Leading Indicators, Bayesian Variable Selection and Nowcasting of Peruvian GDP

XXXVI Encuentro de Economistas del BCRP

Fernando Pérez Forero
fernando.perez@bcrp.gob.pe

Banco Central de Reserva del Perú

The views expressed are those of the author and do not necessarily reflect those of the Central Bank of Peru.

October 31st, 2018

Table of Contents

Motivation

- Information is a valuable item for decision makers.
- In particular, policy makers and private economic agents such as investors need all the time new information related with the aggregate economy in order to take proper decisions for the future.
- A well known issue is the fact that GDP growth data is only available with a lag. The length of this lag is variable across countries, but it is usually the case that the new GDP data is released with more than one month of lag.
- Therefore, given that current GDP is non-observable, economic agents need to perform an exercise of forecasting the present, i.e. *nowcasting*.

Leading Indicators

- There exists a large set of leading indicators and economic variables that
  - are related with the GDP
  - are released almost in real time, i.e. in advance of the GDP.
- As a result, we as econometricians can think on a linear regression model of the GDP onto a subset of these leading indicators in order to forecast the current value of GDP. In particular, previous nowcasting exercises for GDP growth with Peruvian data can be found in Pérez-Forero (2017).
- Leading indicators in Peru: Escobal and Torres (2002), Ochoa and Lladó (2003), Kapsoli and Bencich (2004), among others.

Nowcasting

- *Nowcasting* macroeconomic time series is not straightforward, since tons of potential regressors and model specifications can be spotted by different experts and professional forecasters at any point in time.
- Which is the best linear regression model?
- How can we select the best regressors among a very large set of variables?
- Can we use different models? It is likely that more than one model is popular at any point in time, given the heterogeneity of views across the different experts.
- Most of these experts can claim that they have **'the model'** (non-nested).
- How we can average non-nested models? (this paper).

This paper

- We specify a Structural Time Series model estimated with Bayesian techniques (Scott and Varian, 2015).
- The latter model has been used for *nowcasting* time series using a large set of variables, i.e. Google Trends data.
- We implement the spike-and-slab approach to model selection developed by George and McCulloch (1997) and Madigan and Raftery (1994).
- We use this machinery for finding the best predictors of Peruvian GDP growth.
- Previous *nowcasting* exercises with Peruvian data can be found in Pérez-Forero (2017), among others.
- *Nowcasting* literature: Evans (2005), Giannone (2008), Banbura (2013) among others. Small Open Economies: Brazil (Bragoli , 2015), India (Bragoli and Fosten, 2016) and Japan (Bragoli, 2017).

Table of Contents

A Structural time series model

Scott and Varian (2015): This model is a stochastic generalization of the classic constant-trend regression model:

$$y_t = \mu_t + \gamma_{t,1} + z_t + v_t, \qquad v_t \sim N(0, V)$$

$$\mu_t = \mu_{t-1} + b_{t-1} + w_{1,t}, \qquad w_{1,t} \sim N(0, W_1)$$

$$b_t = b_{t-1} + w_{2,t}, \qquad w_{2,t} \sim N(0, W_2)$$

$$\gamma_{t,1} = -\sum_{i=1}^{S-1} \gamma_{t-1,i} + w_{3,t}, \qquad w_{3,t} \sim N(0, W_3)$$

$$\gamma_{t,i} = \gamma_{t-1,i}, \quad i = 1, \ldots, S-1$$

$$z_t = \sum_{i=1}^{K} \beta_i x_{i,t}$$

Spike and slab variable selection I

- Let γ denote a vector the same length as β that indicates whether or not a particular regressor is included in the regression, where $\gamma_i = 1$ implies that $\beta_i \neq 0$ and $\gamma_i = 0$ indicates $\beta_i = 0$. Let β_γ indicate the subset of for which $\gamma_i = 1$, and let σ^2 be the residual variance from the regression model.
- A spike and slab prior for the joint distribution of $\left(\beta, \gamma, \sigma^{-2}\right)$ can be factored in the usual way:

$$p\left(\beta, \gamma, \sigma^{-2}\right) = p\left(\beta_\gamma \mid \gamma, \sigma^{-2}\right) p\left(\sigma^{-2} \mid \gamma\right) p\left(\gamma\right)$$

