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

1. Motivation

2. A Structural time series model

3. Bayesian Estimation

4. Data Description

5. Nowcasting Peruvian GDP

6. Concluding Remarks
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).

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.
# Table of Contents

1. Motivation  
2. A Structural time series model  
3. Bayesian Estimation  
4. Data Description  
5. Nowcasting Peruvian GDP  
6. Concluding Remarks
A Structural time series model

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

\[ y_t = \mu_t + \gamma_{t,1} + z_t + v_t, \quad v_t \sim N(0, V) \]

\[ \mu_t = \mu_{t-1} + b_{t-1} + w_{1,t}, \quad w_{1,t} \sim N(0, W_1) \]

\[ b_t = b_{t-1} + w_{2,t}, \quad w_{2,t} \sim N(0, W_2) \]

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

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

\[ z_t = \sum_{i=1}^{K} \beta_i x_{i,t} \]
Spike and slab variable selection I

- Let $\gamma$ denote a vector the same length as $\beta$ 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 $\beta_\gamma$ indicate the subset of for which $\gamma_i = 1$, and let $\sigma^2$ be the residual variance from the regression model.

- A spike and slab prior for the joint distribution of $(\beta, \gamma, \sigma^{-2})$ can be factored in the usual way:

\[
p(\beta, \gamma, \sigma^{-2}) = p(\beta_\gamma | \gamma, \sigma^{-2}) p(\sigma^{-2} | \gamma) p(\gamma)
\]

- 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:

\[
\gamma \sim \prod_i \pi_i^{\gamma_i} (1 - \pi_i)^{1-\gamma_i}
\]
Spike and slab variable selection II

- It is convenient to set all $\pi_i$ equal to the same number, $\pi$. 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 $\beta$, let $\Omega^{-1}$ be a prior precision matrix, and let $\Omega^{-1}_\gamma$ denote rows and columns of $\Omega^{-1}$ for which $\gamma_i = 1$. A conditionally conjugate "slab" prior is

$$\beta_\gamma | \gamma, \sigma^{-2} \sim N \left( b_\gamma, \sigma^2 \left( \Omega^{-1}_\gamma \right)^{-1} \right)$$

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

- 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

$$\Omega^{-1} = \frac{\kappa}{T} \left( wX'X + (1 - w) \text{diag} \left( X'X \right) \right)$$
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 \( E(y_{t+h} \mid y_t) \) for that particular draw. Repeating these draws many times gives us an estimate of the posterior distribution of the forecast \( E(y_{t+h} \mid y_t) \).

- 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

1. Motivation

2. A Structural time series model

3. Bayesian Estimation

4. Data Description

5. Nowcasting Peruvian GDP

6. Concluding Remarks
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:

\[ y_t = D_t \alpha_t + Z_t X_t + \epsilon_t, \quad \epsilon_t \sim N(0, H_t) \]

\[ \alpha_t = A_t \alpha_{t-1} + R_t \eta_t, \quad \eta_t \sim N(0, Q_t) \]
Bayesian Estimation

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

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

Analytical computation of the posterior distribution is possible conditional on $\gamma$. 
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)

\[ \alpha_t \mid y^T, \psi_{-\alpha_t} \sim N\left(\bar{\alpha}_T, \bar{P}_T\right), \; t \leq T \]  

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 $\beta$ from $p\left(\beta \mid y^T, \psi_{-\beta}\right)$: Normal

7. Simulate $\sigma^2$ from $p\left(\sigma^2 \mid y^T, \psi_{-\sigma^2}\right)$: Inverse-Gamma

8. Simulate $\gamma$ from $p\left(\gamma \mid y^T, \psi_{-\gamma}\right)$: Metropolis step as in George and McCulloch (1997)
Metropolis-Hastings

