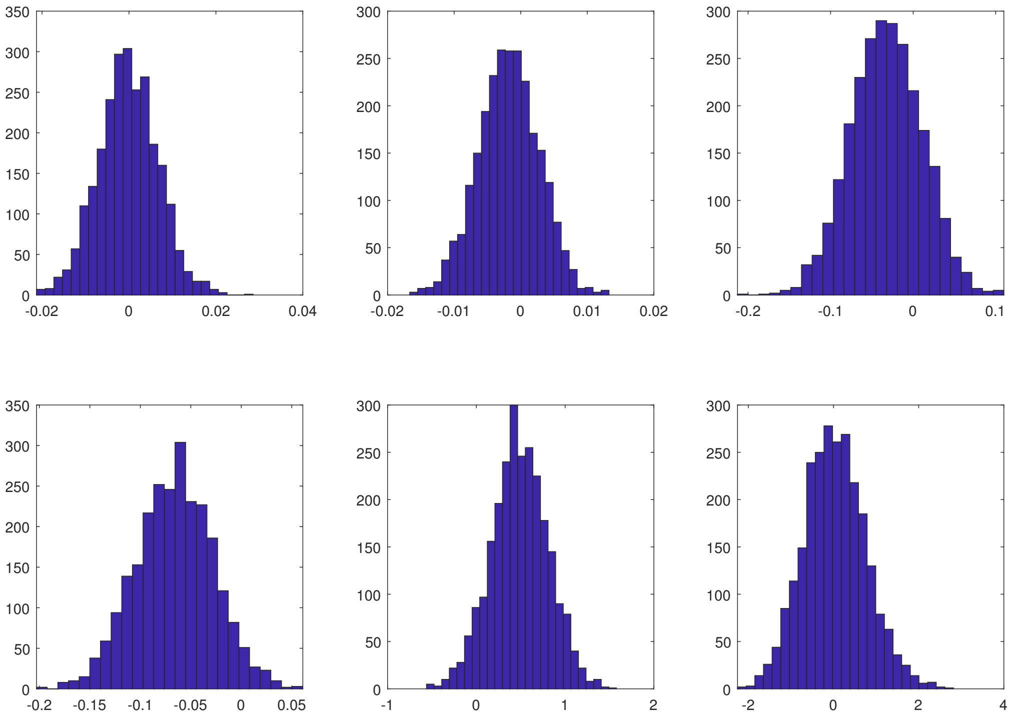


Figure 10: Covariance Parameters α

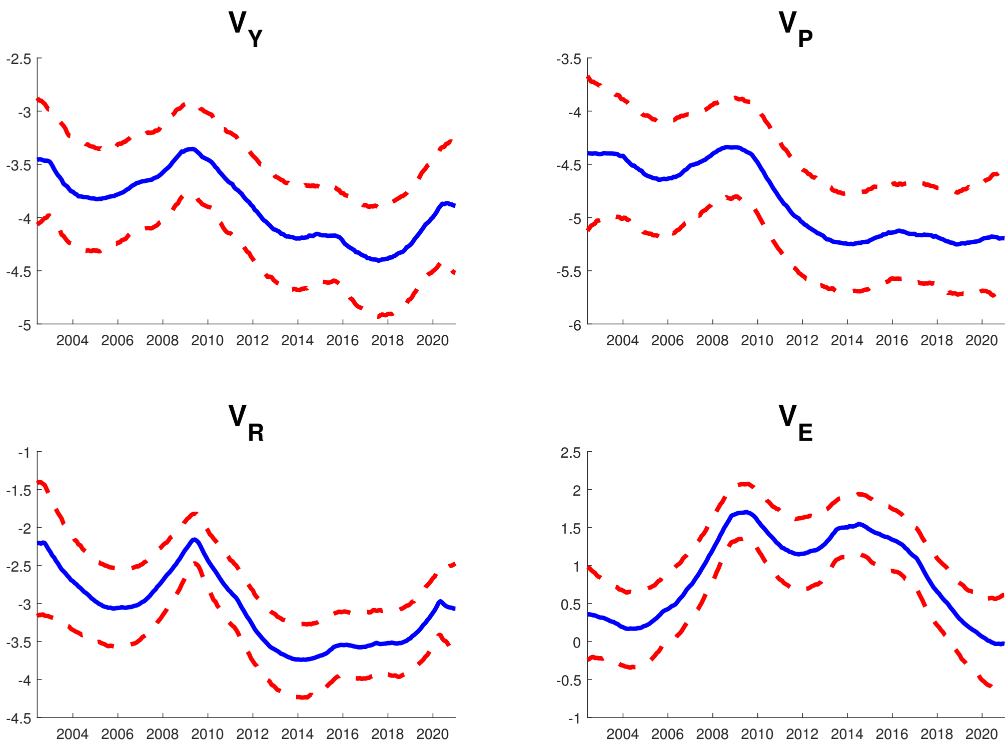


Figure 11: Log-Volatility V

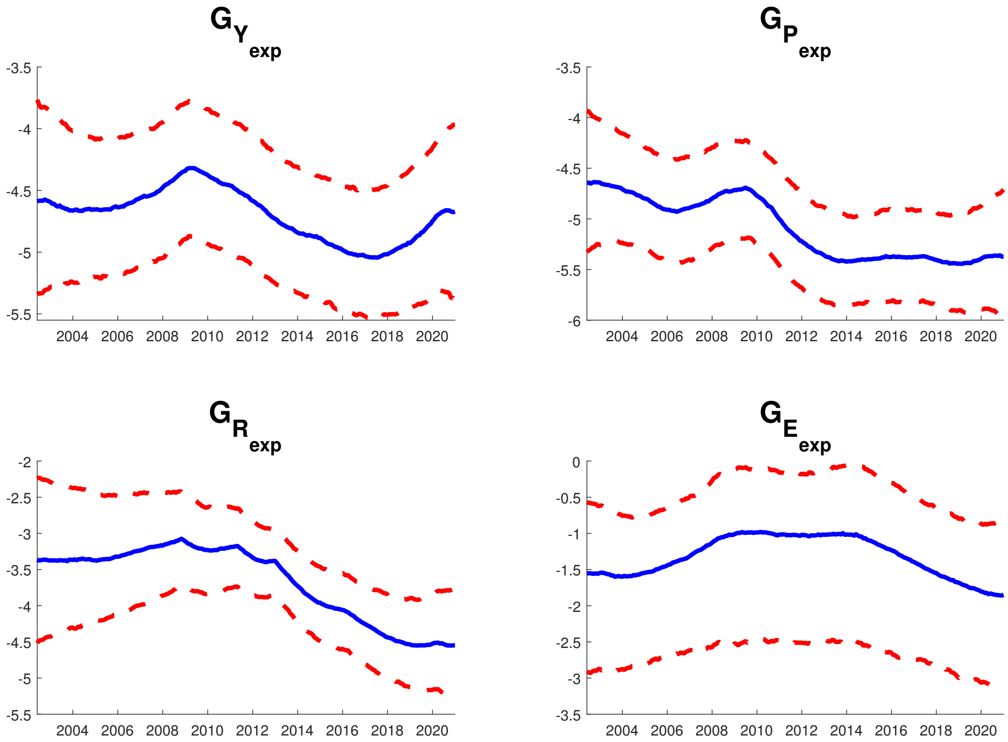


Figure 12: Log-Volatility G

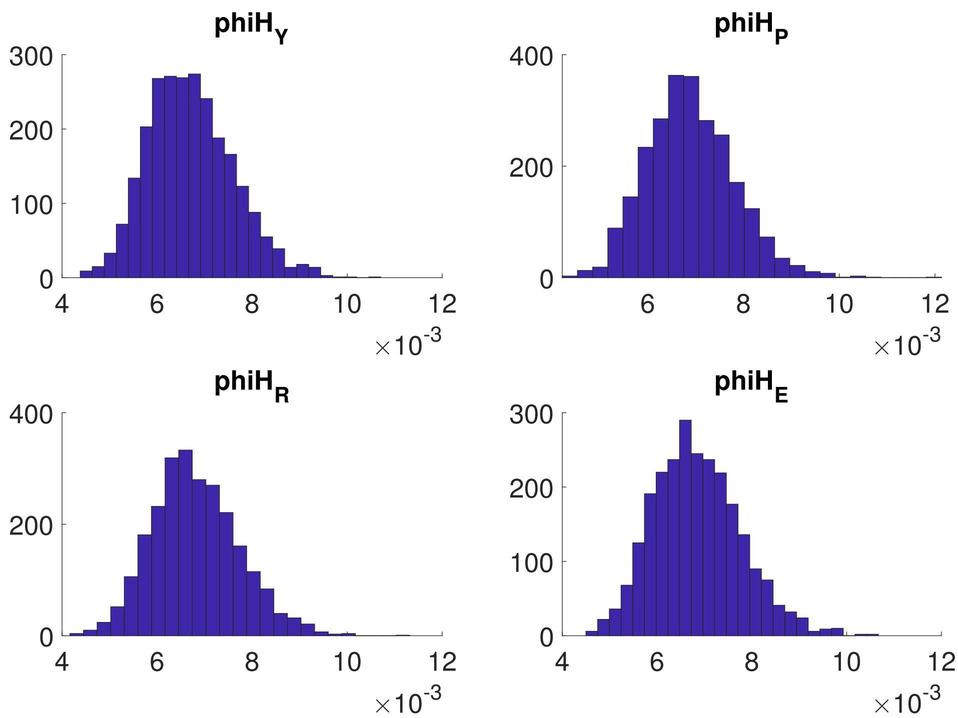


Figure 13: Variance Parameters H

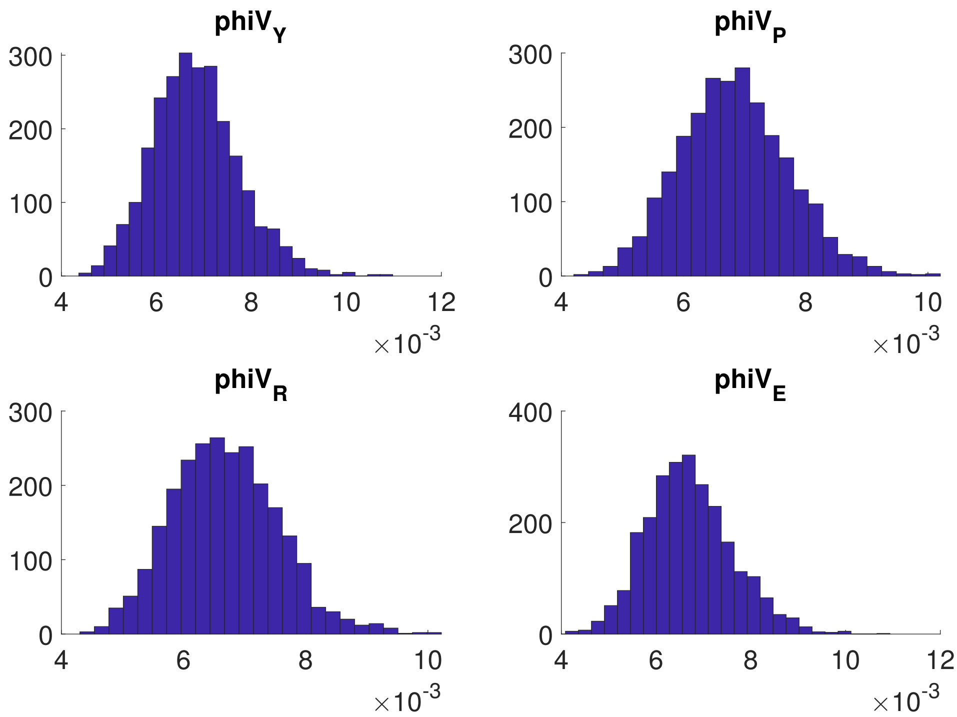


Figure 14: Variance Parameters V

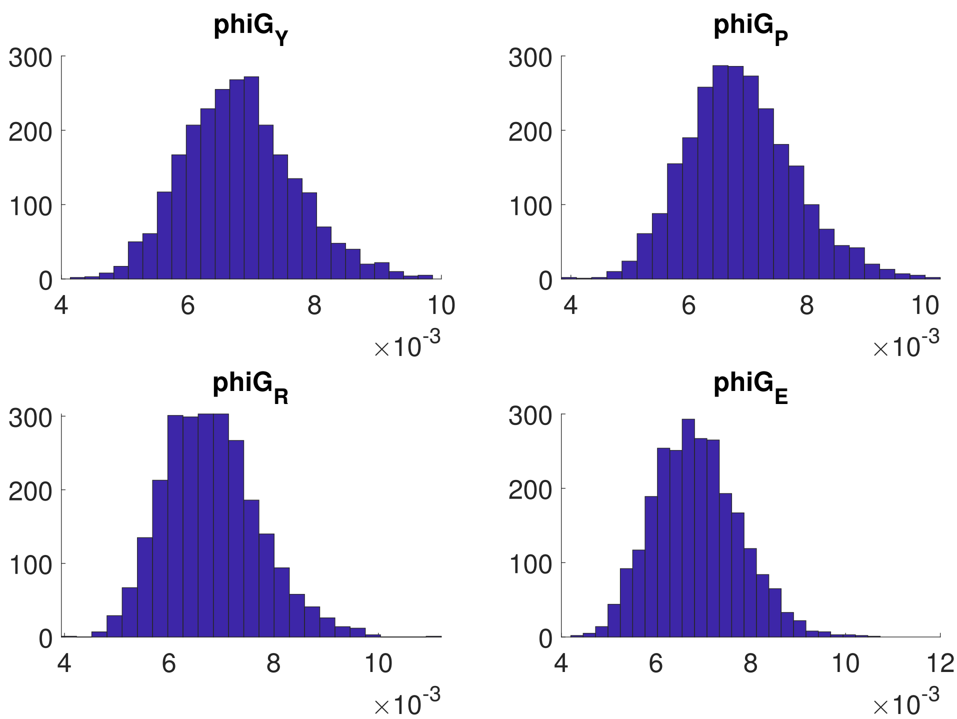


Figure 15: Variance Parameters G

References

AMISANO, G. and GIANNINI, C. (1997). *Topics in Structural VAR Econometrics*. Springer, 2nd edn.

BANBURA, M. and VAN VLODROP, A. (2018). Forecasting with bayesian vector autoregressions with time variation in the mean, tinbergen Institute Discussion Paper, No. TI 2018-025/IV.

CANOVA, F. (2007). *Methods for Applied Macroeconomic Research*. Princeton University Press.

— and GAMBETTI, L. (2009). Structural changes in the US economy: Is there a role for monetary policy? *Journal of Economic Dynamics and Control*, **33**, 477–490.