- The "spike" part of a spike-and-slab prior refers to the point mass at zero, for which we assume a Bernoulli distribution for each i, so that the prior is a product of Bernoullis:

$$\gamma \sim \prod_i \pi_i^{\gamma_i} \left(1 - \pi_i\right)^{1-\gamma_i}$$

Spike and slab variable selection II

- It is convenient to set all π_i equal to the same number, π. If k out of K coefficients are expected to be non-zero then set $\pi = k/K$ in the prior.
- The "slab" component: Let b be a vector of prior beliefs for β, let Ω^{-1} be a prior precision matrix, and let Ω_γ^{-1} denote rows and columns of Ω^{-1} for which $\gamma_i = 1$. A conditionally conjugate "slab" prior is

$$\beta_\gamma \mid \gamma, \sigma^{-2} \sim N\left(b_\gamma, \sigma^2 \left(\Omega_\gamma^{-1}\right)^{-1}\right)$$

$$\frac{1}{\sigma^2} \sim \Gamma\left(\frac{df}{2}, \frac{ss}{2}\right)$$

- Because $X'X/\sigma^2$ is the total Fisher information in the full data, it is reasonable to parametrize $\Omega^{-1} = \kappa X'X/T$. However, since $X'X$ is potentially rank deficient, we assume that

$$\Omega^{-1} = \frac{\kappa}{T}\left(wX'X + (1-w)\, diag\left(X'X\right)\right)$$

Bayesian model averaging (Scott and Varian, 2015)

- Bayesian inference with spike-and-slab priors is an effective way to implement Bayesian model averaging over the space of time series regression models. We will end up drawing from the posterior distribution of the parameters in the model.
- Each draw of parameters from the posterior can be combined with the available data to yield a forecast of $E\left(y_{t+h} \mid y_t\right)$ for that particular draw. Repeating these draws many times gives us an estimate of the posterior distribution of the forecast $E\left(y_{t+h} \mid y_t\right)$.
- This approach is motivated by the Madigan and Raftery (1994) proof that averaging over an ensemble of models does no worse than using the best single model in the ensemble.

Table of Contents

Bayesian Estimation

The model can be re-written as a state-space system with an exogenous component and time varying matrices (Kim and Nelson, 1999), so that:

$$y_t = D_t \alpha_t + Z_t X_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, H_t)$$

$$\alpha_t = A_t \alpha_{t-1} + R_t \eta_t, \qquad \eta_t \sim N(0, Q_t)$$

Bayesian Estimation

Denote $\psi = \left(\theta, \alpha^T\right)$ as the parameter set of the model, then the complete posterior distribution is:

$$p\left(\psi \mid y^T\right) = p\left(\theta, \alpha^T \mid y^T\right) \propto p\left(\theta\right) p\left(\alpha_0\right) \prod_{t=1}^{T} p\left(y_t \mid \alpha_t, \theta\right) p\left(\alpha_t \mid \alpha_{t-1}, \theta\right)$$

Analytical computation of the posterior distribution is possible conditional on γ.

Gibbs Sampling

1. Simulate $\{\alpha_t\}_{t=1}^{T}$ from $p\left(\alpha_t \mid y^T, \psi_{-\alpha_t}\right)$: Carter and Kohn (1994)

$$\alpha_t \mid y^T, \psi_{-\alpha_t} \sim N\left(\overline{\alpha}_{t|T}, \overline{P}_{t|T}\right),\ t \leq T \tag{1}$$

2. Simulate V from $p\left(V \mid Y^T, \psi_{-V}\right)$: Inverse-Gamma
3. Simulate W_1 from $p\left(W_1 \mid y^T, \psi_{-W_1}\right)$: Inverse-Gamma
4. Simulate W_2 from $p\left(W_2 \mid y^T, \psi_{-W_2}\right)$: Inverse-Gamma
5. Simulate W_3 from $p\left(W_3 \mid y^T, \psi_{-W_3}\right)$: Inverse-Gamma
6. Simulate β from $p\left(\beta \mid y^T, \psi_{-\beta}\right)$: Normal
7. Simulate σ^2 from $p\left(\sigma^2 \mid y^T, \psi_{-\sigma^2}\right)$: Inverse-Gamma
8. Simulate γ from $p\left(\gamma \mid y^T, \psi_{-\gamma}\right)$: Metropolis step as in George and McCulloch (1997)