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

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

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

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

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

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distribution</th>
<th>Hyper-parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>Normal</td>
<td>$N \left( 0_{\text{dim} \alpha \times 1}, I_{\text{dim} \alpha} \right)$</td>
</tr>
<tr>
<td>$V$</td>
<td>Inverse-Gamma</td>
<td>$IG \left( \frac{1}{2}, \frac{0.001}{2} \right)$</td>
</tr>
<tr>
<td>$W_{i=1,2,3}$</td>
<td>Inverse-Gamma</td>
<td>$IG \left( \frac{1}{2}, \frac{0.001}{2} \right)$</td>
</tr>
<tr>
<td>$\beta_\gamma$</td>
<td>Normal</td>
<td>$N \left( 0_{\text{dim} \beta_\gamma \times 1}, \sigma^2 \left( \Omega_\gamma^{-1} \right)^{-1} \right)$</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>Inverse-Gamma</td>
<td>$IG \left( \frac{1}{2}, \frac{0.001}{2} \right)$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Spike-slab</td>
<td>$\pi_i = \frac{5}{K}$</td>
</tr>
</tbody>
</table>

Table: Priors for state-space parameters

---

$\kappa = 0.25$, $w = 0.9925$ in $\Omega^{-1} = \frac{\kappa}{T} (wX'X + (1-w) \text{diag}(X'X))$
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 $\gamma$ is around 0.50.
Table of Contents

1 Motivation

2 A Structural time series model

3 Bayesian Estimation

4 Data Description

5 Nowcasting Peruvian GDP

6 Concluding Remarks
Data Description

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

**Figure:** Data Description (1)
<table>
<thead>
<tr>
<th></th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>WTI</td>
</tr>
<tr>
<td>24</td>
<td>Tipo de Cambio Real Multilateral (2009 = 100)</td>
</tr>
<tr>
<td>25</td>
<td>Tipo de Cambio Real Bilateral (2009 = 100)</td>
</tr>
<tr>
<td>26</td>
<td>IPC de EEUU</td>
</tr>
<tr>
<td>27</td>
<td>Tasa Activa MN Promedio</td>
</tr>
<tr>
<td>28</td>
<td>Tasa Activa MN Flujo</td>
</tr>
<tr>
<td>29</td>
<td>Tasa preferencial corporativa a 90 días</td>
</tr>
<tr>
<td>30</td>
<td>Tasa Pasiva MN, depósitos a la vista</td>
</tr>
<tr>
<td>31</td>
<td>Tasa Pasiva MN, depósitos de ahorro</td>
</tr>
<tr>
<td>32</td>
<td>Tasa Pasiva MN, plazo a 30 días</td>
</tr>
<tr>
<td>33</td>
<td>Tasa Pasiva MN, plazo hasta 180 días</td>
</tr>
<tr>
<td>34</td>
<td>Tasa Pasiva MN, plazo hasta 360 días</td>
</tr>
<tr>
<td>35</td>
<td>Tasa Pasiva MN, plazo más de 360 días</td>
</tr>
<tr>
<td>36</td>
<td>Tasa Pasiva MN Promedio</td>
</tr>
<tr>
<td>37</td>
<td>Tasa Pasiva MN Flujo</td>
</tr>
<tr>
<td>38</td>
<td>Tasa MN Interbancaria</td>
</tr>
<tr>
<td>39</td>
<td>Industrial Production Index, Index 2012=100, Monthly, Seasonally Adjusted</td>
</tr>
<tr>
<td>40</td>
<td>Producer Price Index for All Commodities, Index 1982=100, Monthly, Not Seasonally Adjusted</td>
</tr>
<tr>
<td>41</td>
<td>CBOE Volatility Index: VIX®, Index, Monthly, Not Seasonally Adjusted</td>
</tr>
<tr>
<td>42</td>
<td>Índice de expectativas de la economía a 3 meses</td>
</tr>
<tr>
<td>43</td>
<td>Índice de expectativas del sector a 3 meses</td>
</tr>
</tbody>
</table>

**Figure:** Data Description (2)
Table of Contents

1 Motivation

2 A Structural time series model

3 Bayesian Estimation

4 Data Description

5 Nowcasting Peruvian GDP

6 Concluding Remarks
Inclusion Probability for Top Ten predictors

Figure: Top Ten Predictors for GDP
Figure: Predicted GDP growth
Nowcasting GDP

Given the posterior estimates of the parameter set $\psi = (\theta, \alpha^T, \gamma)$, then for each draw $i = 1, \ldots, S$ of $\psi$ 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 $\eta_{T+h} \sim N(0, Q_T)$. 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:

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

where $v_{T+h} \sim N(0, V)$. 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 ($VOL_{INPUT}$), 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


Thanks