— and — (2010). Do expectations matter? the great moderation revisited. *American Economic Journal: Macroeconomics*, **2**, 183–205.

— and PÉREZ FORERO, F. J. (2015). Estimating overidentified, nonrecursive, time-varying coefficients structural vector autoregressions. *Quantitative Economics*, **6**, 359–384.

CARRIERO, A., CLARK, T. E. and MARCELLINO, M. (2016). Common drifting volatility in large bayesian vars. *Journal of Business and Economic Statistics*, **34** (3), 375–390.

CARTER, C. K. and KOHN, R. (1994). On gibbs sampling for state space models. *Biometrika*, **81** (3), 541–553.

CLARK, T. (2011). Real-time density forecasts from bayesian vector autoregressions with stochastic volatility. *Journal of Business and Economic Statistics*, **29** (3), 327–341.

COGLEY, T. and SARGENT, T. J. (2005). Drifts and volatilities: Monetary policies and outcomes in the post WWII u.s. *Review of Economic Dynamics*, **8** (2), 262–302.

DEL NEGRO, M. and PRIMICERI, G. (2015). Time varying structural vector autoregressions and monetary policy: A corrigendum. *Review of Economic Studies*, **82**, 1342–1345.

DEPARTAMENTO DE MODELOS MACROECONÓMICOS (2009). Modelo de proyección trimestral del bcrp, working Papers 2009-006, Banco Central de Reserva del Perú.

DIEBOLD, F. X., SCHORFHEIDE, F. and SHIN, M. (2017). Real-time forecast evaluation of dsge models with stochastic volatility. *Journal of Econometrics*, **201** (2), 322–332.

DOAN, T., LITTERMAN, R. and SIMS, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, **3** (1), 1–100.

DURBIN, J. and KOOPMAN, S. (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika*, **89** (3), 603–615.

GIANNONE, D., LENZA, M. and PRIMICERI, G. E. (2015). Prior selection for bayesian vector autoregressions. *The Review of Economics and Statistics*, **97** (2), 436–451.

JACQUIER, E., POLSON, N. G. and ROSSI, P. E. (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, **20** (1), 69–87.

KIM, S., SHEPHARD, N. and CHIB, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies*, **65** (3), 361–393.

KOOP, G. and KOROBILIS, D. (2010). Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends in Econometrics*, **3** (4), 267–358.

LITTERMAN, R. (1986). Forecasting with bayesian vector autoregressions: Five years of experience. *Journal of Business and Economic Statistics*, **4**, 25–38.

LLOSA, G., TUESTA, V. and VEGA, M. (2005). Un modelo de proyección bvar para la inflación peruana, working Papers 2005-007, Banco Central de Reserva del Perú.

PÉREZ FORERO, F. J. (2018). Nowcasting peruvian gdp using leading indicators and bayesian variable selection, working Papers 2018-010, Banco Central de Reserva del Perú.

PRIMICERI, G. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, **72**, 821–852.

WAGGONER, D. F. and ZHA, T. (1999). Conditional forecasts in dynamic multivariate models. *The review of economics and statistics*, **81** (4), 639–651.

WINKELRIED, D. (2013). Modelo de proyección trimestral del bcrp: Actualización y novedades. *Revista Estudios Econiómicos*, **26**, 9–60.