Metropolis-Hastings

In order to sample γ from $p\left(\gamma \mid y^T, \psi_{-\gamma}\right)$ as in George and McCulloch (1997), we implement a Metropolis-Hastings step. The algorithm is as follows:

1. Generate a candidate value γ^* with probability distribution $q\left(\gamma^{(j)}, \gamma^*\right)$.
2. Set $\gamma^{(j+1)} = \gamma^*$ with probability:

$$\alpha^{MH}\left(\gamma^{(j)}, \gamma^*\right) = \min\left\{\frac{q\left(\gamma^*, \gamma^{(j)}\right)}{q\left(\gamma^{(j)}, \gamma^*\right)} \frac{g\left(\gamma^*\right)}{g\left(\gamma^{(j)}\right)}, 1\right\}$$

   Otherwise $\gamma^{(j+1)} = \gamma^{(j)}$.

In particular, $q\left(\gamma^{(j)}, \gamma^*\right)$ is such that γ^* is generated by randomly changing one component of $\gamma^{(j)}$. As a consequence, $q\left(.\right)$ is a symmetric proposal.

Priors

Parameter	Distribution	Hyper-parameters
α_0	Normal	$N\left(0_{\dim\alpha\times 1}, I_{\dim\alpha}\right)$
V	Inverse-Gamma	$IG\left(\frac{1}{2}, \frac{0.001}{2}\right)$
$W_{i=1,2,3}$	Inverse-Gamma	$IG\left(\frac{1}{2}, \frac{0.001}{2}\right)$
β_γ	Normal	$N\left(0_{\dim\beta_\gamma\times 1}, \sigma^2\left(\Omega_\gamma^{-1}\right)^{-1}\right)$ [1]
σ^2	Inverse-Gamma	$IG\left(\frac{1}{2}, \frac{0.001}{2}\right)$
γ	Spike-slab	$\pi_i = 5/K$

Table: Priors for state-space parameters

[1] where $\kappa = 0.25, w = 0.9925$ in $\Omega^{-1} = \frac{\kappa}{T}\left(wX'X + (1-w)\, diag\left(X'X\right)\right)$

Estimation Setup

- We run the Gibbs sampler for $K = 1,000,000$ and discard the first $500,000$ 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 100. As a result, we have $5,000$ draws for conducting inference.
- The acceptance rate of the metropolis-step associated with γ is around 0.50.

Table of Contents

Data Description

- Monthly data from January 2003 until June 2018.
- More than 80 regressors related with economic activity indicators, interest rates, money aggregates, stock markets, price indexes and also external variables.
- Given the specified model, we take year-to-year growth rates when it is convenient, except for interest rates and some particular indexes.

	Variable
1	Emisión primaria
2	Circulante
3	Tipo de cambio nominal
4	PBI total
5	Producción de electricidad (COES)
6	Consumo Interno de Cemento
7	IGV Interno
8	Ventas de Pollos
9	Empleo mensual en Lima Metropolitana (miles de personas) - PEA Ocupada
10	Empleo mensual en Lima Metropolitana (miles de personas) - Ingreso Mensual
11	Empleo mensual en Lima Metropolitana (porcentaje) - Tasa de Desempleo (%)
12	GNF - Gobierno General
13	FBK - Gobierno General
14	Volumen de Importaciones de Insumos Industriales
15	Bolsa de Valores de Lima - Índices Bursátiles - Índice General BVL (base 31/12/91 = 100)
16	Bolsa de Valores de Lima - Índices Bursátiles - Índice Selectivo BVL (base 31/12/91 = 100)
17	Índice de Precios al Consumidor
18	Índice de precios al consumidor sin Alimentos ni Enenrgía
19	Índice de Precios al por Mayor
20	Términos de Intercambio
21	Tasa LIBOR a 3 meses
22	EMBI Perú

Figure: Data Description (1)

	Variable
23	WTI
24	Tipo de Cambio Real Multilateral (2009 = 100)
25	Tipo de Cambio Real Bilateral (2009 = 100)
26	IPC de EEUU
27	Tasa Activa MN Promedio
28	Tasa Activa MN Flujo
29	Tasa preferencial corporativa a 90 días
30	Tasa Pasiva MN, depósitos a la vista
31	Tasa Pasiva MN, depósitos de ahorro
32	Tasa Pasiva MN, plazo a 30 días
33	Tasa Pasiva MN, plazo hasta 180 días
34	Tasa Pasiva MN, plazo hasta 360 días
35	Tasa Pasiva MN, plazo más de 360 días
36	Tasa Pasiva MN Promedio
37	Tasa Pasiva MN Flujo
38	Tasa MN Interbancaria
39	Industrial Production Index, Index 2012=100, Monthly, Seasonally Adjusted
40	Producer Price Index for All Commodities, Index 1982=100, Monthly, Not Seasonally Adjusted
41	CBOE Volatility Index: VIX©, Index, Monthly, Not Seasonally Adjusted
42	Índice de expectativas de la economía a 3 meses
43	Índice de expectativas del sector a 3 meses

Figure: Data Description (2)

Table of Contents

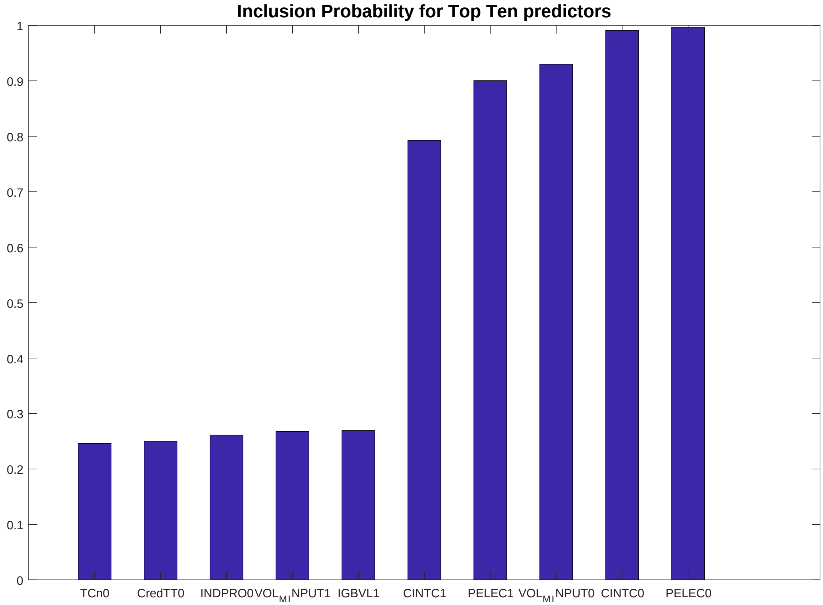

Figure: Top Ten Predictors for GDP

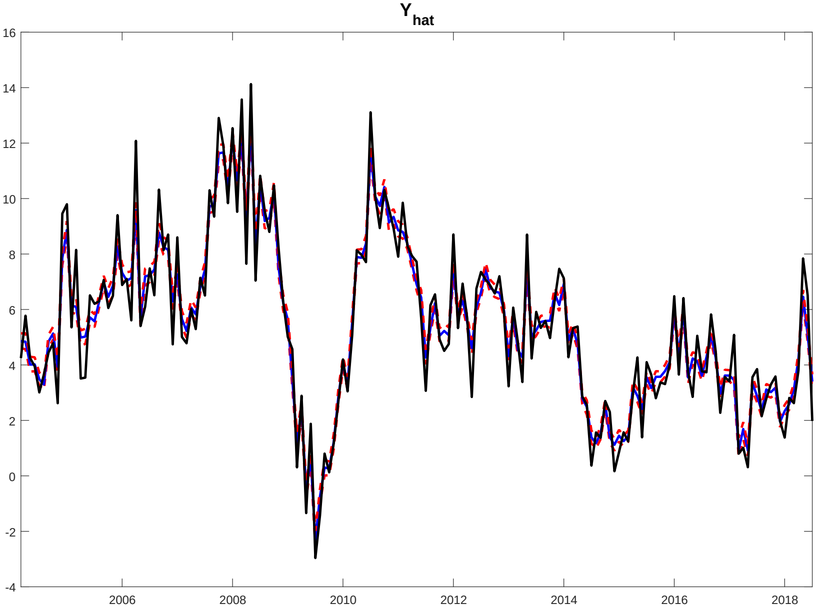

Figure: Predicted GDP growth

Nowcasting GDP

Given the posterior estimates of the parameter set $\psi = \left(\theta, \alpha^T, \gamma\right)$, then for each draw $i = 1, \ldots, S$ of ψ we can forecast the latent variable such that:

$$\alpha_{T+h|T}^{(i)} = \left[A_T^{(i)}\right]^h \alpha_{T|T}^{(i)} + \eta_{T+h}$$

where $\eta_{T+h} \sim N\left(0, Q_T\right)$. The latter, together with the data available of exogenous regressors up to a horizon h, is useful in order to forecast the dependent variable using the measurement equation:

$$y_{T+h|T}^{(i)} = D_{T+h}^{(i)} \alpha_{T+h|T}^{(i)} + \left(\beta \mid \gamma\right)^{(i)} x_{T+h} + v_{T+h}$$

where $v_{T+h} \sim N\left(0, V\right)$. That is, as long as we have out of sample data available of the vector x_t, then it is possible to compute the conditional forecast.

Concluding Remarks

- Peruvian GDP short term forecasting and nowcasting is not straightforward.
- We have selected and ranked regressors among a large set of variables using Bayesian techniques.
- Among the main regressors we have detected the following variables: electricity production (PELEC), internal consumption of cement (CINTC) and the volume of imported input goods (VOL_{MINPUT}), all of them in contemporaneous form (t).
- Model averaging using the method suggested by Scott and Varian (2015) allows us to produce density forecasts and quantify the uncertainty associated with the estimation, i.e. the outcome is not only a point forecast of GDP growth. This seems to be very powerful and promising for policymakers interested in producing risk scenarios.

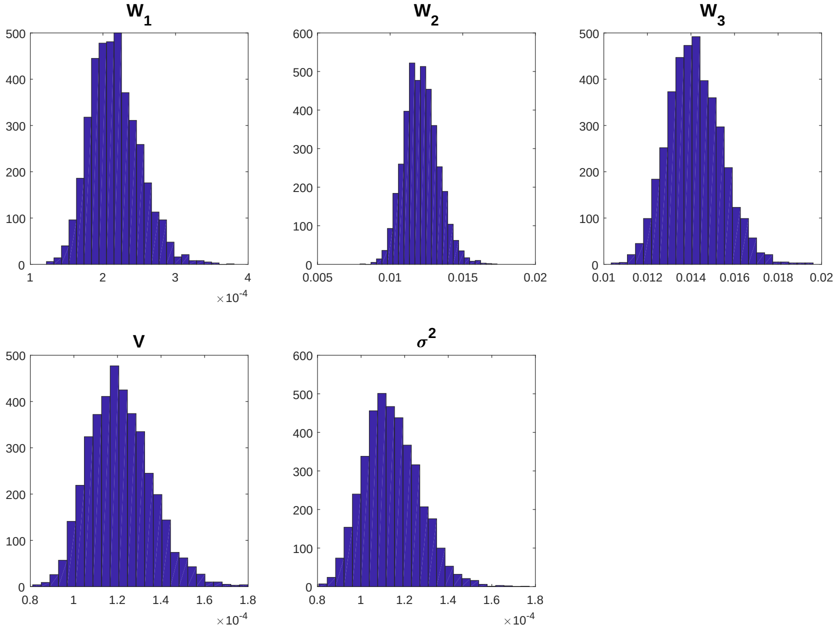

Figure: Posterior Distribution of hyper-parameters

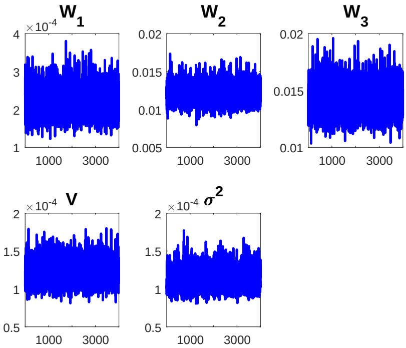

Figure: Posterior Draws of hyper-parameters

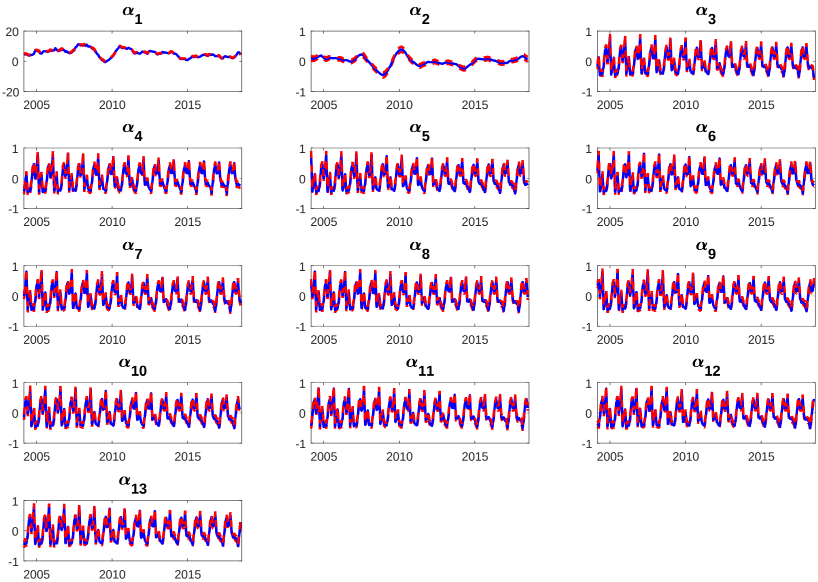

Figure: Posterior Distribution of Latent Factors

References I

BANBURA, M., GIANNONE, D., MODUGNO, M. and REICHLIN, L. (2013). *Now-Casting and the Real-Time Data Flow*, Elsevier-North Holland, chap. 2, pp. 2–195.

BRAGOLI, D. (2017). Now-casting the japanese economy. *International Journal of Forecasting*, **33** (2), 390–402.

— and FOSTEN, J. (2016). Nowcasting indian gdp, university of East Anglia School of Economics Working Paper Series 2016-06, School of Economics, University of East Anglia, Norwich, UK. June 2016.

—, METELLI, L. and MODUGNO, M. (2015). The importance of updating: Evidence from a brazilian nowcasting model. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, **2015** (1), 5–22.

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

ESCOBAL, J. and TORRES, J. (2002). Un sistema de indicadores líderes del nivel de actividad para la economía peruana. *Lima*.

References II

EVANS, C. (2005). Where are we now? real-time estimates of the macroeconomy. *International Journal of Central Banking*, **1** (2), 127–175.

GEORGE, E. I. and MCCULLOCH, R. E. (1997). Approaches for bayesian variable selection. *Statistica Sinica*, **49** (428), 339–373.

GIANNONE, D., REICHLIN, L. and SMALL, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, **55**, 665–676.

KAPSOLI, J. and BENCICH, B. (2004). Indicadores líderes, redes neuronales y predicción de corto plazo. *Pontificia Universidad Católica del Perú, Revista Economía*, **27**.

KIM, C.-J. and NELSON, C. R. (1999). *State-Space Models with Regime-Switching: Classical and Gibbs-Sampling Approaches with Applications*. MIT Press.

References III

Madigan, D. and Raftery, A. (1994). Model selection and accounting for model uncertainty in graphical models using occam's window. *Journal of American Statistical Association*, **82**, 1535–1546.

Ochoa, E. and Lladó, J. (2003). Modelos de indicadores líderes de actividad económica para el perú. *Revista Estudios Económicos*, **10**.

Pérez-Forero, F., Ghurra, O. and Grandez, R. (2017). Un indicador líder de actividad real para el perú, bCRP Documento de Trabajo Documento de Trabajo 2017-001.

Scott, S. and Varian, H. (2015). Bayesian variable selection for nowcasting economic time series. In A. Goldfarb, S. M. Greenstein and C. E. Tucker (eds.), *Economic Analysis of the Digital Economy*, University of Chicago Press by the National Bureau of Economic Research, pp. 119–135.

Thanks